

The Semantic Web Landscape

Dean Allemang
Working Ontologist, LLC

Semantic Web – what it is and isn't

Lot's of cool technologies that could be seen as “Semantic”:



Translate



What does “Semantic” mean?

Definition of **SEMANTIC**

1: of or relating to meaning in language



What is the meaning of meaning?



What is the “Web”?



A way to share **documents** on a global scale.
Anyone can **read** any **document** on the web

What is the “Semantic Web”?



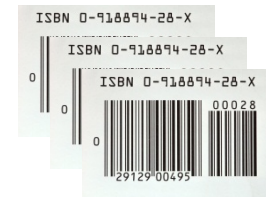
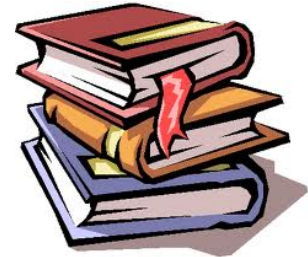
A way to share **data** on a global scale.
Anyone can **query** any **data** on the web
all

Problems you can solve on the web

- **Search** (help me find what I am looking for)
- **Comparison** (help me compare two things)
- **Route planning** (get me from here to there)
- **Diagnosis** (what's wrong with me?)
- **Measurement** (how much? How far?)
- **Broadcast** (let everyone know)
- ...

Semantic Web doesn't do any of these – it allows data to be shared, to facilitated these functions and more

Example: search

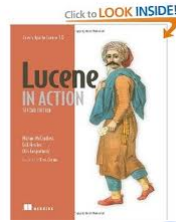


An index to guide the search

Stuff to search through (information, products, services, ...)

Example: search

- Text analytics?
- Link analytics?
- Crowdsourcing?
- Self-annotation?
- Expert curation?



