Semantic Web: an Introduction

Jian-hua Yeh (葉建華)

真理大學資訊科學系助理教授

au4290@email.au.edu.tw

Textbook

- **A Semantic Web Primer**
- **By Grigoris Antoniou, Frank van Harmelen**
- **天瓏圖書公司**

History of the Semantic Web

- **Web was "invented" by Tim Berners-Lee (amongst others), a physicist working at CERN**
- **TBL's original vision of the Web was much more ambitious than the reality of the existing (syntactic) Web:**

"... a goal of the Web was that, if the interaction between person and hypertext could be so intuitive that the machine-readable information space gave an accurate representation of the state of people's thoughts, interactions, and work patterns, then machine analysis could become a very powerful management tool, seeing patterns in our work and facilitating our working together through the typical problems which beset the management of large organizations."

- **TBL (and others) have since been working towards realising this vision, which has become known as the Semantic Web**
  - **E.g., article in May 2001 issue of Scientific American…**

Scientific American, May 2001:

Beware of the Hype

THE SEMANTIC WEB

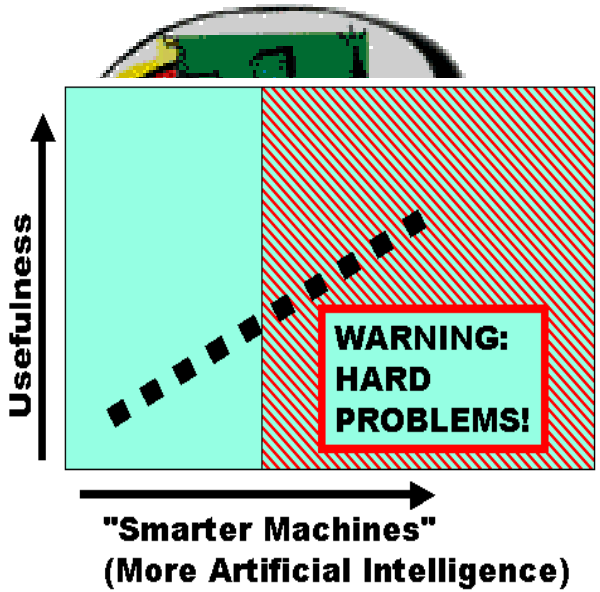
A new form of Web content
that is meaningful to computers
will unleash a revolution of new abilities

