Robert Engelmore Memorial Lecture

Knowledge is Power: A view from the Semantic Web

Jim Hendler

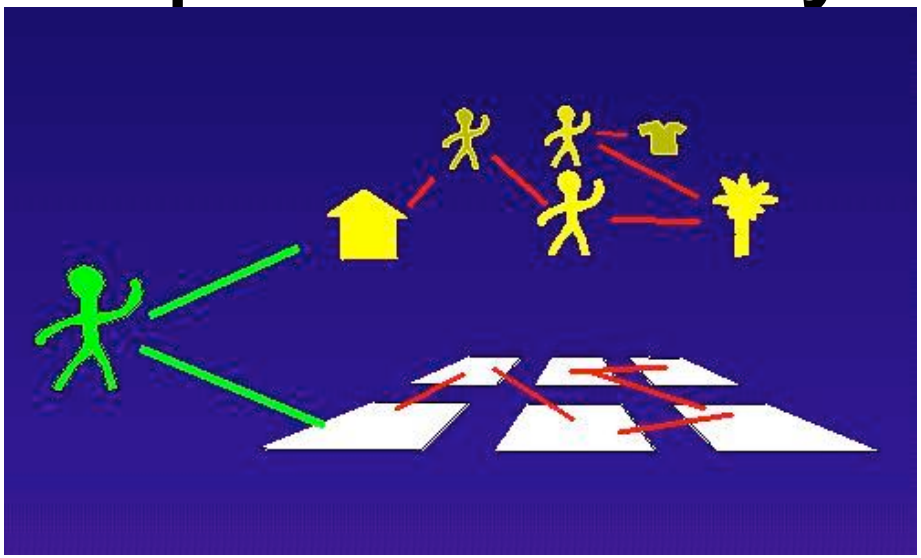
[w/thanks to the *incredible* MINDSWAP research group
visit us on the (Semantic) Web at http://www.mindswap.org]

Knowledge is Power

- But in knowledge resides the power. Because of the importance of knowledge in expert systems and because the current knowledge acquisition method is slow and tedious, much of the future of expert systems depends on breaking the knowledge acquisition bottleneck and in codifying and representing a large knowledge infrastructure.

  -- R. Engelmore, E. Feigenbaum, 1993

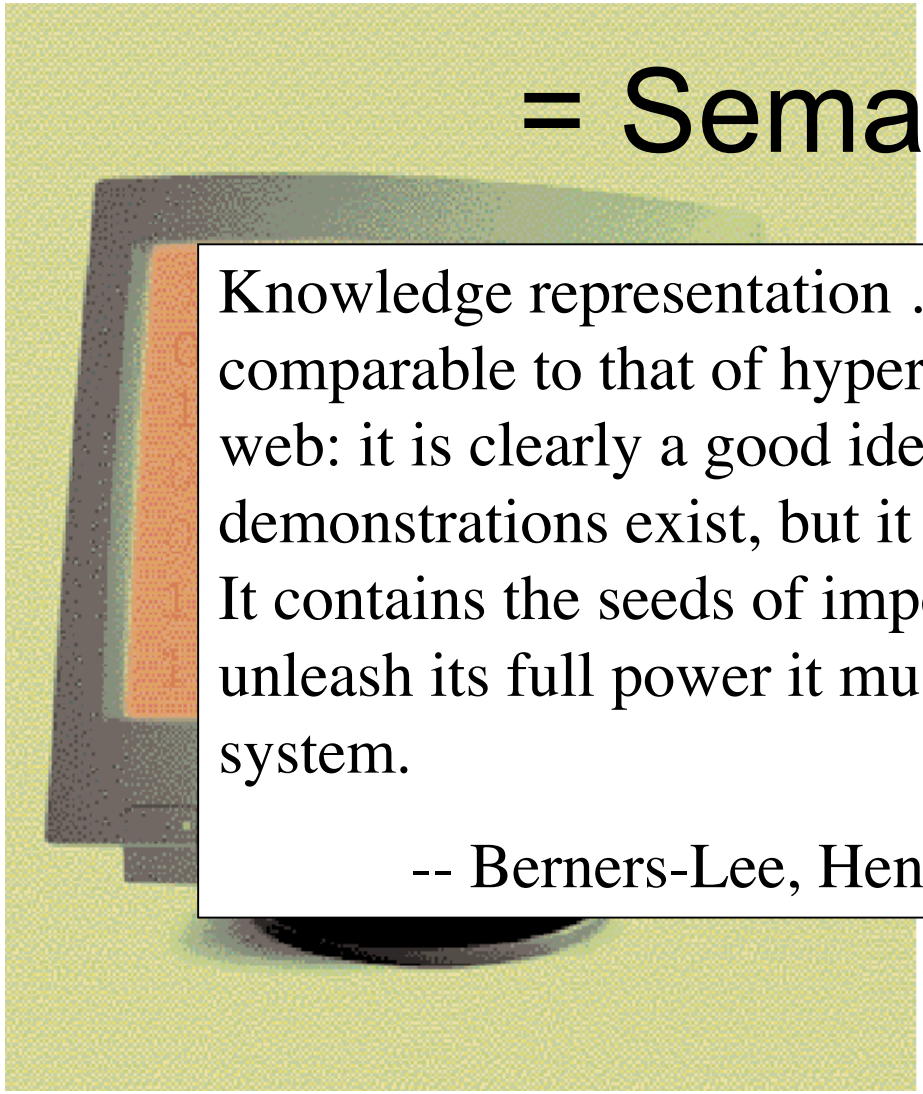
Concepts are Everywhere



"… documents on the web describe real objects and imaginary concepts, and give particular relationships between them... The title document to a house describes a house and also the ownership relation with a person. ... This means that machines, as well as people operating on the web of information, can do real things.

Tim Berners-Lee, 1994

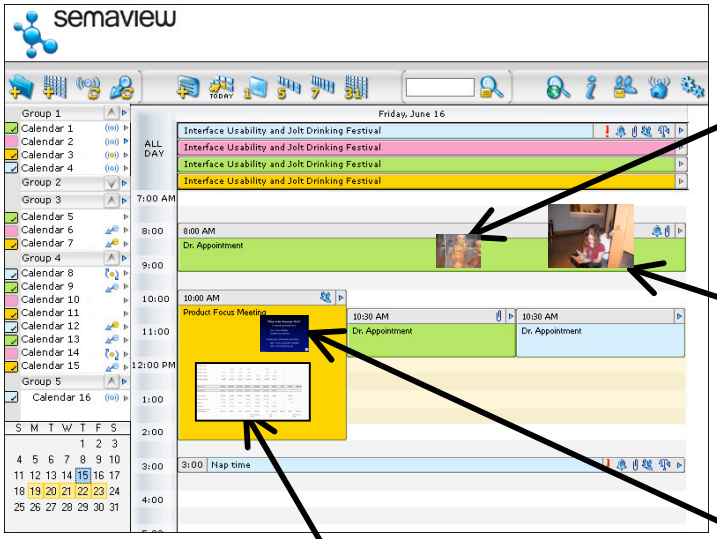
= Semantic Web



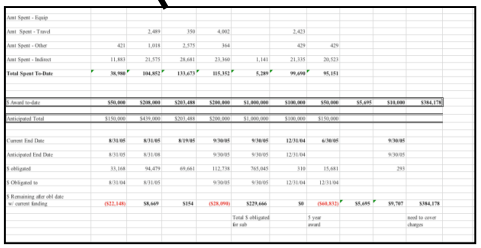
Knowledge representation … is currently in a state comparable to that of hypertext before the advent of the web: it is clearly a good idea, and some very nice demonstrations exist, but it has not yet changed the world. It contains the seeds of important applications, but to unleash its full power it must be linked into a single global system.

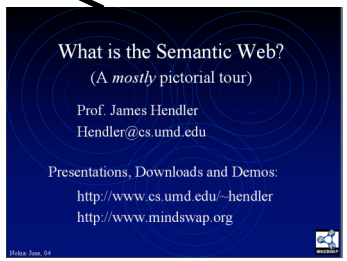
-- Berners-Lee, Hendler, Lassila, 2001.