Google

amazon mechanical turk
beta Artificial Intelligence

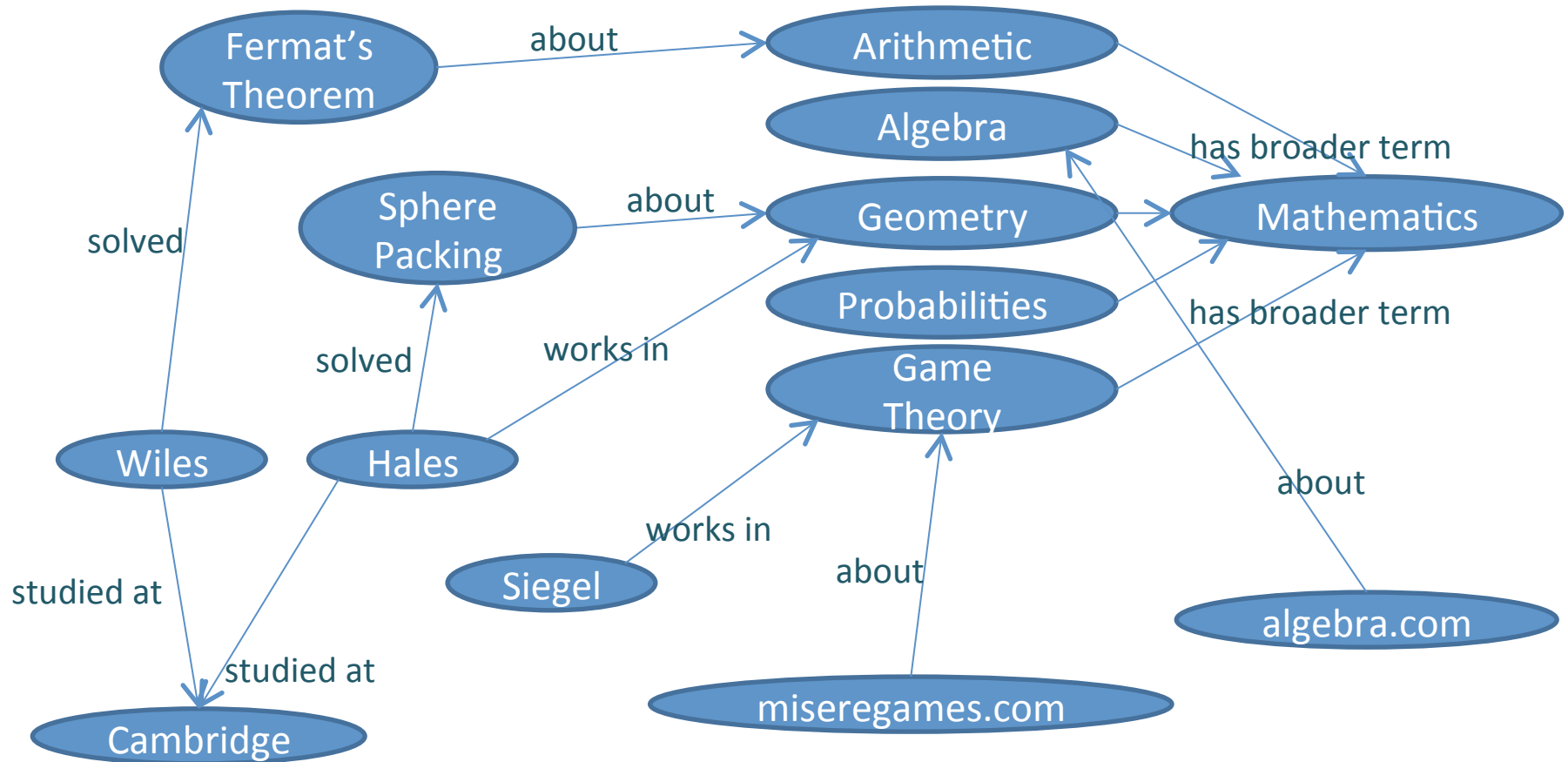


Where does that index come from?