by
TIM BERNERS-LEE,
JAMES HENDLER and
ORA LASSILA

PHOTO CREDIT HERE

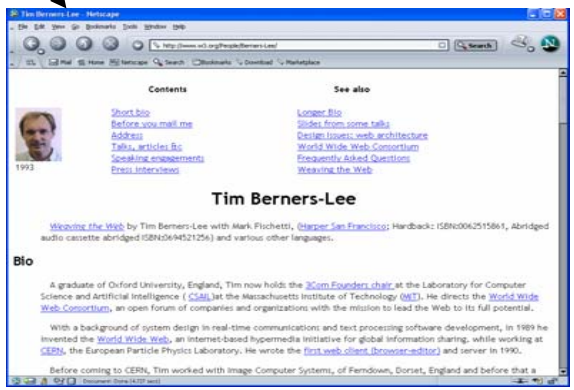
Beware of the Hype

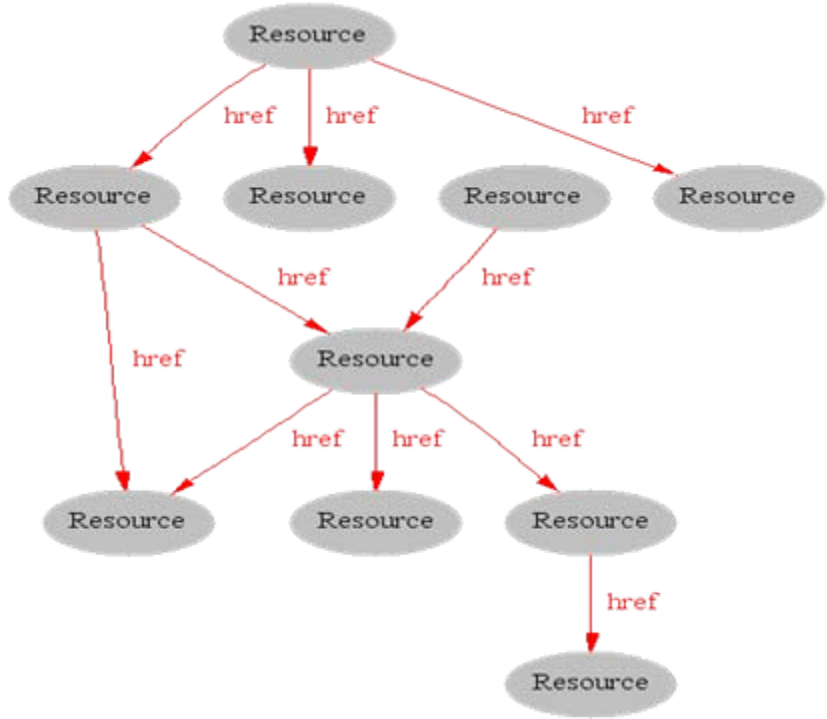
- **Hype seems to suggest that Semantic Web means: "semantics + web = AI"**
  - **"A new form of Web content that is meaningful to computers will unleash revolution of new abilities"**
- **More realistic to think of it as meaning "semantics + web + AI = more useful web"**
  - **Realising the complete "vision" is too hard for now (probably)**
  - **But we can make a start by adding semantic annotation to web resources**



Images from Christine Thompson and David Booth

Where we are Today: the Syntactic Web





[Hendler & Miller 02]

The Syntactic Web is…

- **A hypermedia, a digital library**
  - **A library of documents called (web pages) interconnected by a hypermedia of links**
- **A database, an application platform**
  - **A common portal to applications accessible through web pages, and presenting their results as web pages**
- **A platform for multimedia**
  - **BBC Radio 4 anywhere in the world! Terminator 3 trailers!**
- **A naming scheme**
  - **Unique identity for those documents**

A place where computers do the presentation (easy) and people do the linking and interpreting (hard).

Why not get computers to do more of the hard work?

[Goble 03]

Hard Work using the Syntactic Web…

Find images of Peter Patel-Schneider, Frank van Harmelen and Alan Rector…

Rev. Alan M. Gates, Associate Rector of the Church of the Holy Spirit, Lake Forest, Illinois

Impossible (?) using the Syntactic Web…

- **Complex queries involving background knowledge**
  - **Find information about "animals that use sonar but are not either bats or dolphins", e.g., Barn Owl**
- **Locating inform epositories**
  - **Travel enquiries**
  - **Prices of goods**
  - **Results of huma ments**
- **Finding and usi es"**
  - **Visualise surfac tween two proteins**
- **Delegating com eb "agents"**
  - **Book me a holic somewhere warm, not too far away, an ak French or English**

What is the Problem?

- **Consider a typical web page:**

- **Markup consists of:**
  - **rendering information (e.g., font size and colour)**
  - **Hyper-links to related content**
- **Semantic content is accessible to humans but not (easily) to computers…**

What information can we see…

WWW2002

The eleventh international world wide web conference

Sheraton waikiki hotel

Honolulu, hawaii, USA

7-11 may 2002

1 location 5 days learn interact

Registered participants coming from

australia, canada, chile denmark, france, germany, ghana, hong kong, india, ireland, italy, japan, malta, new zealand, the netherlands, norway, singapore, switzerland, the united kingdom, the united states, vietnam, zaire

Register now

On the 7th May Honolulu will provide the backdrop of the eleventh international world wide web conference. This prestigious event …

Speakers confirmed

Tim berners-lee

Tim is the well known inventor of the Web, …

Ian Foster

Ian is the pioneer of the Grid, the next generation internet …

What information can a machine see…

Solution: XML markup with “meaningful” tags?