As we publish more info- how do we control access …

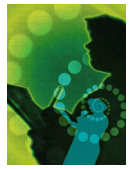


Who can see What??



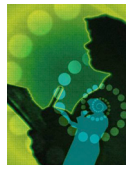


Who can see What??



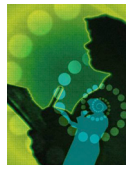
Imagine if

- Every PDF document on the Web contained metadata in a machine readable form as to who created it, when, with what tool, and some of the key document features



Imagine if

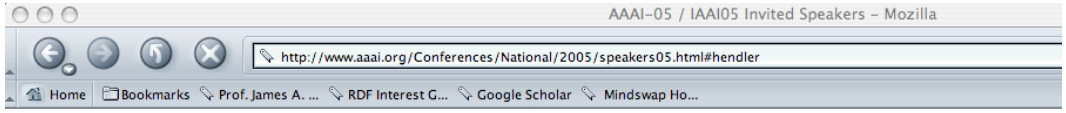
- We could
  - Take a photo from our phone
  - Upload it to the Web
  - Search for a vocabulary relating to the *content* of the photo
  - Easily use that vocabulary to annotate the image
    - Relating it to existing concepts in that space
  - Send it to a Web portal
  - And have it show up indexed under appropriate terms



Imagine if

- There was a distributed knowledge base, accessible through almost every computer on Earth, with information about millions of people, places, things, transactions, processes, services, …. Indexed against thousands of ontologies in a standard KR language with dozens of open-source (and commercial) tools for parsing, serializing, browsing, editing, storing, searching and inferencing these …

Now stop imagining

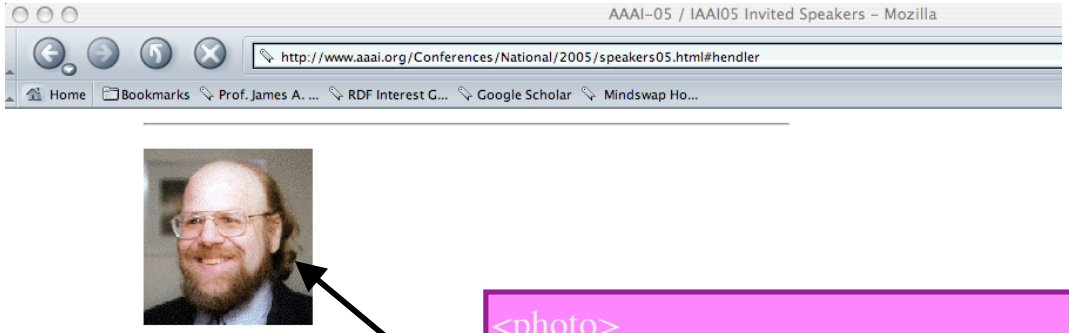


IAAI-05 Robert S. Engelmore Memorial Lecture
Knowledge as Power: A View from the Semantic Web
James Hendler, *University of Maryland*

The emerging semantic web focuses on bringing KR-like capabilities to Web applications in a Web-friendly way. The ability to put knowledge on the Web, share it, and reuse it through standard Web mechanisms provides new and interesting challenges to artificial intelligence. In this talk, Hendler explores the similarities and differences between the semantic web and traditional AI knowledge representation systems, and sees if he can validate the analogy "the semantic web is to KR as the Web is to hypertext."

Jim Hendler is a professor and director of the Joint Institute for Knowledge Discovery at the University of Maryland. One of the inventors of the semantic web, Hendler was the recipient of a 1995 Fulbright Foundation Fellowship, is a former member of the US Air Force Science Advisory Board, and is a Fellow of the American Association for Artificial Intelligence. He is also the former chief scientist of the Information Systems Office at the US Defense Advanced Research Projects Agency (DARPA), was awarded a US Air Force Exceptional Civilian Service Medal in 2002, and is a member of

The Web



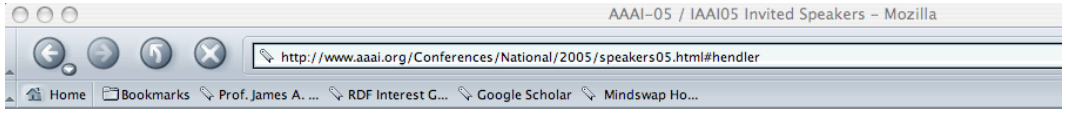
IAAI-05 Robert S. Engelmore M
Knowledge as Power: A View
James Hendler, *University of M*

The emerging semantic web focuses
to Web applications in a Web-fr
knowledge on the Web, share it, an
mechanisms provides new and interesting challenges to artificial intelligence. In this talk, Hendler explores the similarities and differences between the semantic web and traditional AI knowledge representation systems, and sees if he can validate the analogy "the semantic web is to KR as the Web is to hypertext."

Jim Hendler is a professor and director of the Joint Institute for Knowledge Discovery at the University of Maryland. One of the inventors of the semantic web, Hendler was the recipient of a 1995 Fulbright Foundation Fellowship, is a former member of the US Air Force Science Advisory Board, and is a Fellow of the American Association for Artificial Intelligence. He is also the former chief scientist of the Information Systems Office at the US Defense Advanced Research Projects Agency (DARPA), was awarded a US Air Force Exceptional Civilian Service Medal in 2002, and is a member of

```
<photo>
  <subject> http://www.cs.umd.edu/~hendler </subject>
  <name> jim Hendler</name>  </name>
  …
</photo>
```

+ Micro-language
(or XML schema)



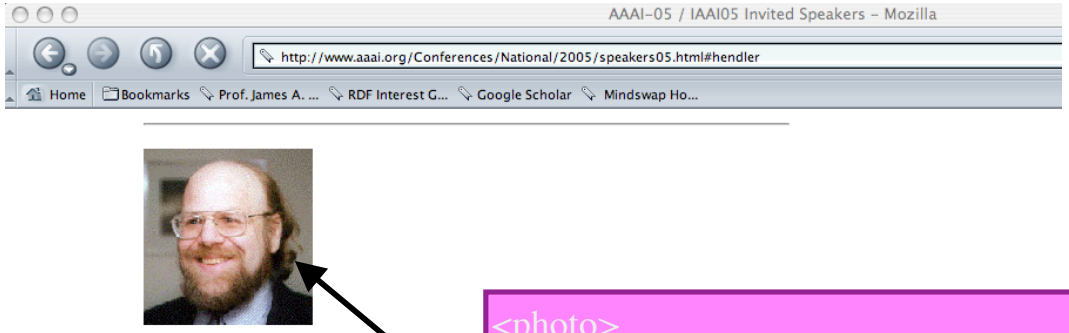
IAAI-05 Robert S. Engelmore M
Knowledge as Power: A View
James Hendler, *University of M*

The emerging semantic web focuses
to Web applications in a Web-fr
knowledge on the Web, share it, an
mechanisms provides new and interesting challenges to artificial intelligence. In this talk, Hendler explores the similarities and differences between the semantic web and traditional AI knowledge representation systems, and sees if he can validate the analogy "the semantic web is to KR as the Web is to hypertext."