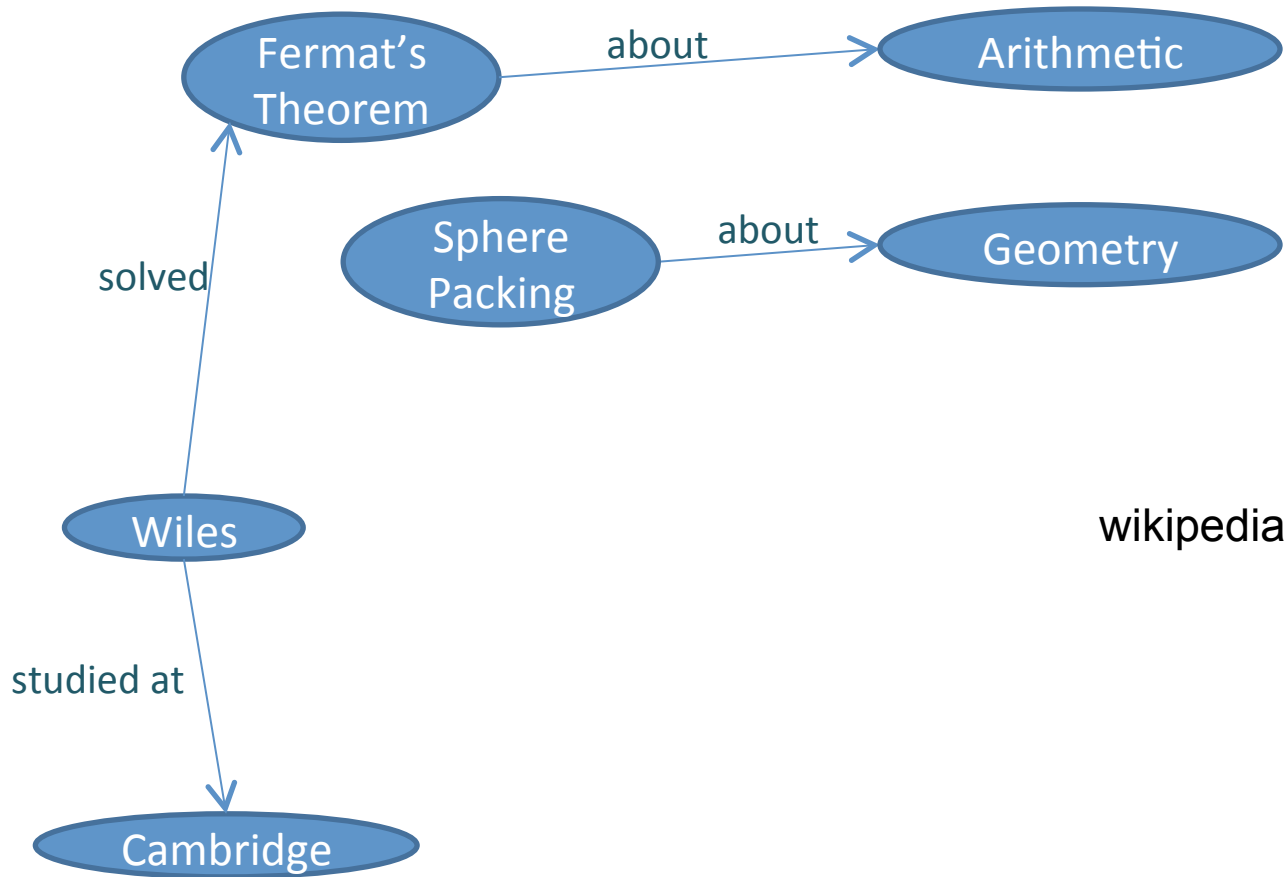
Sharing data on the Semantic Web

Semantic Web description of Mathematics (abridged)



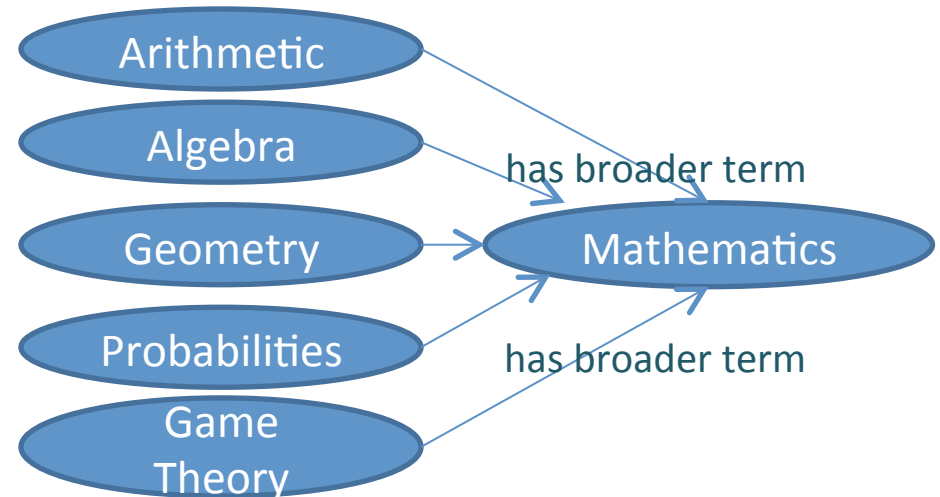
Semantic Web – emphasis is on the “**WEB**”

Semantic Web description of Mathematics (abridged)



Semantic Web – emphasis is on the “**WEB**”