<name>⌖⌖⌖🗎🗀🗀🗎
❄≋♏ ♏●♏❖♏■♦≋ ♓■♦♏❒■♋♦♓□■♋● ⬩□❒●♎ ⬩♓♎♏ ⬩♏♌♍□■</name>

<location>♦≋♏❒♋♦□■ ⬩♋♓&♓&♓ ≋□♦♏●
☝□■□●◆●◆🖃 ≋♋⬩♋♓♓🖃 ✞♦✌</location>

<date>🖮🖃🗁🗁 ❍♋🖂 🗎🗀🗀🗎</date>

<slogan>🗁 ●□♍♋♦♓□■ 🎛 ♎♋🖂⬩ ●♏♋❒■ ♓■♦♏❒♋♍♦</slogan>

<participants>☼♏♑♓⬩♦♏❒♏♎ ◘♋❒♦♓♍♓◘♋■♦⬩ ♍□❍♓■♑ ↗❒□❍
♋◆⬩♦❒♋●♓♋🖃 ♍♋■♋♎♋🖃 ♍≋♓●♏ ♎♏■❍♋❒&🖃 ↗❒♋■♍♏🖃
♑♏❒❍♋■🖂🖃 ♑≋♋■♋🖃 ≋□■♑ &□■♑🖃 ♓■♎♓♋🖃 ♓❒♏●♋■♎🖃
♓♦♋●🖂🖃 ☞♋◘♋■🖃 ❍♋●♦♋🖃 ■♏⬩ ⌘♏♋●♋■♎🖃 ♦≋♏
■♏♦≋♏❒●♋■♎⬩🖃 ■□❒⬩♋🖂🖃 ⬩♓■♑♋◘□❒♍🖃 ⬩⬩♓♦⌘♏❒●♋■♎🖃
♦≋♏ ◆■♓♦♏♎ &♓■♑♎□❍🖃 ♦≋♏ ◆■♓♦♏♎ ⬩♦♋♦♏⬩🖃 ❖♓♏♦■♋❍🖃
⌘♋♓❒♏</participants>

<introduction>☼♏♑♓⬩♦♏❒ ■□⬩
🏳■ ♦≋♏ 🖮⬩≋ 💣♋🖂 ☝□■□●◆●◆ ⬩♓●● ◘❒□❖♓♎♏ ♦≋♏
♌♋♍&♎❒□◘ □↗ ♦≋♏ ♏●♏❖♏■♦≋ ♓■♦♏❒■♋♦♓□■♋● ⬩□❒●♎
⬩♓♎♏ ⬩♏♌ ♍□■↗♏❒♏■♍♏🖃 ❄≋♓⬩ ◘❒♏⬩♦♓♑♓□◆⬩ ♏❖♏■♦ ⑤
♦◘♏♋&♏❒⬩ ♍□■↗♓❒❍♏♎</introduction>

<speaker>❄♓❍ ♌♏❒■♏❒⬩🖃●♏♏</speaker>

<bio>❄♓❍ ♓⬩ ♦≋♏ ⬩♏●● &■□⬩■ ♓■❖♏■♦□❒ □↗ ♦≋♏
⌖♏♌🖃</bio>...