Jim Hendler is a professor and director of the Joint Institute for Knowledge Discovery at the University of Maryland. One of the inventors of the semantic web, Hendler was the recipient of a 1995 Fulbright Foundation Fellowship, is a former member of the US Air Force Science Advisory Board, and is a Fellow of the American Association for Artificial Intelligence. He is also the former chief scientist of the Information Systems Office at the US Defense Advanced Research Projects Agency (DARPA), was awarded a US Air Force Exceptional Civilian Service Medal in 2002, and is a member of

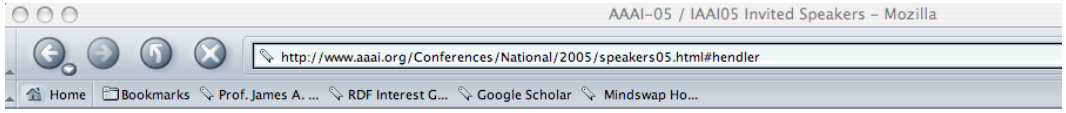
```
<photo>
  <subject> http://www.cs.umd.edu/~hendler </subject>
  <name> jim Hendler</name>  </name>
  …
</photo>
```

Adding Semantics -- NOT!




```
<photo>
  <subject> http://www.cs.umd.edu/~hendler </subject>
  <name> jim Hendler</name>  </name>
  …
</photo>
```

What is the context? We need information about Talks, Subjects, People, Events, etc. and the roles this item plays in them



Event:WebPage

IAAI-05 Robert S. Engelmore Memor
Knowledge as Power: A View fro
James Hendler, *University of Marylo*

The emerging semantic web focuses on b
to Web applications in a Web-friendly
knowledge on the Web, share it, and reu
mechanisms provides new and interest
intelligence. In this talk, Hendler exp
differences between the semantic web an
representation systems, and sees if he ca
semantic web is to KR as the Web is to hyp

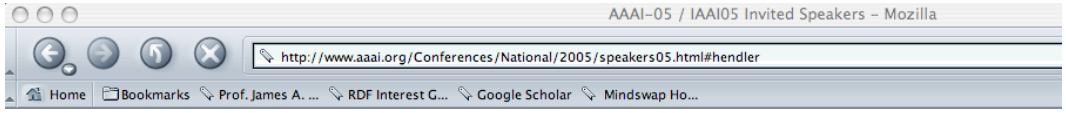
Jim Hendler is a professor and directo
Knowledge Discovery at the University
inventors of the semantic web, Hendler
Fulbright Foundation Fellowship, is a for
Force Science Advisory Board, and is
Association for Artificial Intelligence. H
scientist of the Information Systems Office at the US Defense
Advanced Research Projects Agency (DARPA), was awarded a US Air
Force Exceptional Civilian Service Medal in 2002, and is a member of

<> rdf:type photo:Photograph,
Photo:File http://…/images#image1,
Photo:topic :event1#event:speaker.
Event1 a Event:event;
date “2005-07-13”,
speaker http://…#hendler.html
Title “IAAI 2005…”
:JimH rdf:type foaf:person;
name “Jim Hendler”

```
<daml:ObjectProperty rdf:ID="photograph">
<rdfs:domain rdf:resource="#Picture"/>
<rdfs:range rdf:resource= …#person"/>
</daml:ObjectProperty>
```

```
<s:Class
rdf:about="http://www.semanticweb.org/ont
ologies/swrc-onto-2000-09-
10.daml#Conference">
<s:comment>
describes a generic conceptabout events
        </s:comment>
<s:subClassOf
rdf:resource="http://www.semanticweb.org/o
ntologies/swrc-onto-2000-09-
10.daml#Event"/>
<a:disjointFrom
rdf:resource="http://www.semanticweb.org/o
ntologies/swrc-onto-2000-09-
10.daml#Workshop"/>
<a:restrictedBy
rdf:resource="http://www.semanticweb.org/o
ntologies/swrc-onto-2000-09-
10.daml#genid18"/>
```

```
<rdf:Description rdf:about="http://www.w3.org/2001/03/earl/0.95#Person">
<rdf:type rdf:resource="http://www.w3.org/2000/01/rdf-schema#Class"/>
<rdfs:subClassOf rdf:resource="http://www.w3.org/2001/03/earl/0.95#Assertor"/>
</rdf:Description>
```



Event:WebPage

<daml:ObjectProperty rdf:ID="photograph">
<rdfs:domain rdf:resource="#Picture"/>
<rdfs:range rdf:resource= …#person"/>
</daml:ObjectProperty>

<> rdf:type photo:Photograph,

<s:Class
rdf:about="http://www.semanticweb.org/ont

about events

emanticweb.org/o
09-

emanticweb.org/o
09-

emanticweb.org/o
09-

IAAI-05 Robert S. E
Knowledge as Pov
James Hendler, *Uni*

The emerging semanti
to Web applications
knowledge on the Wel
mechanisms provides
intelligence. In this
differences between th
representation system
semantic web is to KR

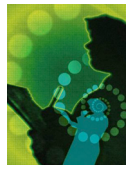
Jim Hendler is a pro
Knowledge Discovery
inventors of the sema
Fulbright Foundation
Force Science Adviso
Association for Artifi
scientist of the Infor
Advanced Research Pr
Force Exceptional Civilian Service Medal in 2002, and is a member of

Semantics is about "real objects and imaginary concepts and the particular relations between them!" (cf. Berners-Lee 94)

And the field that knows how to do this is AI!

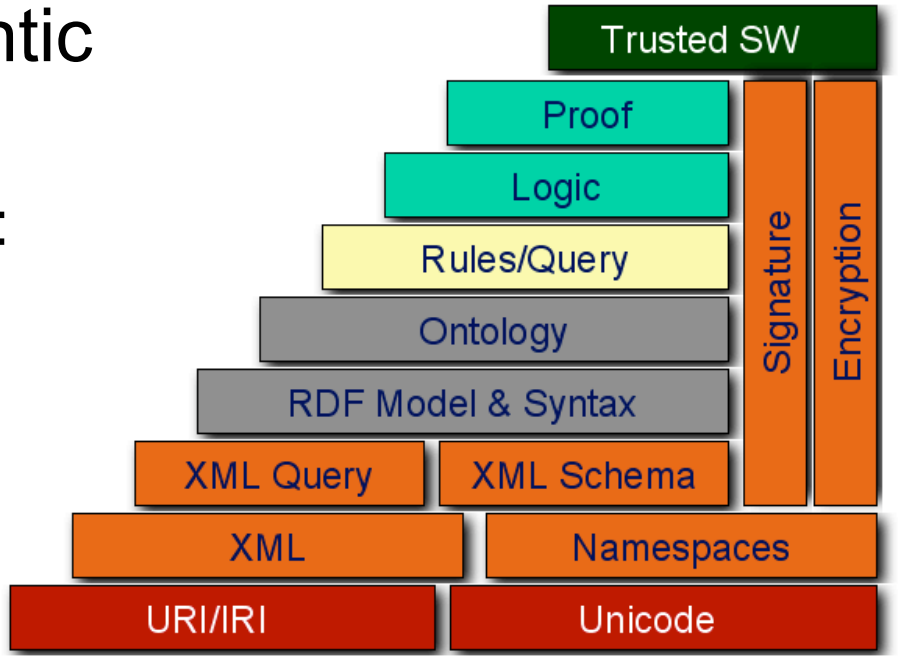
(We call it "knowledge representation"!)

<rdf:Description rdf:about="http://www.w3.org/2001/03/earl/0.95#Person">
<rdf:type rdf:resource="http://www.w3.org/2000/01/rdf-schema#Class"/>
<rdfs:subClassOf rdf:resource="http://www.w3.org/2001/03/earl/0.95#Assertor"/>
</rdf:Description>

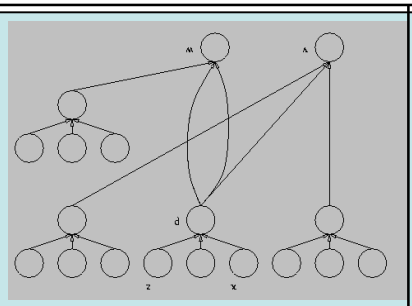
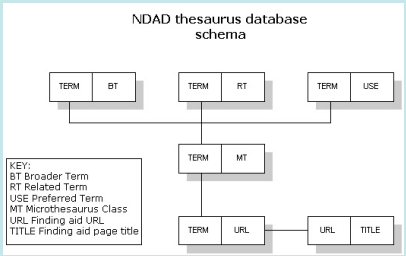
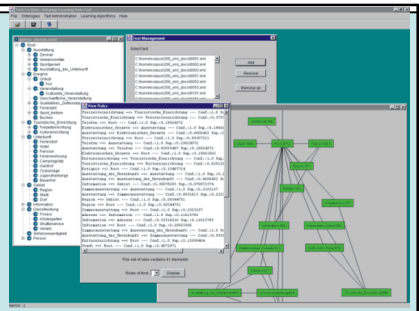
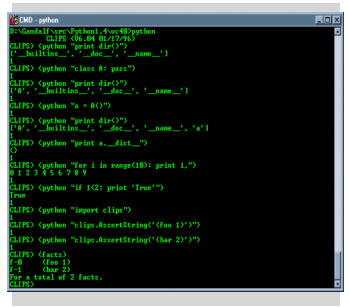


