The Semantic Web and the Role of Information & Knowledge Professionals

Marcia Lei Zeng
Kent State University

The Semantic Web and Libraries, Archives and Museums - - trends for the future

2008-09-17 Stockholm

The Questions:

The Semantic Web

- What is it?
- Why is it important?

There are many ways to answer these questions.

Let’s look at the trends from three perspectives …

1. The Web:

   From the Web of Documents

   to the Web of Data

2. Our Work:

   From [bibliographic and authority] control

   to knowledge organization, resource discovery, and delivery

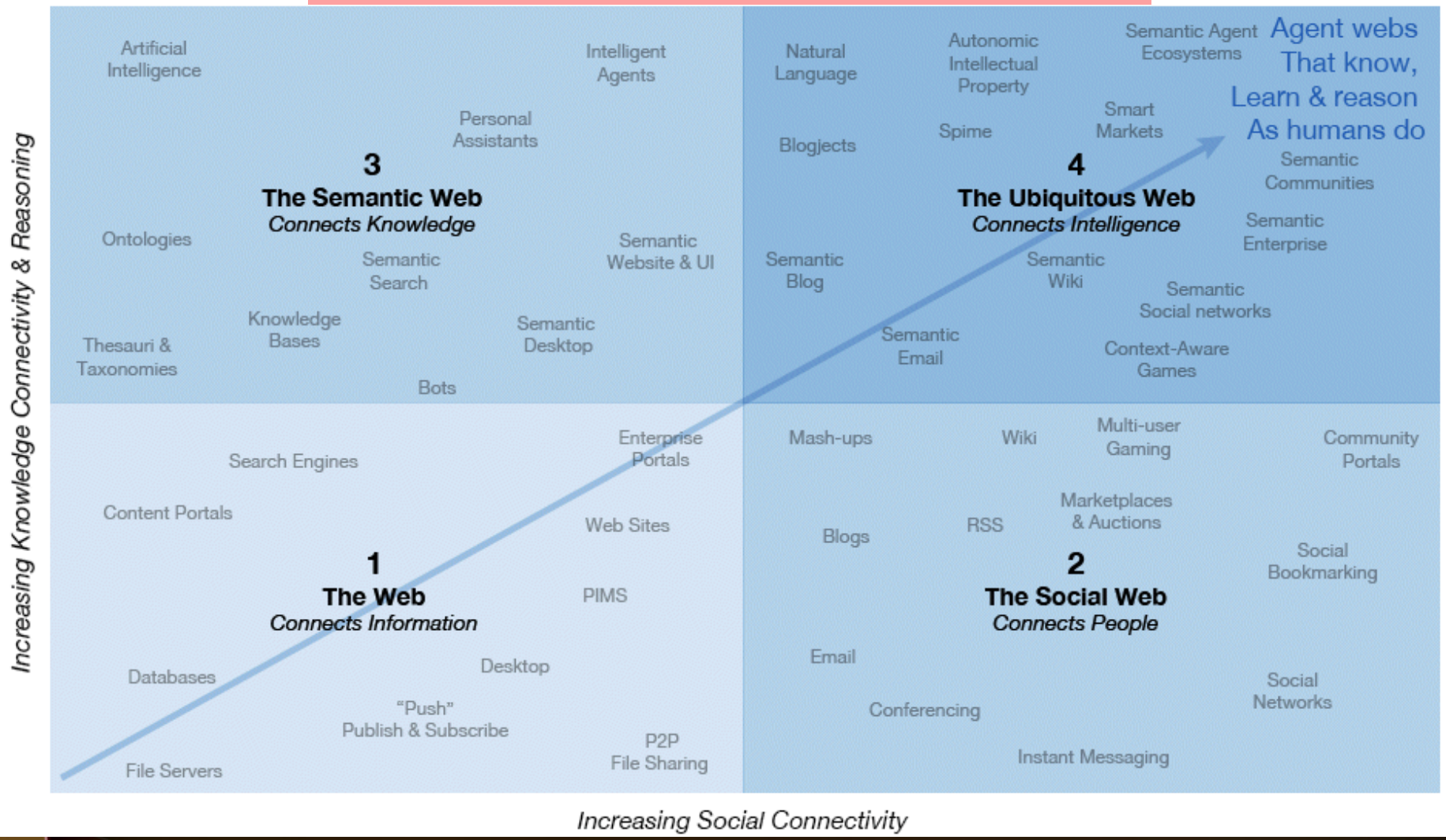
3. Our Data:

   From machine-readable

   to machine-processable and become Web resources

1. The Web:
From
the Web of Documents
to
the Web of Data

What is the Evolution of the Internet to 2020?



Web 1.0: connecting information and getting on the net.
Web 2.0: connecting people — putting the "I" into user interface, and the "we" into Webs of social participation.
Web 3.0 (starting now): representing meanings, connecting knowledge, and putting these to work in ways that make our experience of internet more relevant, useful, and enjoyable.
Web 4.0 (will come later): connecting intelligences in a ubiquitous Web where both people and things reason and communicate together.