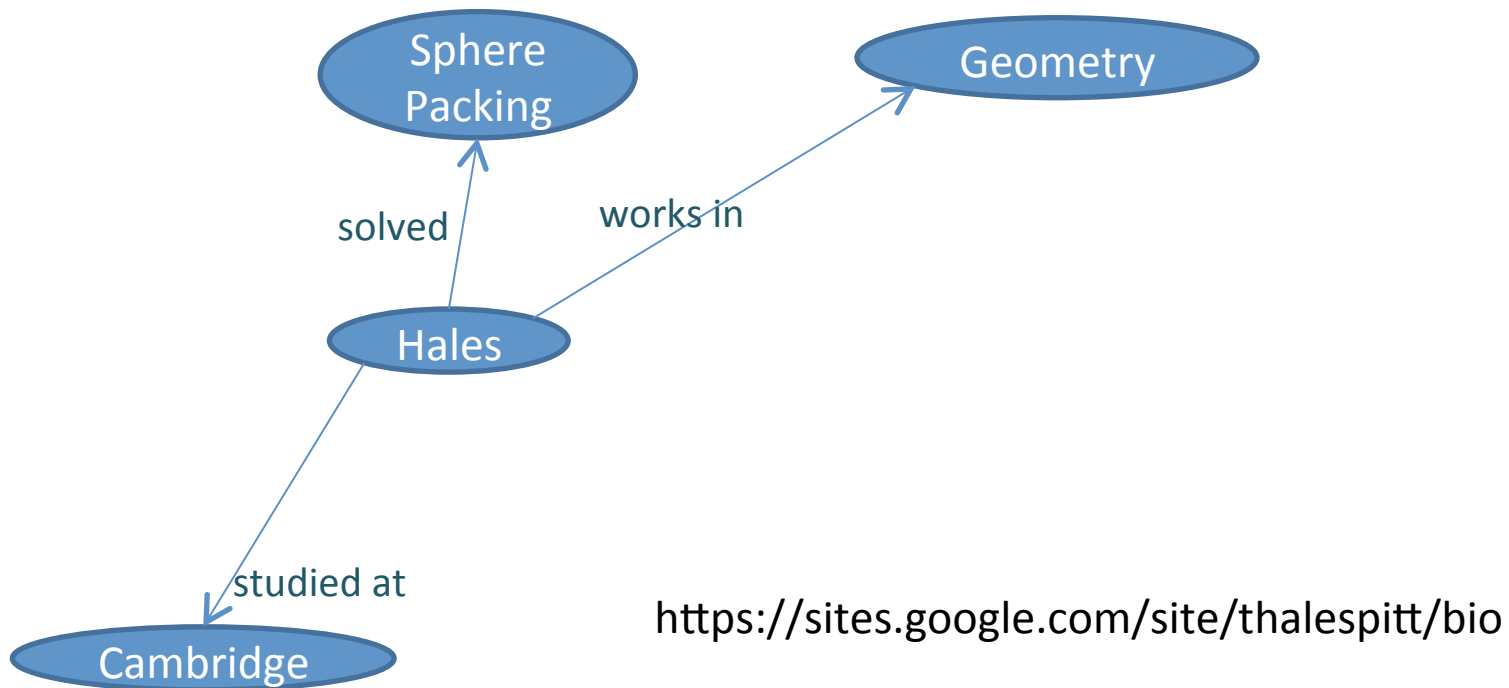
Semantic Web description of Mathematics (abridged)



Library of Congress

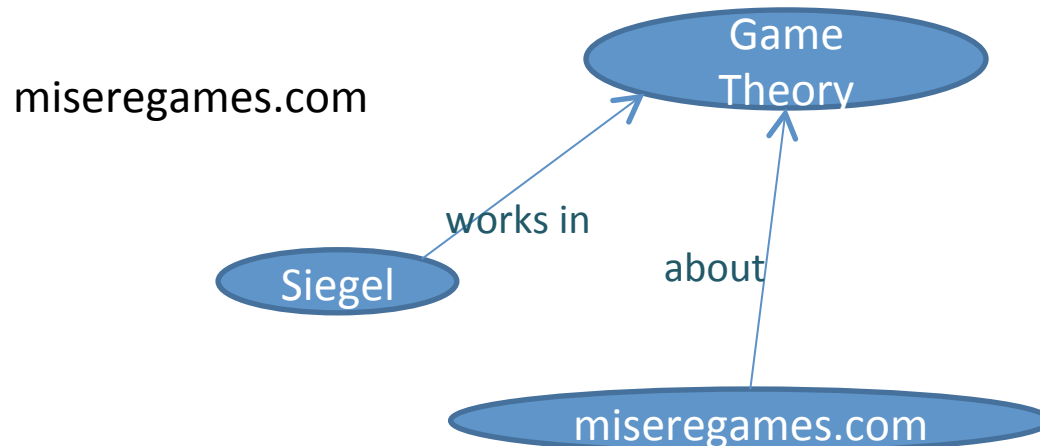
Semantic Web – emphasis is on the “**WEB**”