Sem Web Langauge Stack

- The most used slide about the Semantic Web
  - Berners-Lee, 01: Semantic Web Layercake (this version ca. 03)

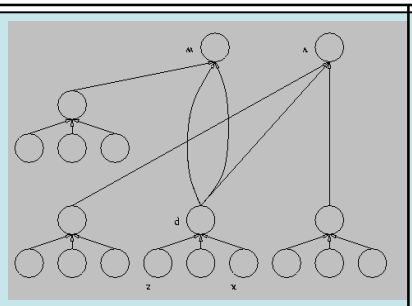
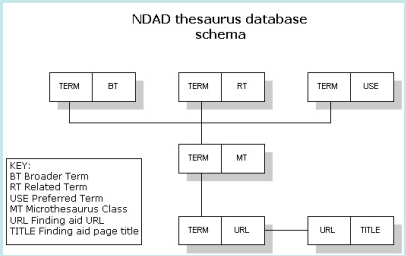
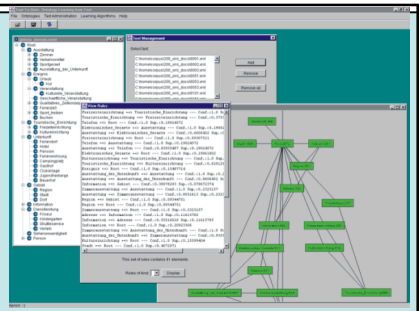
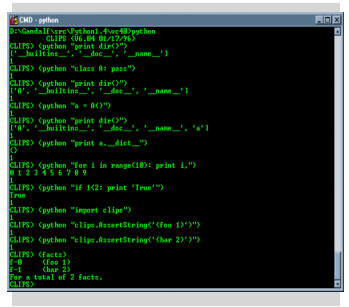


AI languages for the Web

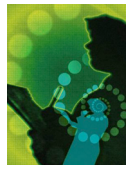
			...	
Semantic Networks (BUT…)	Frame Language (BUT…)	KR Logic-Lite (BUT…)		Next up: OPS5-ish (BUT…)
RDF	RDF Schema	OWL	...	Sem Web Rules Language

All of these languages add semantic modeling primitives to XML.

AI STANDARDS for the Web

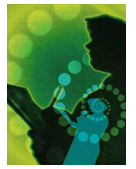
			…	
Semantic Networks (BUT…)	Frame Language (BUT…)	KR Logic-Lite (BUT…)		Next up: OPS5-ish (BUT…)
RDF	RDF Schema	OWL	…	Sem Web Rules Language

RDF, RDFS and OWL are W3C Recommendations!!!



BUT…

- **Using World Wide Web technology!**
  - **Standard Languages** (W3C: Feb, 2004)
    - If it doesn't import/export RDF/OWL, don't buy it!
  - **Web-based approach**
    - No new servers and boxes - use your current Web infrastructure!!!
      - My FOAF file is http://www.cs.umd.edu/~hendler/2003/foaf.rdf
    - Documents for exchange -- no need for exotic new security protocols, etc.
  - **Eventual network effect**
    - When it is time to open up beyond intranets, we know the technology can scale!
- **Don't underestimate the importance of these!**



This is not your father's KR…

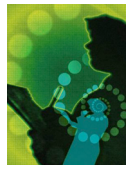
Q. How is OWL different from earlier ontology languages?

A. OWL is a *Web* Ontology language. Where earlier languages have been used to develop tools and ontologies for specific user communities (particularly in the sciences and in company-specific e-commerce applications), they were not defined to be compatible with the architecture of the World Wide Web in general, and the Semantic Web in particular.

OWL rectifies this by providing a language which uses the linking provided by RDF to add the following capabilities to ontologies:

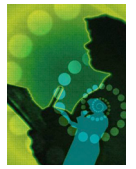
- Ability to be distributed across many systems
- Scalable to Web needs
- Compatible with Web standards for accessibility and internationalization.
- Open and extensible

(OWL FAQ - W3C)



Open and extensible

- Example: I have created the world's most comprehensive ontology on feline leukemia
  - Over 45,000 cancer-related classes
  - Over 47,000 common sense terms



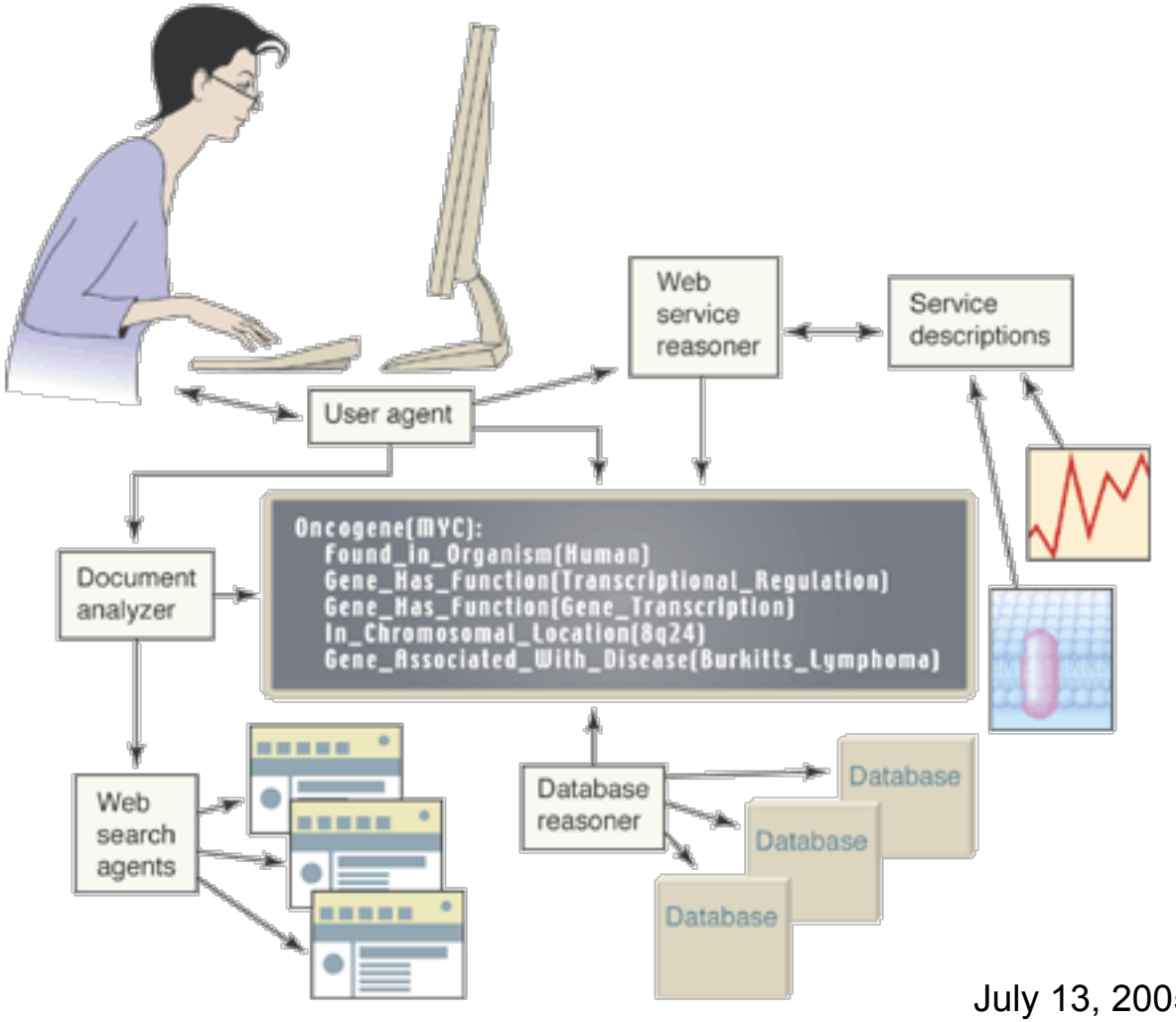
… And it fits on this slide

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE rdf:RDF [
  <!ENTITY feleuk.owl "http://www.mindswap.org/ontologies/feleuk.owl">
  <!ENTITY owl "http://www.w3.org/2002/07/owl#">
  <!ENTITY rdf "http://www.w3.org/1999/02/22-rdf-syntax-ns#">
  <!ENTITY rdfs "http://www.w3.org/2000/01/rdf-schema#">
  <!ENTITY NCI "http://www.ncibi.nih.gov/NCIT/NCIT.owl#">
  <!ENTITY CYC="http://www.cyc.com/2004/06/04/cyc#">
  ]>
<rdf:RDF xml:base="&feleuk.owl;"
      xmlns:owl="&owl;"
      xmlns:rdf="&rdf;"
      xmlns:rdfs="&rdfs;"
      xmlns:NCI="&NCI;"
      xmlns:CYC="&CYC;">