Source: Nova Spivak, Radar Networks; John Breslin, DERI; & Mills Davis, Project10X, 2007, 2008 Copyright MILLS•DAVIS

- “The Semantic Web isn't inherently complex. The Semantic Web language, at its heart, is very, very simple. It's just about the relationships between things.”

-- Q&A with Tim Berners-Lee

BusinessWeek Special Report April 9, 2007

The Current Web
-- *"web of documents"*

Resources:

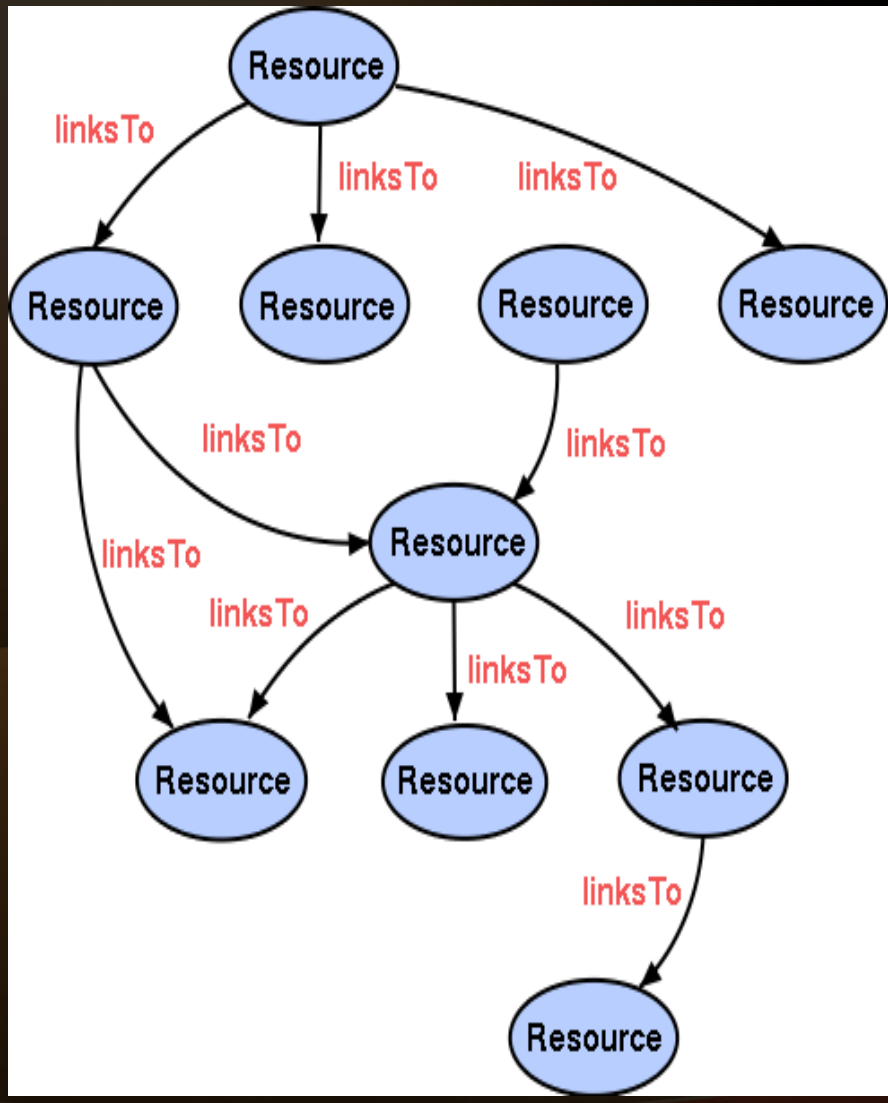
- **identified by** URIs**;**
- **relationship between the** current **and the** target **resource is defined by a** hyperlink**;**
- **resource** types **are not specified**.

Links:

- *href, src*, **… limited, non-descriptive;**
- links **are not named;**
- role of a link **is deduced by the** human **reader.**