But What About…

Need to Add "Semantics"

- **External agreement on meaning of annotations**
  - **E.g., Dublin Core**
    - **Agree on the meaning of a set of annotation tags**
  - **Problems with this approach**
    - **Inflexible**
    - **Limited number of things can be expressed**
- **Use Ontologies to specify meaning of annotations**
  - **Ontologies provide a vocabulary of terms**
  - **New terms can be formed by combining existing ones**
  - **Meaning (semantics) of such terms is formally specified**
  - **Can also specify relationships between terms in multiple ontologies**

Ontology: Origins and History

Ontology in Philosophy

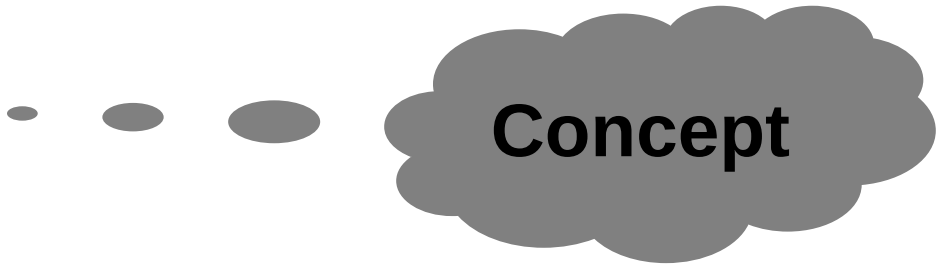
a philosophical discipline—a branch of philosophy that deals with the nature and the organisation of reality

- **Science of Being (Aristotle, Metaphysics, IV, 1)**
- **Tries to answer the questions:**

  What characterizes being?

  Eventually, what is being?

Ontology in Linguistics



activates

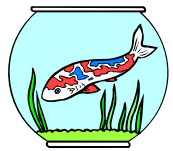
Relates to

Form

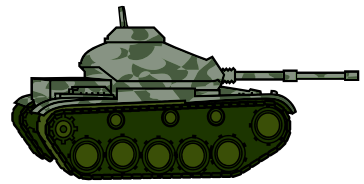
Stands for

Referent

“Tank“



?



[Ogden, Richards, 1923]

Ontology in Computer Science

- **An ontology is an engineering artifact:**
  - **It is constituted by a specific vocabulary used to describe a certain reality, plus**
  - **a set of explicit assumptions regarding the intended meaning of the vocabulary.**

- **Thus, an ontology describes a formal specification of a certain domain:**
  - **Shared understanding of a domain of interest**
  - **Formal and machine manipulable model of a domain of interest**

“An explicit specification of a conceptualisation” [Gruber93]

Structure of an Ontology

Ontologies typically have two distinct components:

- **Names for important concepts in the domain**
  - **Elephant is a concept whose members are a kind of animal**
  - **Herbivore is a concept whose members are exactly those animals who eat only plants or parts of plants**
  - **Adult_Elephant is a concept whose members are exactly those elephants whose age is greater than 20 years**
- **Background knowledge/constraints on the domain**
  - **Adult_Elephants weigh at least 2,000 kg**
  - **All Elephants are either African_Elephants or Indian_Elephants**
  - **No individual can be both a Herbivore and a Carnivore**

A Semantic Web — First Steps

Make web resources more accessible to automated processes

- **Extend existing rendering markup with semantic markup**
  - **Metadata annotations that describe content/funtion of web accessible resources**
- **Use Ontologies to provide vocabulary for annotations**
  - **“Formal specification” is accessible to machines**

- **A prerequisite is a standard web ontology language**
  - **Need to agree common syntax before we can share semantics**
  - **Syntactic web based on standards such as HTTP and HTML**

Ontology Design and Deployment

- **Given key role of ontologies in the Semantic Web, it will be essential to provide tools and services to help users:**
  - **Design and maintain high quality ontologies, e.g.:**
    - **Meaningful — all named classes can have instances**
    - **Correct — captured intuitions of domain experts**
    - **Minimally redundant — no unintended synonyms**
    - **Richly axiomatised — (sufficiently) detailed descriptions**
  - **Store (large numbers) of instances of ontology classes, e.g.:**
    - **Annotations from web pages**
  - **Answer queries over ontology classes and instances, e.g.:**
    - **Find more general/specific classes**
    - **Retrieve annotations/pages matching a given description**
  - **Integrate and align multiple ontologies**

Example Ontology