  <owl:Ontology rdf:about=""
            rdfs:label="Feline Leukemia"
            owl:versionInfo="Feline Leuk 1.0"/>

  <owl:Class rdf:about="#Feline-Leukemia">
    <rdfs:subClassOf rdf:resource="NCI:Leukemia"/>
    <rdfs:subClassOf>
      <owl:Restriction>
        <owl:allValuesFrom rdf:resource="CYC:cat"/>
        <owl:onProperty rdf:resource="#NCI:diseased-organism"/>
      </owl:Restriction>
    </rdfs:subClassOf>
  </owl:Class>

</rdf:RDF>
```

Link to 45000 terms at NCI

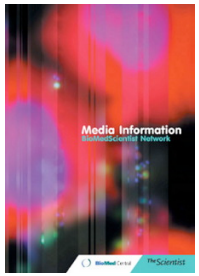
Link to 47000 (Open)CYC ters

- Sem Web languages allow linking of
  - multimedia
  - databases
  - services
    - web
    - Grid
  - meta-data repositories
- Or any other Web resource!



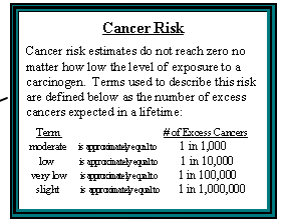
- Including other Semantic Web resources
  - partial mappings just fine
  - this creates a web of models (semantics) much like the current web is a web of texts
- Network effect as mappings provide links to linked resources

BioMedCentral Article

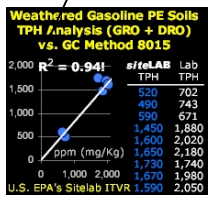


Oncogene(MYC):
Found_In_Organism(Human).
Gene_Has_Function(Transcriptional_Regulation).
Gene_Has_Function(Gene_Transcription).
In_Chromosomal_Location(8q24).
Gene_Associated_With_Disease(Burkitts_Lymphoma).

NCI Cancer Ontology (OWL)



EPA Vocabulary (RDFS)



EPA data set (XML)

```
<meta>
  <classifications>
    <classification type="MYC" subtype="old_arx_id">bcr-2-1-059</classification>
  </classifications>
</meta>
```

BioMedCentral Metadata (XML)

And Web services come to the party too

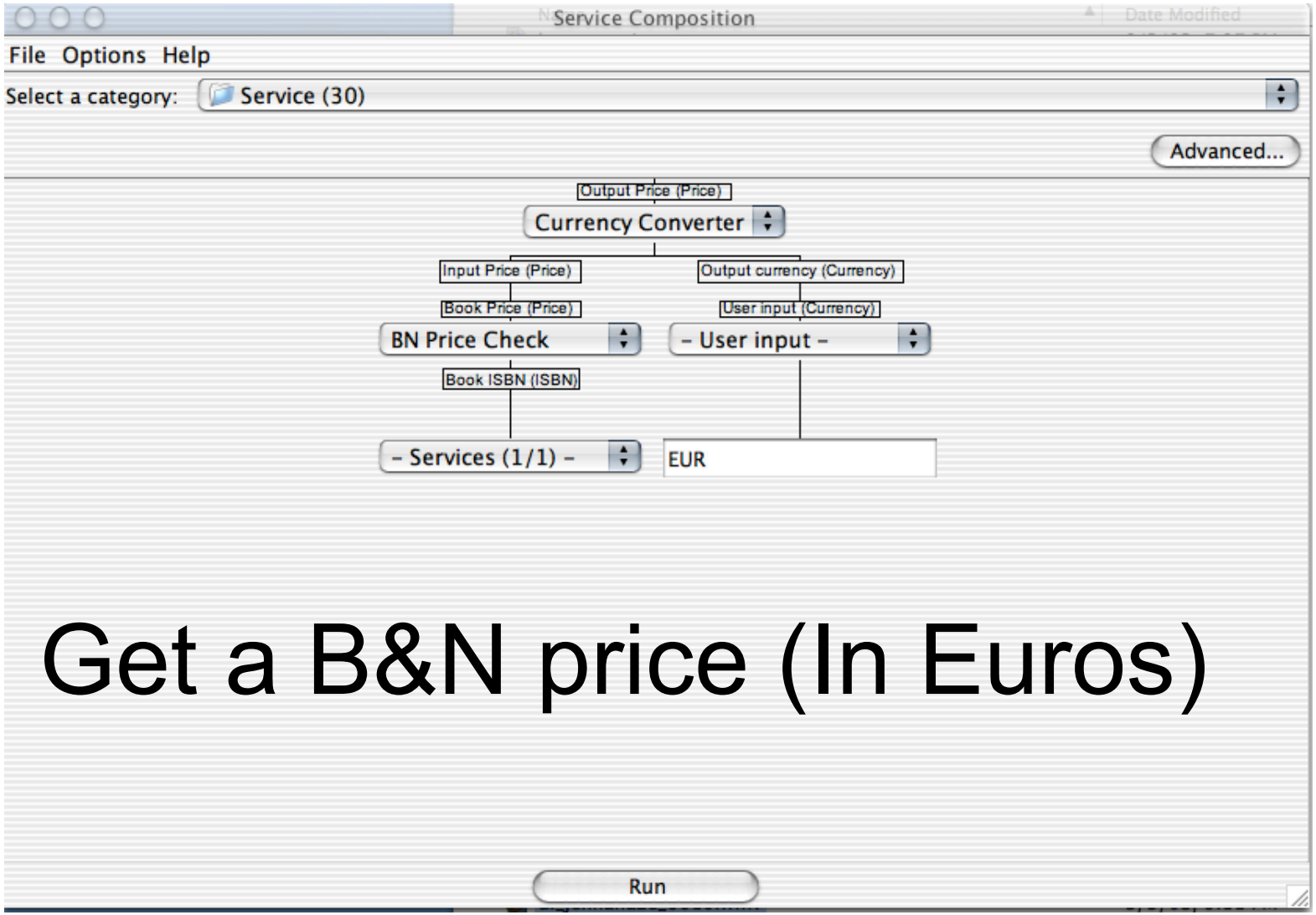
input xsd:complex=”oncogene”