The language**: html**

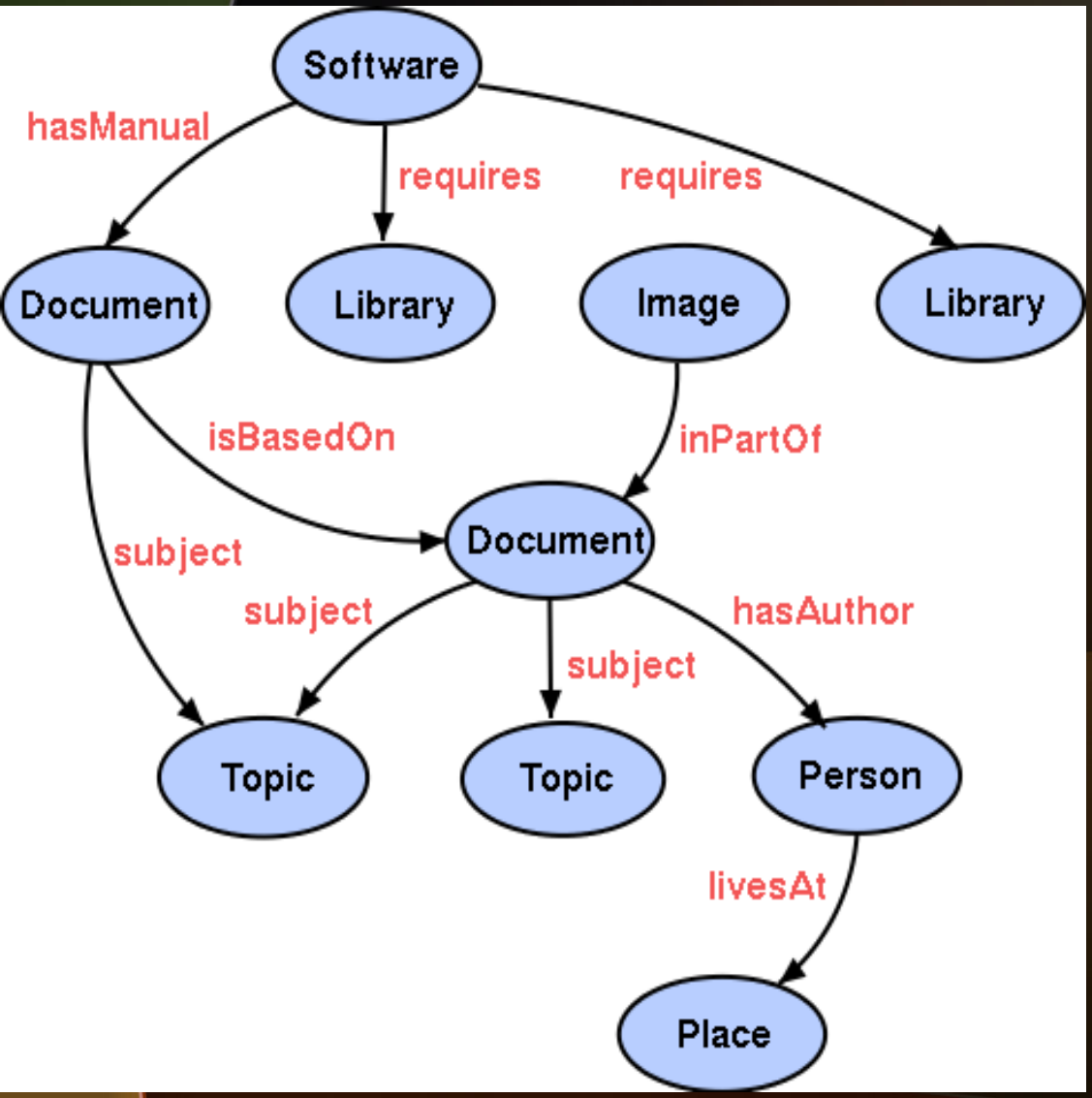
- human-understandable **content is delivered and hyperlinked;**
- **very little** machine understandable **information - significance of the links only evident from the context around the anchor.**



Source: W3C Semantic Web Activity - http://www.w3.org/2001/sw/

The Semantic Web

"web of data"



Resources:

- **Globally Identified by URIs; types are specified**

Links:

- **Identified by URIs; descriptive**

Machine:

- **More processable information is available**

Source: W3C Semantic Web Activity - http://www.w3.org/2001/sw/

- **The vision of the Semantic Web is to extend principles of the Web** from documents to data**.**

http://www.w3.org/2001/sw/SW-FAQ

The semantic Web is a web of machine understandable content.