Semantic Web description of Mathematics (abridged)



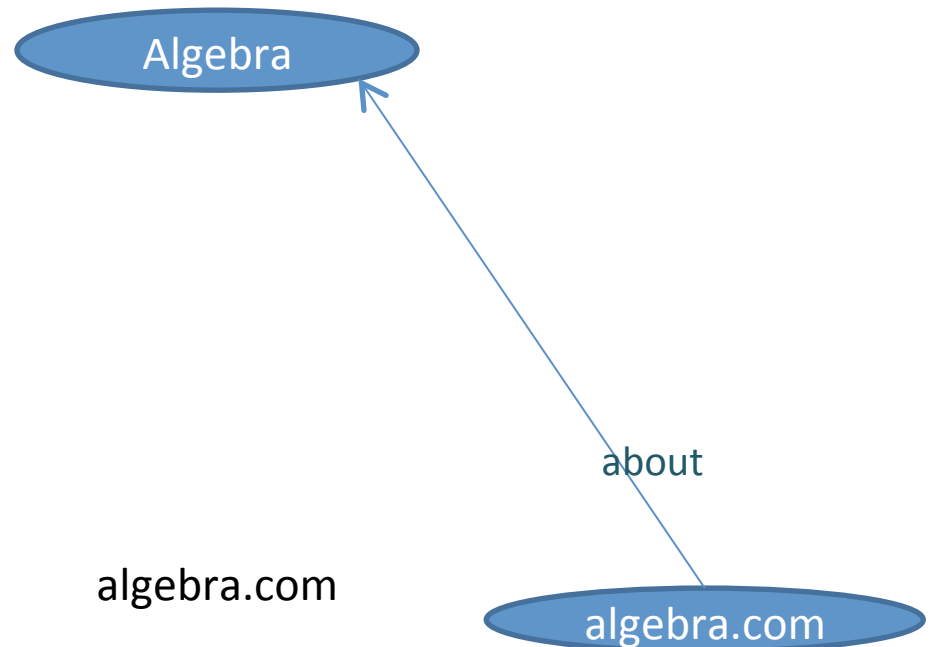
Semantic Web – emphasis is on the “**WEB**”

Semantic Web description of Mathematics (abridged)




Semantic Web – emphasis is on the “**WEB**”

Semantic Web description of Mathematics (abridged)



Semantic Web – emphasis is on the “**WEB**”



On the Semantic Web, metadata is as easy
to manage as data

Example: Structured Search

“Find me someone who has ‘Semantic Web’ as part of their job title”

What kind of index would we need to answer this sort of question?

Interlingua

The image shows a screenshot of a website with red annotations mapping data to a central schema. The website is the Stanford Center for Biomedical Informatics Research, formerly Stanford Medical Informatics. The schema lists the following fields: name, position, work phone, email, and photo. Red boxes and arrows connect these fields to specific data points on the website. For example, 'name' points to 'Mark A. Musen', 'position' points to 'Professor of Medicine (Biomedical Informatics) Division Head (BMIR), Co-Director, Biomedical Informatics Training Program', 'work phone' points to '650-725-3390 (office)', 'email' points to 'musen@stanford.edu', and 'photo' points to a portrait of Mark A. Musen. Another set of annotations points to a profile for Ivan Herman, with 'name' pointing to 'Ivan Herman', 'position' pointing to 'I am Semantic Web Activity Lead', 'work phone' pointing to 'As part of my work, I also participate in lots of outreach activities...', 'email' pointing to 'ivan@w3.org', and 'photo' pointing to a portrait of Ivan Herman.