```
<types>
  <xsd:schema xmlns="http://www.w3.org/2001/XMLSchema"
    targetNamespace="urn:GoogleSearch">
    <xsd:complexType name="GoogleSearchResult">
      <xsd:all>
        <xsd:element name="documentFiltering" type="xsd:boolean" />
        <xsd:element name="searchComments" type="xsd:string" />
        <xsd:element name="estimatedTotalResultsCount" type="xsd:int" />
        <xsd:element name="estimateIsExact" type="xsd:boolean" />
        <xsd:element name="resultElements" type="typens:ResultElementAr
        <xsd:element name="searchQuery" type="xsd:string" />
        <xsd:element name="startIndex" type="xsd:int" />
        <xsd:element name="endIndex" type="xsd:int" />
        <xsd:element name="searchTips" type="xsd:string" />
        <xsd:element name="directoryCategories" type="typens:DirectoryCa
        <xsd:element name="searchTime" type="xsd:double" />
      </xsd:all>
    </xsd:complexType>
  </xsd:schema>
</types>
<message name="doGoogleSearch">
  <part name="key" type="xsd:string" />
  <part name="q" type="xsd:string" />
  <part name="start" type="xsd:int" />
  <part name="maxResults" type="xsd:int" />
  <part name="filter" type="xsd:boolean" />
  <part name="restrict" type="xsd:string" />
  <part name="safeSearch" type="xsd:boolean" />
  <part name="lr" type="xsd:string" />
  <part name="ie" type="xsd:string" />
  <part name="oe" type="xsd:string" />
</message>
<message name="doGoogleSearchResponse">
  <part name="return" type="typens:GoogleSearchResult" />
</message>
<operation name="doGoogleSearch">
  <input message="typens:doGoogleSearch" />
  <output message="typens:doGoogleSearchResponse" />
</operation>
</definitions>
```

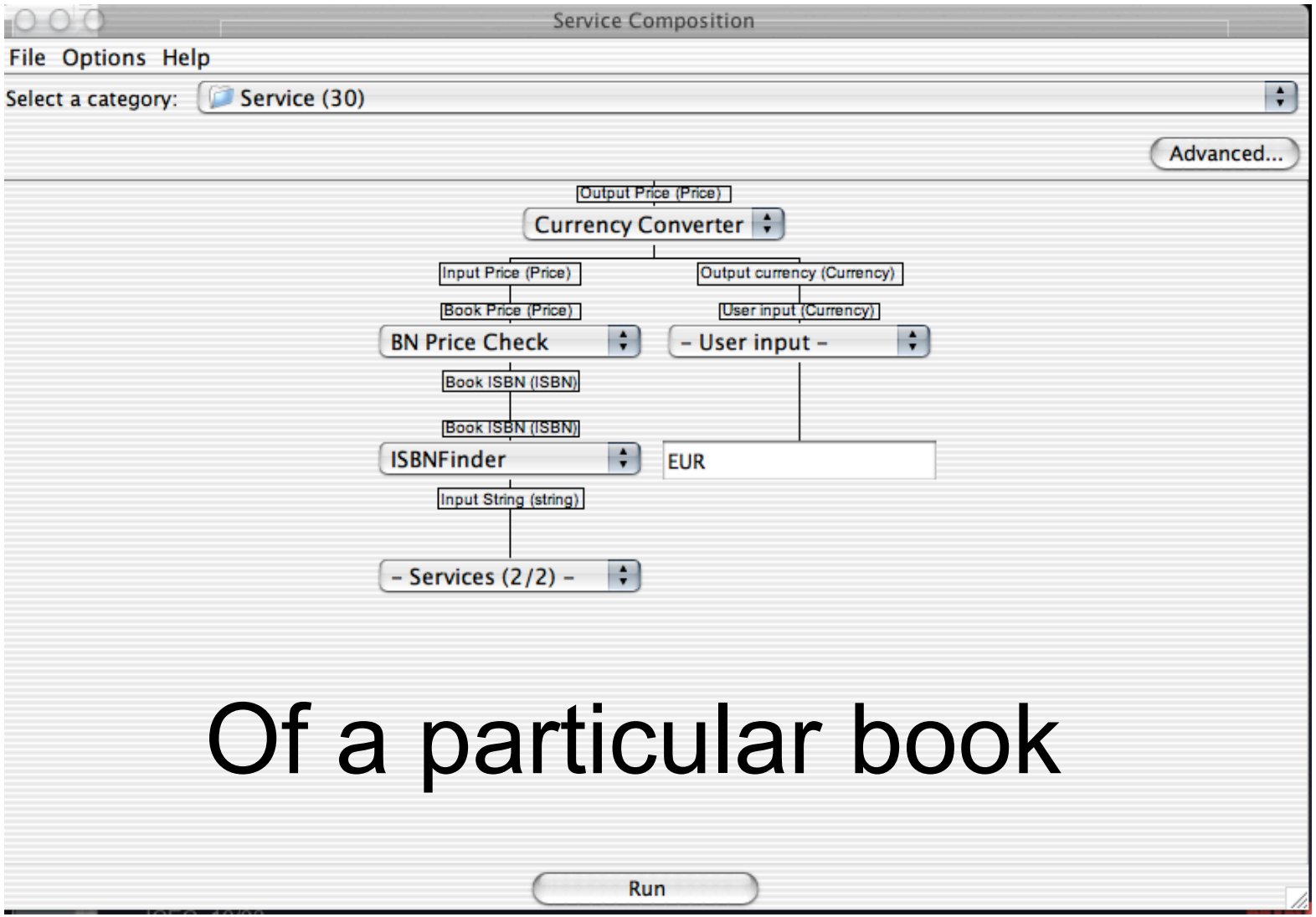
Oncogene(MYC):
Found_In_Organism(Human).
Gene_Has_Function(Transcriptional_Regulation).
Gene_Has_Function(Gene_Transcription).
In_Chromosomal_Location(8q24).
Gene_Associated_With_Disease(Burkitts_Lymphoma).

output xsd:complex=”RiskType”