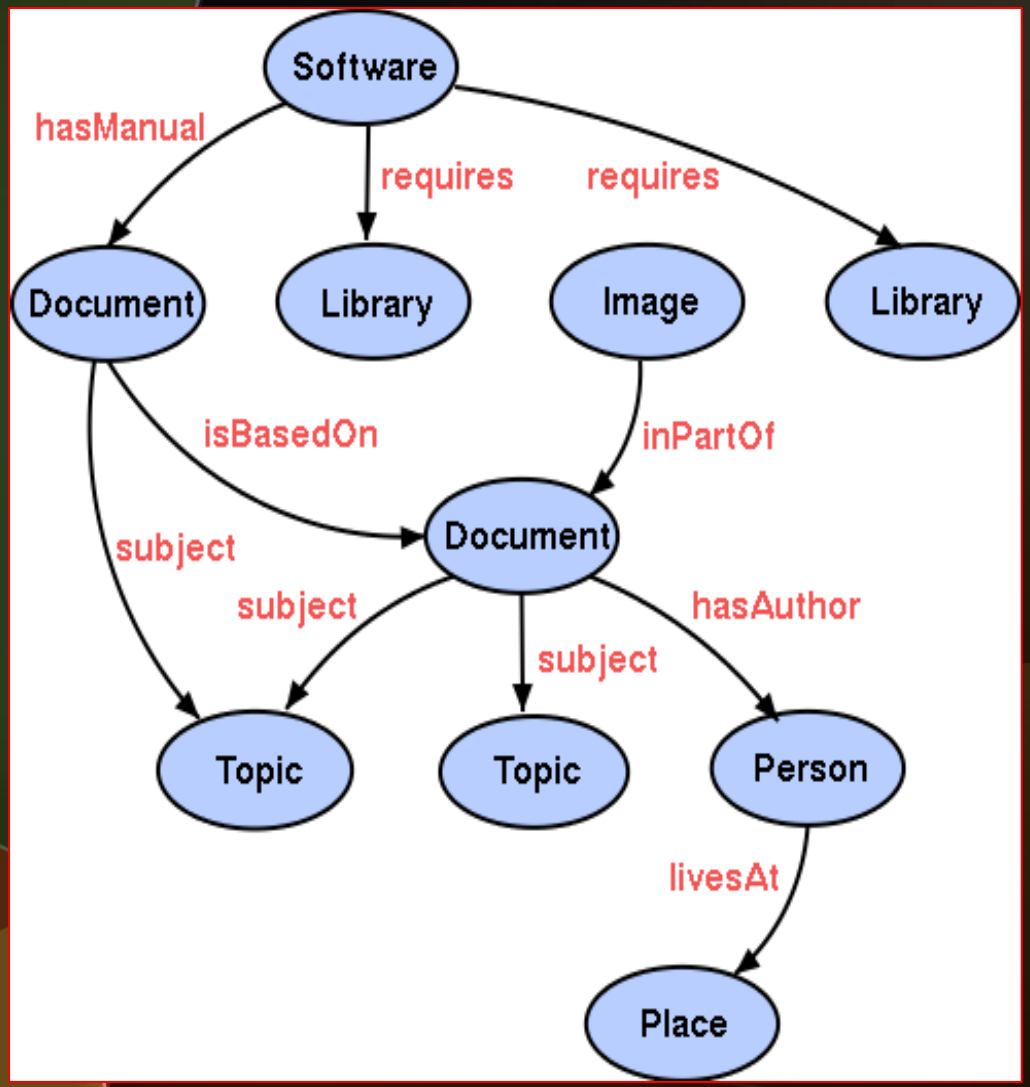
- “Today, the Web is quite effective at helping us to publish and discover documents, but the individual information elements within those documents … cannot be handled directly as data.”
  - managing
  - integrating
  - analyzing
  - ...

Timothy Berners-Lee, 2007-03-01
http://dig.csail.mit.edu/2007/03/01-ushouse-future-of-the-web.html

- **“Today you can see the data with your browser, but can't get other computer programs to manipulate or analyze it without going through a lot of manual effort yourself.”**
- **“As this problem is solved, we can expect that Web as a whole to** look more like a large database or spreadsheet**, rather than just a set of linked documents.”**

Timothy Berners-Lee, 2007-03-01
http://dig.csail.mit.edu/2007/03/01-ushouse-future-of-the-web.html

Revisit the graph …



- **The Semantic Web**
  - **Web of** data **(resources)**
  - **Relationships can be established** between *any* two resources.
  - **The relationship (i.e, the link) itself is** *named***.**
    - **The definition of those relations allow for a better and** automatic **interchange of data.**
  - **The language: RDF**
    - **RDF allows** machine-understandable **content to be distributed across the web and 'hyperlinked'.**

And compare again …

- Current Web:
  - Web of documents
  - the hyperlinks define a relationship between the current page and the target.
  - the link used by a human on the (traditional) Web is not named; and its role is deduced by the human reader.
  - The language: html
    - HTML allows human-understandable content to be distributed across the web and 'hyperlinked'.

- The Semantic Web
  - Web of data (resources)
  - relationships can be established between *any* two resources; there is no notion of "current" page.
  - the relationship (i.e, the link) itself is *named*
    - The definition of those relations allow for a better and automatic interchange of data.
  - The language: RDF
    - RDF allows machine-understandable content to be distributed across the web and 'hyperlinked'.

Based on: Alistair Miles, Taxonomies and the Semantic Web, CISTRANA Workshop 02/05

Examples of application areas of the Semantic Web technologies

- **Merge**
- **Analyse**
- **Repurpose**
- **…**

- **Used in *data integration*, whereby data in various locations and various formats can be integrated in one, seamless application;**

- **Used in *resource discovery and classification,* to provide better, domain-specific search engine capabilities;**

- **Used in *cataloging* for describing the content and content relationships available at a particular Web site, page, or digital library;**
- **…**

-- W3C: How would you define the main goals of the Semantic Web?