STANFORD CENTER FOR BIOMEDICAL INFORMATICS RESEARCH
FORMERLY STANFORD MEDICAL INFORMATICS

HOME NEWS & ANNOUNCEMENTS NATIONAL RESOURCES RESEARCH LABS PROJECTS

Musen Lab
Shah Lab

ABOUT US
PEOPLE
EVENTS
PUBLICATIONS
PROJECTS
AFFILIATES PROGRAM
EDUCATION
RESEARCH LABS
NATIONAL RESOURCES
COMMUNITY
CONTACT US

BMIR Home > People > Mark A. Musen

People

Mark A. Musen, M.D., Ph.D.
Professor of Medicine (Biomedical Informatics) Division
Head (BMIR), Co-Director, Biomedical Informatics Training
Program

Contact information
musen@stanford.edu
650-725-3390 (office)
650-725-7944 (fax)
See Med Profile

Ivan Herman

My Work at W3C | Contact information | Short CV | Upcoming trips | Public presentations

My Work at W3C

I am Semantic Web Activity Lead, that is my main work at W3C. I am member of IW3C2 (International World Wide Web Conference Committee) (the committee coordinating the yearly WWW conference series), serving as a liaison for W3C, and of SWSA (Semantic Web Science Association), the committee responsible for the International Semantic Web Conferences series.

As part of my work, I also participate in lots of outreach activities, and I regularly make presentations, tutorials, etc. You can consult my list of presentations for further details.

Contact information

Email:
ivan@w3.org
(sha1sum: 5c80032d5f6012aa1775ea2f63e1676bafd5e80b)

Postal address:
C/o Centre for Mathematics and Computer Sciences (CWI)
Science Park 123, P.O. Box 94079, 1090 GB Amsterdam, The Netherlands (see map).

Phone numbers:
phone: +31-20-5924163
mobile: +31-64-1044153

Suppose we all agreed on how to represent information about a person. Suppose that we all used that to mark up our pages. Now we could search all the information about all the people in a uniform manner.

This is the basis of microformats and schema.org

Interlingua

STANFORD CENTER FOR BIOMEDICAL INFORMATICS RESEARCH
FORMERLY STANFORD MEDICAL INFORMATICS

HOME | ABOUT US | PEOPLE | EVENTS | PUBLICATIONS | PROJECTS | AFFILIATE | EDUCATION | RESEARCH LABS | NATIONAL RESOURCES | COMMUNITY | CONTACT US

name
position
work phone
email
photo

Mark A. Musen, M.D., Ph.D.
Professor of Medicine (Biomedical Informatics) Division
Head (BMIR), Co-Director, Biomedical Informatics Training
Program

Contact Information
musen@stanford.edu
650-725-3390 (office)
650-725-7944 (fax)
See Med Profile

Ivan Herman

My Work at W3C | Contact information | Short CV | Upcoming trips | Public presentations

My Work at W3C

I am Semantic Web Activity Lead, that is my main work at W3C. I am member of IW3C2 (International World Wide Web Conference Committee) (the committee coordinating the yearly WWW conference series), serving as a liaison for W3C, and of SWSA (Semantic Web Science Association), the committee responsible for the International Semantic Web Conferences series.

As part of my work, I also participate in lots of outreach activities, and I regularly make presentations, tutorials, etc. You can consult my list of presentations for further details.

Contact information

Email:
ivan@w3.org
(sha1sum: 5c0032d5f6012aa1775ea2f63e1676bfd5e30b)

Postal address:
C/o Centre for Mathematics and Computer Sciences (CWI)
Science Park 123, P.O. Box 94079, 1090 GB Amsterdam, The Netherlands (see map).

Phone numbers:
phone: +31-20-5924163
mobile: +31-6-1044153

full name
title
office phone
email
image

What if there are competing formats?
Easy – get the other guy to change!!

... or have a way to talk about how they relate.