```
<owl:Class rdf:about="http://annotation.semanticweb.org/iswc/iswc.daml#RiskIndicator">
  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty rdf:resource="http://annotation.semanticweb.org/iswc/iswc.daml#name"/>
      <owl:allValuesFrom rdf:resource="http://www.w3.org/2000/10/XMLSchema#string"/>
    </owl:Restriction>
  </rdfs:subClassOf>
</:Class>
```

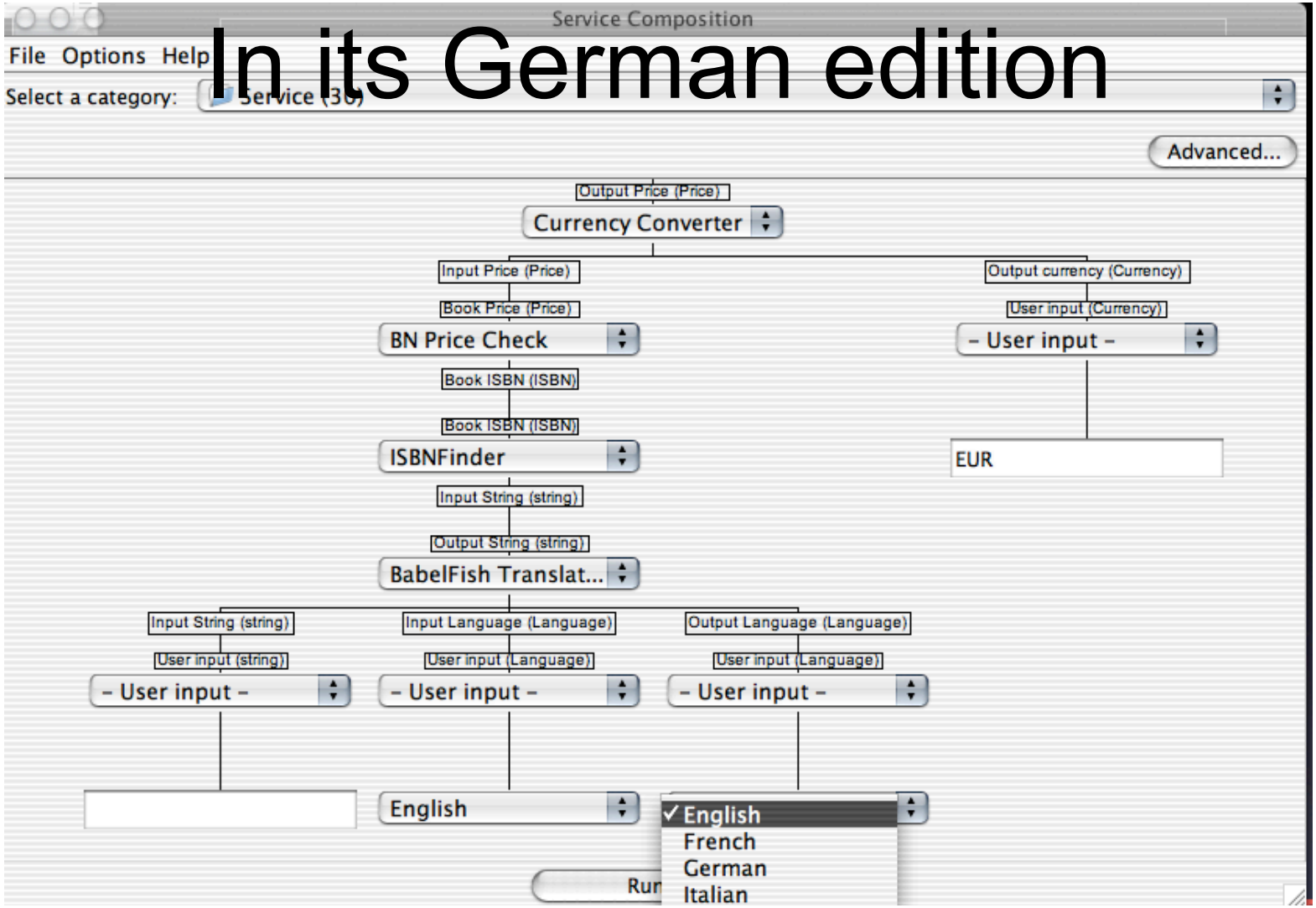


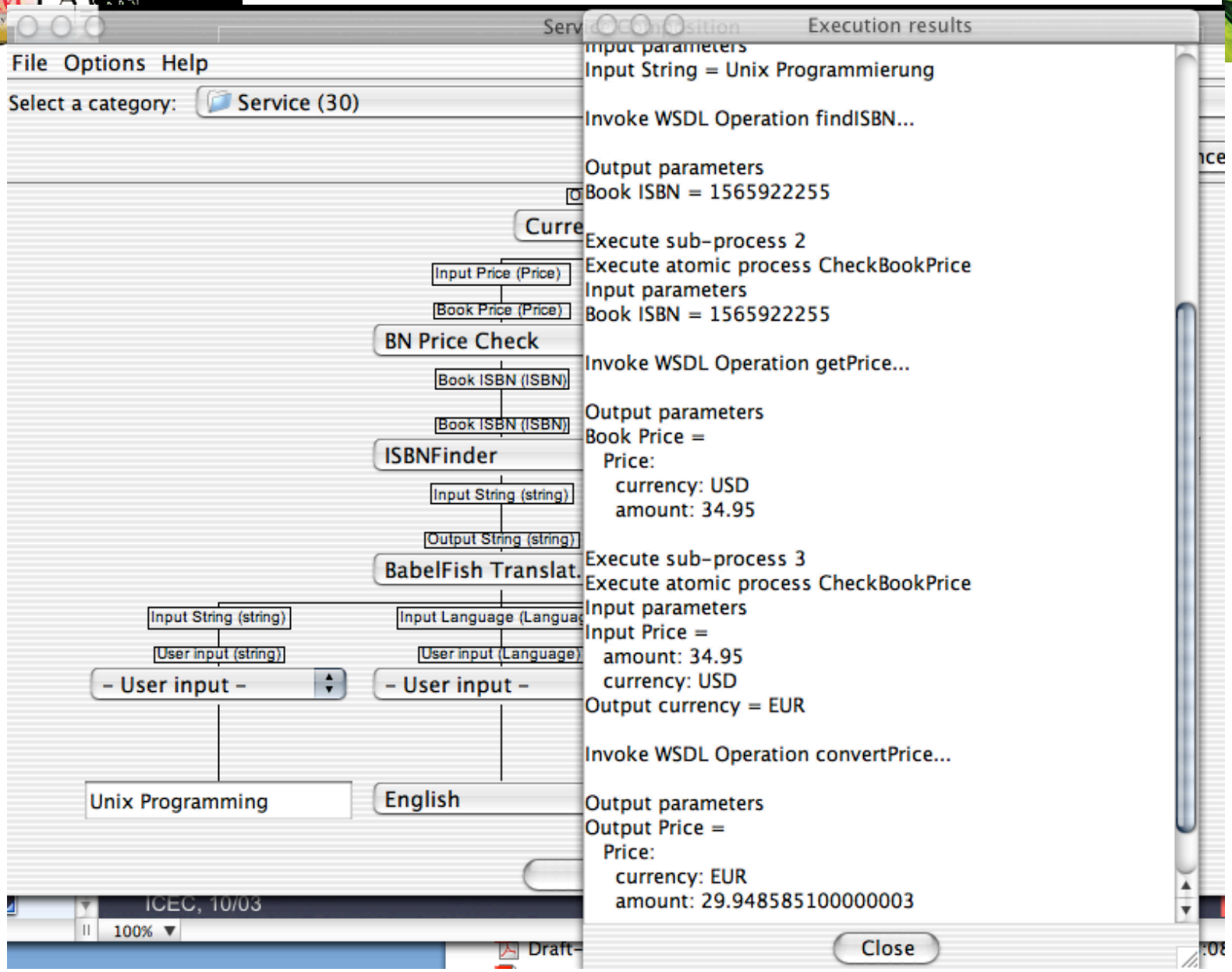
Get a B&N price (In Euros)

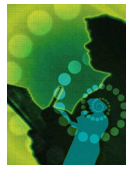


Of a particular book

In its German edition

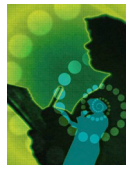



File Options Help
Select a category:
Service (30)
Input Price (Price)
Book Price (Price)
BN Price Check
Book ISBN (ISBN)
Book ISBN (ISBN)
ISBNFinder
Input String (string)
Output String (string)
BabelFish Translat.
Input String (string)
Input Language (Languag
User input (string)
User input (Language)
- User input -
- User input -
Unix Programming
English
Execution results
Input String = Unix Programmierung
Invoke WSDL Operation findISBN...
Output parameters
Book ISBN = 1565922255
Execute sub-process 2
Execute atomic process CheckBookPrice
Input parameters
Book ISBN = 1565922255
Invoke WSDL Operation getPrice...
Output parameters
Book Price =
Price:
currency: USD
amount: 34.95
Execute sub-process 3
Execute atomic process CheckBookPrice
Input parameters
Input Price =
amount: 34.95
currency: USD
Output currency = EUR
Invoke WSDL Operation convertPrice...
Output parameters
Output Price =
Price:
currency: EUR
amount: 29.948585100000003
Close



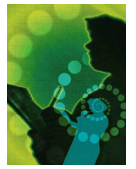
And it's real…

- Companies getting into the act
  - Oracle to support RDF in database 10.2
  - IBM SNObase ontology management system
  - Adobe embeds RDF in all content
  - HP, Cisco, Nokia, Sun … announcements/use in '05
  - Start ups in the space (Cerebra, Siderean, SandPiper, …)
    - And failure already (Tucana)
      - Whose IP was bought by large contractor (Northrup Grumman)
- Many open source tools available
  - Open source (Kowari) scalable triple store
    - 100,000,000+ triples
    - Supports RDFS, OWL support coming
  - RDFLib, 3Store…
  - Jena, Sesame …
  - Protégé, SWOOP, Onto(xxx)…



Organization	Usage	Generic Capability
US Intel Agency	Unstructured intel document analysis	Scalable cross-document co-reference & link analysis
US Dept. of Defense	Federated search and analysis	Ontology-enhanced search & discovery with cross-document co-reference & link analysis
Legal Research Ctr.	Unstructured legal document search and analysis	Ontology-enhanced search with cross-document co-reference & link analysis
Viewpoint Systems General Motors	Quality test & measurement analysis	Advanced manufacturing data analysis with rapidly changing data schemas
Enterprise Software Company	Centralized meta-data repository; EAI software platform	Change impact analysis & configuration management
Sun Microsystems	RDF/OWL representation of products and services	Dynamic Web site navigation (SwoRDFish)

(From David Wood)



Semantic Web Commercial Roadmap

Short-term:
EII

RDF triple stores
OWL as "semantic technology" standard
Database and Image Markup

"RDF/OWL in the business enterprise"

"intranet" uses

Coming:

Personal Data Integration

Home enterprise
Social Network-based apps
Small business portals

"Rules and policy"