Example: Explore Data on Locations of Marine Animals and Plants

OBIS
OCEAN BIOGEOGRAPHIC INFORMATION SYSTEM

Name Search | Search »

Advanced Search | Browse by taxonomic groups

SPECIES INFORMATION

Back to previous page

Carcharodon carcharias

Name verified: Catalogue of Life; FishBase; Compagno, Leonard J.V.
Organism type: a fish FishBase

Data Extent Map (from OBIS Australia/ C Square Mapper)

For more options, view full-size map

You searched for the distribution of *Carcharodon carcharias*
Your search returned 197 records

View all 197 results as .TXT (download) **HTML** (download)

External Mappers

KGS »

The KGSMapper is a dynamic habitat analysis and inference tool fro
Biogeoinformatics of Hexacorals. Click here for for more informatio
about the KGS mapper.

Note, if your request contains < 1,000 records, taxon names will be
transmitted to the KGS mapper along with the locations; if > 1,000
records, only locations will be sent.

ACON »

The ACON mapper is a dynamic multi-species mapper from Bedford
Institute of Oceanography. Use for species with less than 100,000
records. Click here for more information about the ACON mapper.

The data are from the following sources:

- Australian Museum (OBIS Australia)
- Marine and Coastal Management - Linefish Dataset (Second Semester of 1985) (AfrOBIS)
- SAM Ichthyology (OBIS Australia)
- Bureau of Rural Sciences National commercial fisheries half-degree data set 2000-2002 (OBIS Australia)
- Marine and Coastal Management - Linefish Dataset (Second Semester of 1989) (AfrOBIS)
- Marine and Coastal Management - Linefish Dataset (Second Semester of 1986) (AfrOBIS)
- MV Ichthyology (OBIS Australia)
- FishBase DiGIR Provider - Philippine Server
- iziko South African Museum - Shark Collection (AfrOBIS)
- Canadian Museum of Nature - Fish Collection (OBIS Canada)
- IndOBIS, Indian Ocean Node of OBIS (IndOBIS)

A registry of many kinds of (mapped) data

- So, what are the impacts of the Semantic Web to our profession?
- What are the roles of the information and knowledge professionals?