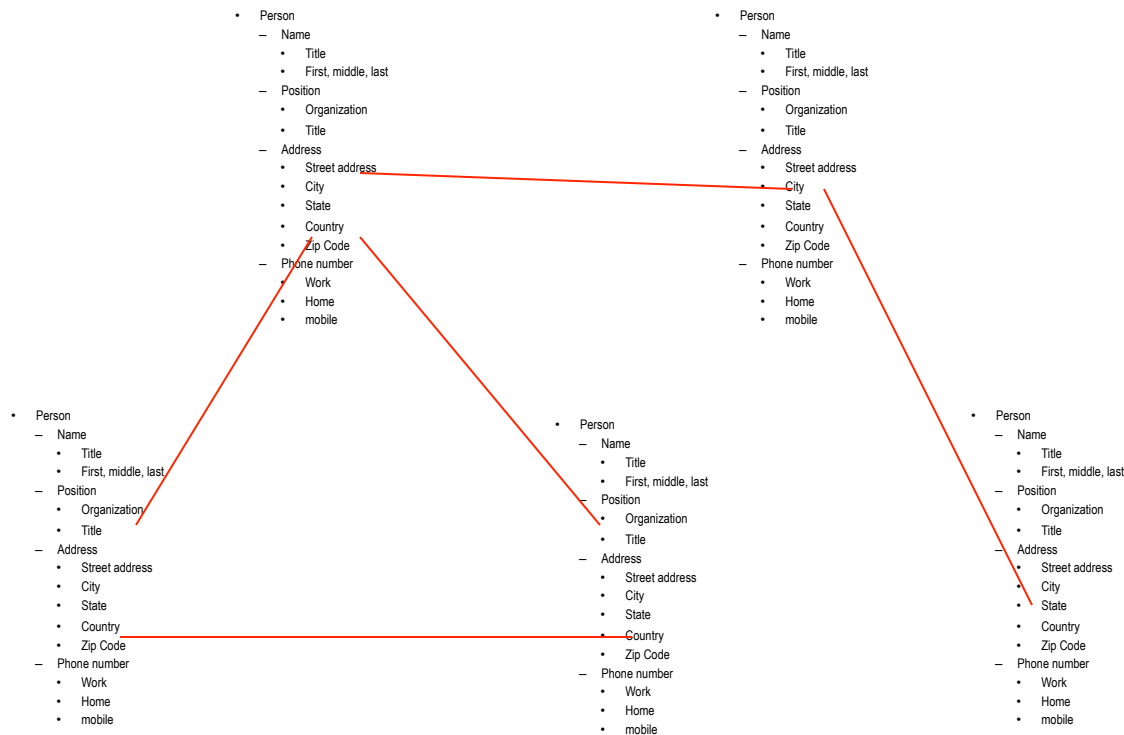
What is an “Ontology”?

- Person
 - Name
 - Title
 - First, middle, last
 - Position
 - Organization
 - Title
 - Address
 - Street address
 - City
 - State
 - Country
 - Zip Code
 - Phone number
 - Work
 - Home
 - mobile
- Employee
 - Name
 - Employee number
 - First, middle, last
 - Job
 - Division
 - Job Description
 - Address
 - Street address
 - City
 - County
 - Country
 - Post code
 - Phone number
 - Work
 - Home
 - mobile

- Sharable, modular piece of metadata.
 - Sharable = referenceable from other places
 - Modular means self-contained
 - Metadata describes other data

Web of metadata

Ontologies refer to one another just as web pages do



Semantic Agents on the web



- Agents provide services for people
- Agents read data from pages
- Agents merge data from multiple pages
- Examples:
 - Mashups
 - Auto-fill forms
 - Facebook “like” or “share” buttons

Life in the Data Cathedral

- Everything follows strict rules
- Only civilized things allowed in
- Everything under control



Living in the data wilderness

- Caveat emptor
- Data is in multiple formats, multiple schemas, unknown quality
- The web grew as a wilderness



Two Strategies for How The Web was Won

1. First, get everything into the same form, get everyone to agree, get them all into the Cathedral. Then, data exchange is easy.



2. Learn to live in the wilderness, making use of the valuable things there, while dealing with the dangerous ones.