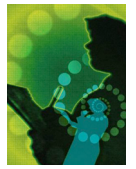
"OEM" market

And even:
Small business EAI

Every doctor, dentist, laundromat, etc. has a "supply chain"
Do business your way, map to their ASPs

"Like visicalc did to reports"

Disruptive technology



Sem Web "Technology" Roadmap

Long-term

Automation Focus
learning/modifying, model-based interaction

"information factory"

Mid-term

"Rule" Focus
Development/
personalization
Access control

Help users explore/analyze relationships in data

Short-term

Ontology Focus
Use, creation, tools, editing, collaboration, sharing, learning/modifying

Support collaboration/sharing

Research Example: Access and Privacy Control

URI
variable

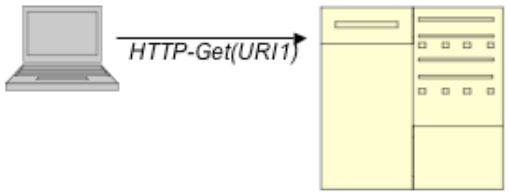
1) If X is AC rep of Y, X can delegate W3C member access rights in Y.
2) *Kari* is AC rep of *Elisa* .

1) If X is employee of *Elisa*,
X has W3C member access rights.
2) *Tiina* is employee of *Elisa*.

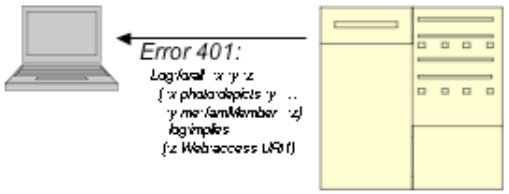
Tiina: I have W3C member access rights
Proof: Alan 1, Alan 2, Kari 1, Kari 2

W3C MEMBER site

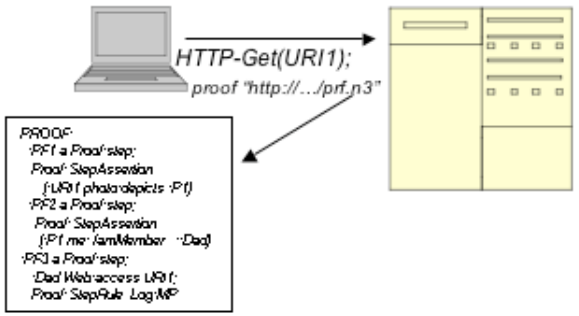
Policy Aware WEB



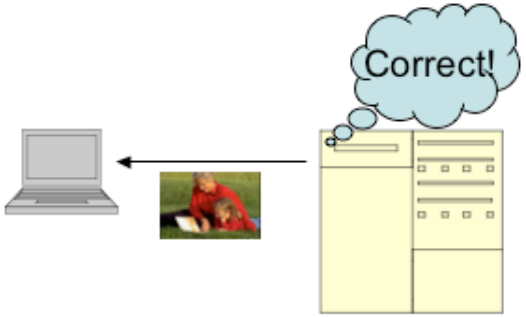
(A) User requests a resource.



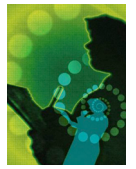
(B) 401 error provides access rules.



(C) Proof is generated and pointer is sent in new HTTP-Get request.

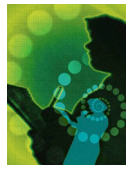


(D) Proof is checked, and confirmed, and the transaction succeeds.



You are here…

Documents, linked to
  Images, annotated with
    Ontologies, linked to
      Other ontologies, describing
        Databases, exported as
          RDF graphs, inputted to
            Services, which designate
Documents, linked to
      (*ad infinitum*)



The other shoe falls

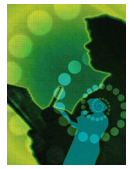
- I gave a talk at KR called "The Semantic Web: KR's worst nightmare"
  - KBs in AI are assumed to be consistent
  - KR&R stresses expressivity over simplicity

< > :responseCode
[rdfs:subClassOf http:Error404] .

--- Ontology Not Found

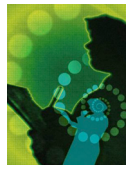
You've encountered a "Ontology Not Found" error while trying to access a semantic term grounded on the University of Maryland Computer Science Deparment semantic web server.

  - On the Web, you cannot assume these!



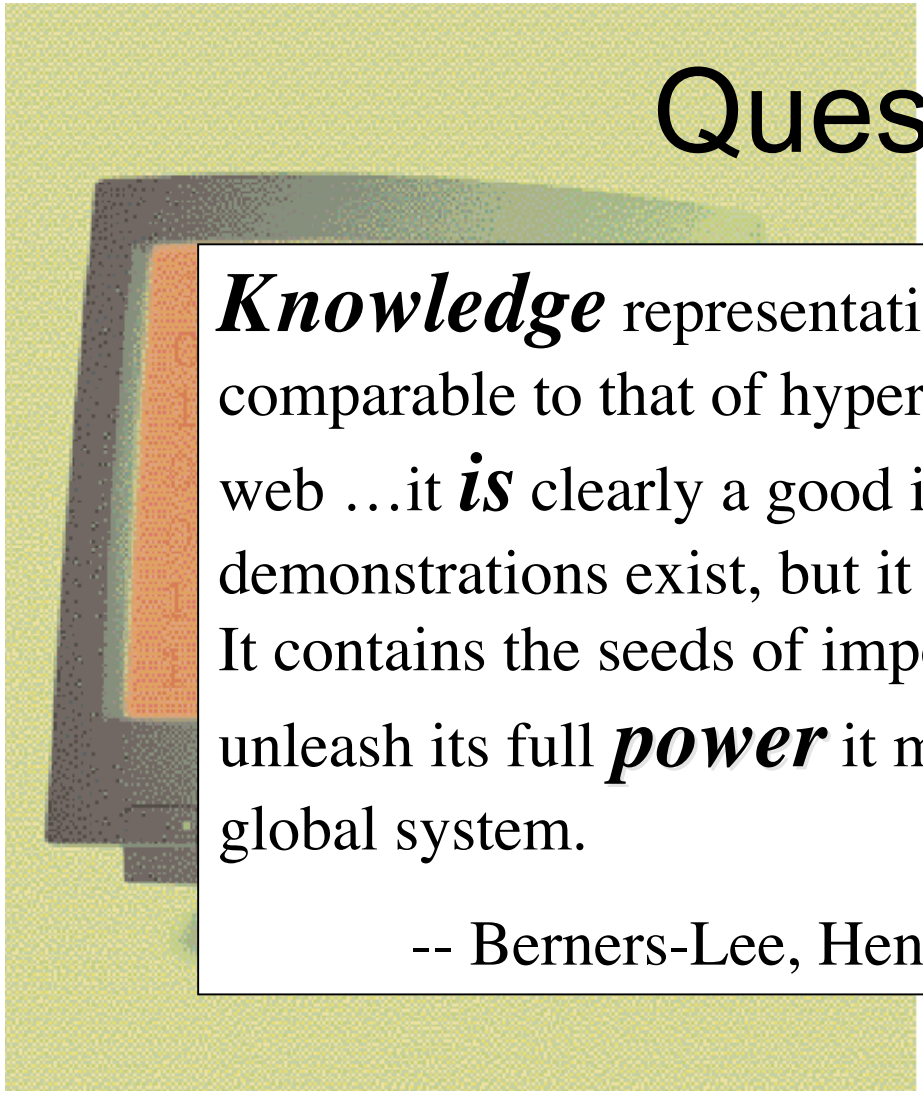
So

- Semantic Web Languages (RDF/RDFS/OWL)
  - Are not-very-expressive-KR-language standards
    - Not KIF, or even KL-ONE
  - Create non-persistent KBS
    - Servers come and go
    - Ontologies change over time
  - And can't be kept consistent
    - Disagreement, error, dishonesty…



In short

- Semantic Web Languages (RDF/RDFS/OWL)
  - Are not-very-expressive-KR-language standards
    - Like HTML is to SGML
  - Create non-persistent KBs
    - Like the 404 error (w/o which there is no Web)
  - And can't be kept consistent
    - Like the blog-space and Web 2.0
- Which is why they might be just what the doctor ordered…



Questions?

Knowledge representation … is currently in a state comparable to that of hypertext before the advent of the web …it ***is*** clearly a good idea, and some very nice demonstrations exist, but it has not yet changed the world. It contains the seeds of important applications, but to unleash its full ***power*** it must be linked into a single global system.

-- Berners-Lee, Hendler, Lassila, 2001.