2. Our Work:

From

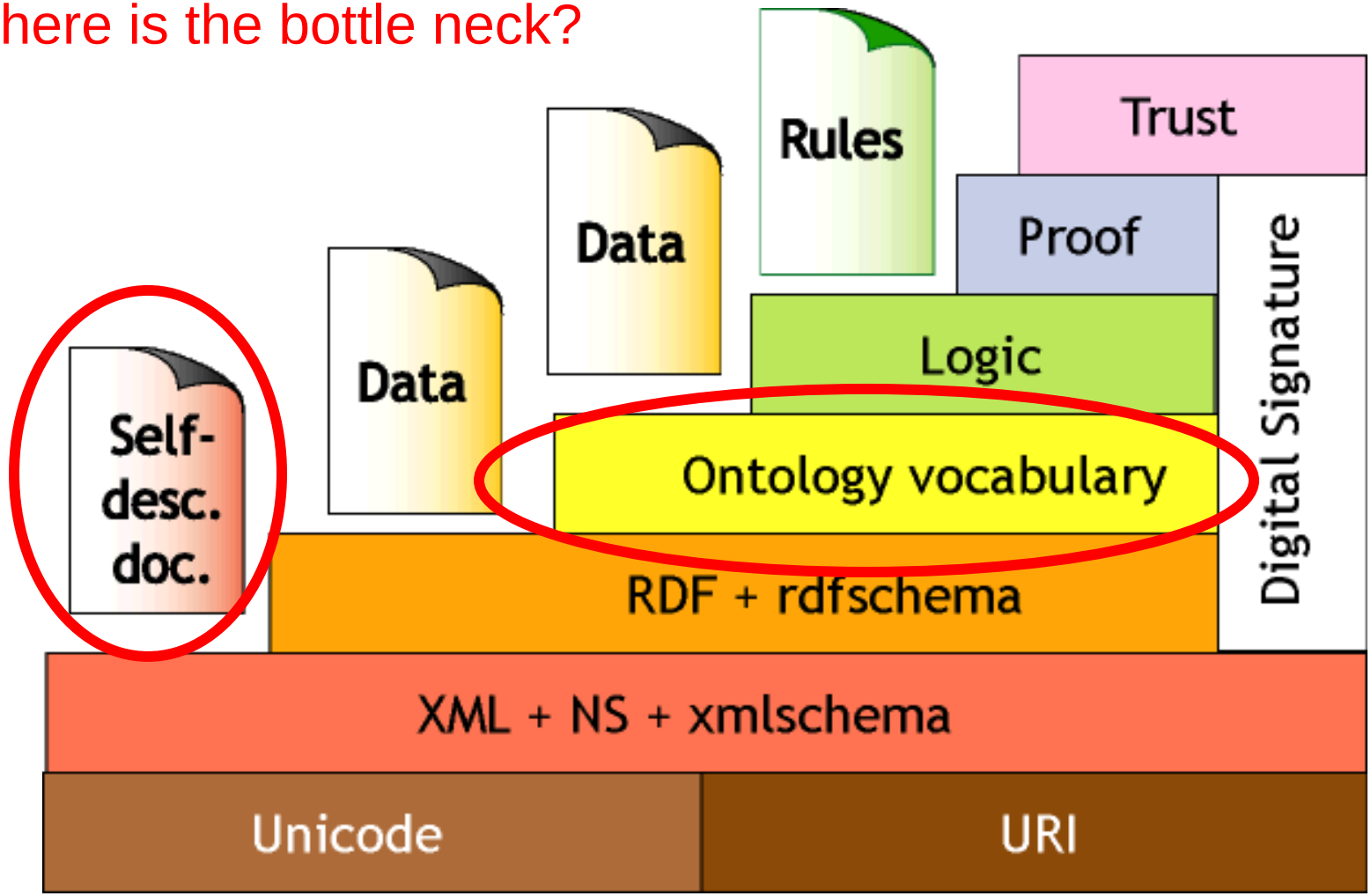
[bibliographic] Control

to

Knowledge Organization and Resource Discovery & Delivery

Semantic Web Architecture

Where is the bottle neck?



Source: Tim Berners-Lee, 2000

http://www.w3.org/2000/Talks/1206-xml2k-tbl

Why do we still need
[standardized / interoperable]
metadata
for this
Web of Data (Resources)?

So much data (resources) on the Web

Table 1.13: The size of the Internet in terabytes.	
Medium	**2002 Terabytes**
Surface Web	167
Deep Web	91,850
Email (originals)	440,606
Instant messaging	274
TOTAL	532,897

Source: How much information 2003

10-100 TB = total collection of Library Of Congress

So much data (resources) on the Web

Resources must be self-descriptive.

However, semantic conflicts could occur at any data communication processes

Semantic Conflicts – Examples (1)

"071210"

- **It may be correctly *transferred* (passed) from one system to another,**
- **but its meaning may be *interpreted* in any number of different ways.**

“071210”

an integer?

a string?

- **a date?**
  - **birth date?**
  - **publication date?**
  - **discovery date?**
  - **creation date?**
  - **restoration date?**
  - **yy-mm-dd (1907 Dec. 10)?**
  - **mm-dd-yy (July 12, 1910)?**
  - **mm-yy-dd (July 1912 10th)?**
  - **dd-mm-yy (7 Dec. 1910)?**
- **a phone number?**
- **an area code?**
- **a hexadecimal number (#071210) for a color?**

If: Start-time= 2001, duration=6

Then: Earliest-date=2001, latest-date=2007 ?

Semantic Conflicts – Examples (2)

"title"

catalog.xsd→ title of a book

journalIndex.xsd → title of a journal

articleIndex.xsd → title of an article

employee.xsd → title of a position

autoInsurance.xsd → ownership certificate

(note: .xsd is the file extension for XML schemas)

Shareable metadata

will ensure

semantic interoperability

Metadata standards:

• for data structures (e.g., Dublin Core Metadata Element Sets)

• for data contents (e.g., AACR2, Cataloging Cultural Objects (CCO), *Describing Archives*: *A Content Standard* (DACS))

• for data values (e.g., Art and Architecture Thesaurus, name authorities, country codes)

[VRA Core 4.0 elements]:

DATE

Subelements:

earliestDate

latestDate

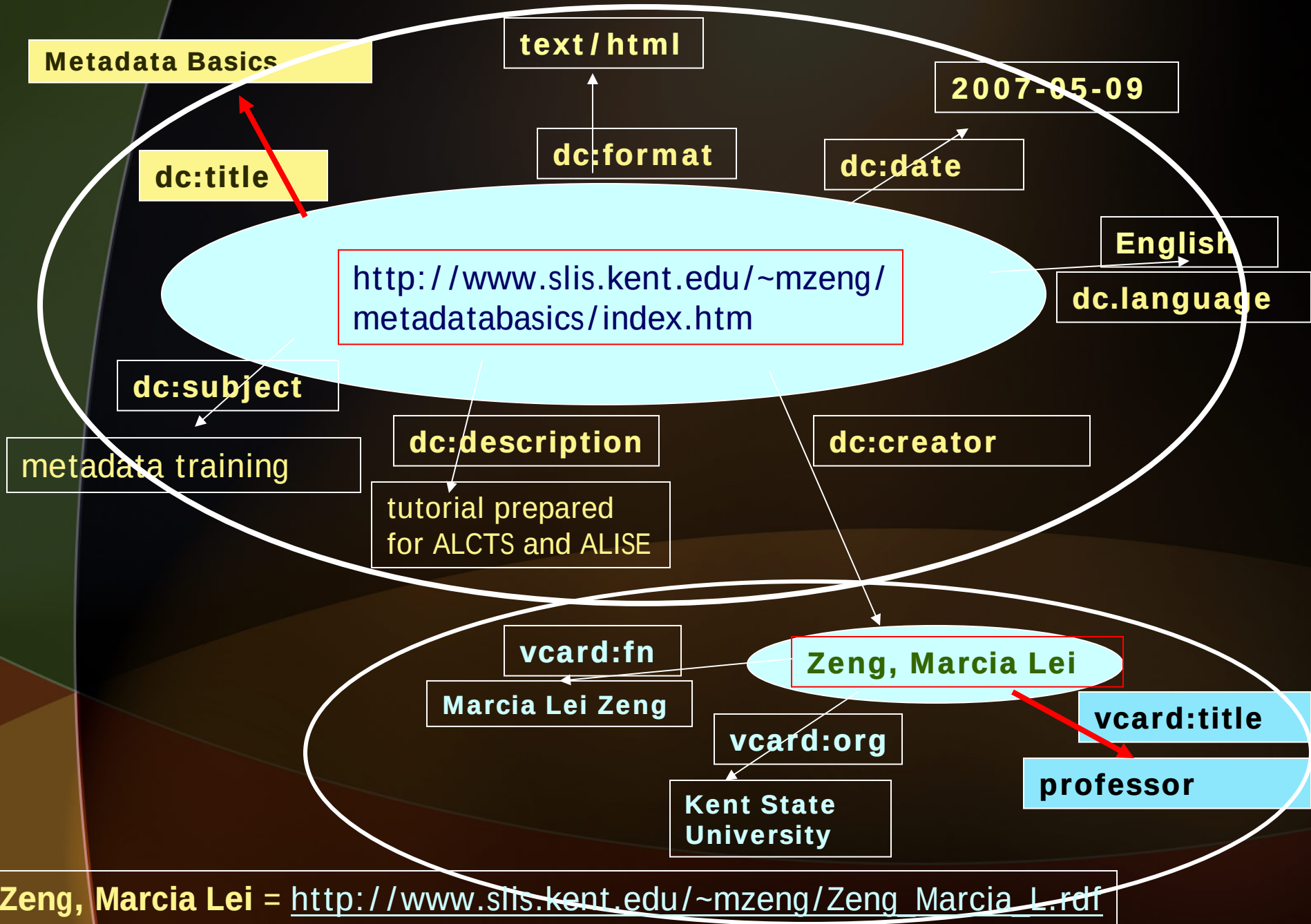
[Tagging example]:

```
<dateSet>

<display>created 1520-1525</display>

<date type="creation" source="Grove
    Dictionary of Art Online"
    href="http://www.groveart.com"
     dataDate="2005-06-08">

  <earliestDate>1520</earliestDate>
  <latestDate>1525</latestDate>
</date>
```

VRA Core 4.0 =Visual Resources Association Core Categories, Version 4

How will this be part of the Semantic Web?

A resource descried with RDF graphs:



```
<?xml version="1.0"?>
<rdf: RDF
    xmlns:rdf= "http://www.w3.org/1999/02/22-rdf-syntax-ns#"
    xmlns:dc= "http://purl.org/dc/elements/1.1/"
    xmlns:vcard = "http://www.w3.org/2001/vcard-rdf/3.0#">

  <rdf:Description rdf:about =
    "http://www.slis.kent.edu/~mzeng/metadatabasics/index.htm">
  <dc:creator rdf:href= "Zeng, Marcia Lei"/>
  <dc:title> Metadata Basics </dc:title>
  <dc:subject> metadata standards; metadata training; metadata value
    space </dc:subject>
  <dc:date scheme="W3CDTF">2007-05-09</dc:date>
  <dc:format>text/html</dc:format>
  <dc:language>en</dc:language>
  <dc:description> tutorial prepared for ALCTS and ALISE
    </dc:description>
  </rdf:Description>
  <rdf:Description ID= "Zeng, Marcia Lei">
    <vcard:fn>Marcia Lei Zeng </vcard:fn>
    <vcard:title>professor</vcard:title>
   <vcard:org>Kent State University </vcard:org>
  </rdf:Description>
</rdf:RDF>
```

Shareable metadata

will ensure

semantic interoperability

which need

Knowledge Organization Systems (KOS)

to enable

resource self-description
and auto-reasoning

KOS enable resources' self-description and auto-reasoning

We know the facts that:

- *Anders Cato* **is Head of Cataloguing, Kungl. biblioteket (isA)**
- **Cataloging Department is a unit at the Kungl. Biblioteket (instantOf)**
- **Kungl. Biblioteket is the National Library of Sweden (hasName)**

- **Anders Cato is the chair of Cataloguing Section**
- **Cataloguing Section is a section of IFLA**

- **Chairs of IFLA sections chaired standing committee meetings at Quebec City**
- **Quebec City is a part of Canada**

National Library of Sweden Staff
. Head of Cataloging
. . Anders Cato

IFLA
. Sections
. . Cataloging Section
… Standing Committees

World
. Canada
. . Quebec City

We can have many reasonable results:

A staff member of the National Library of Sweden chairs a IFLA Section.

The head of Cataloging at the National Library of Sweden chairs IFLA Cataloguing Section.

Anders Cato chaired standing committee meetings in IFLA in Canada.

Example: SWED (Semantic Web Environmental Directory) -- a prototype of a kind of directory of environmental **organizations** and **projects**

http://www.swed.org.uk/swed/index.html

[SWED Ontology components slide]

Based on: Alistair Miles, Taxonomies and the Semantic Web, CISTRANA Workshop 02/05,

SWED is a portal that harvests it's content from ***a semantic web*** (i.e., ***web of machine-understandable content)***

Where are the differences? (1)

authority data

become

discovery tools

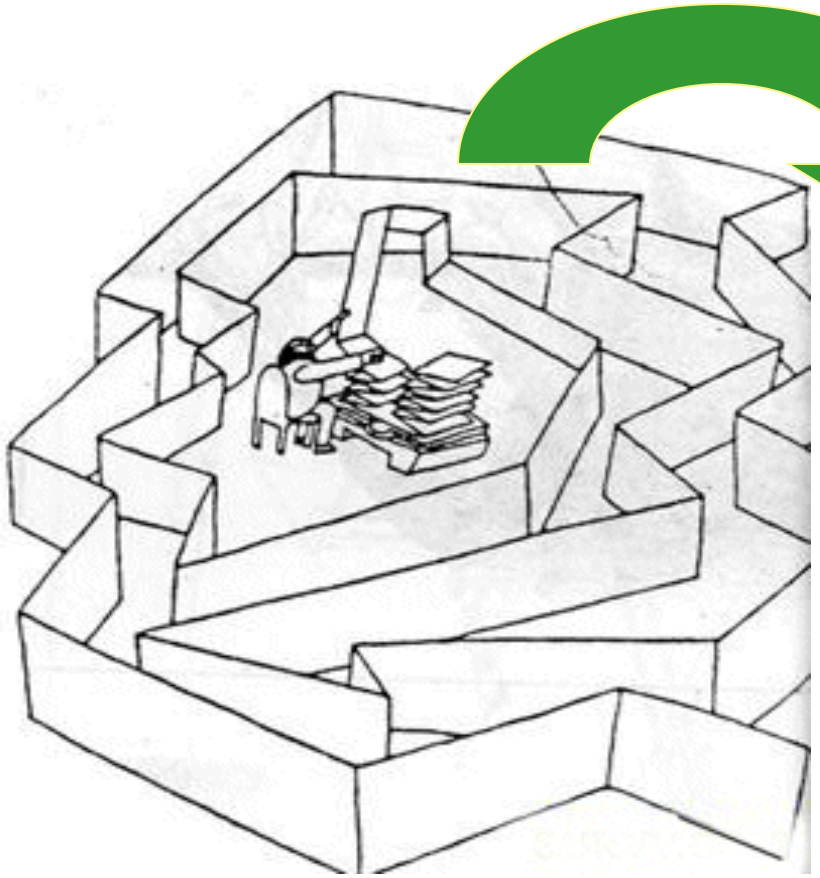
Examples: LC name authority record for W. Shakespeare

WorldCat Identity graphic display based on this data

From http://orlabs.oclc.org/Identities/lccn-n78-95332, choose 2004-2009 bar

Where are the differences? (2)

systems used for
bibliographic control **and**
collection management



transfer into

semantic tools for Web resources

Tree of Life web project

http://www.tolweb.org/tree/

core Caryophyllales - Mic

File Edit View Favorites Tools

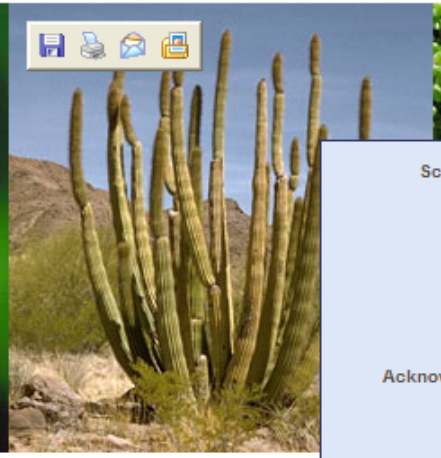
Back

Address http://tolweb.org/core_Cary

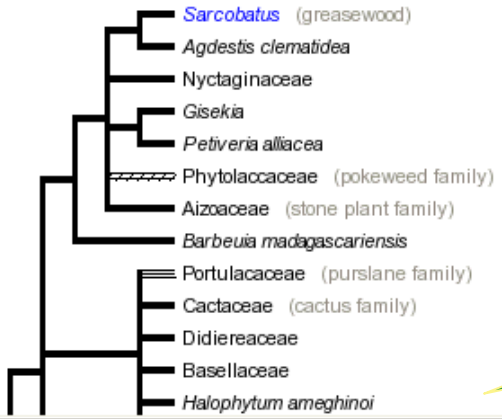
Go Links

Google Search 237 blocked Check AutoLink AutoFill Options

Pam Soltis, Doug Soltis, and Monica Arakaki



Scientific Name	Stenocerus
Location	Organ Pipe Cactus National Monument (Arizona, US)
Comments	Organ pipe cactus (Cactaceae)
Creator	Photograph by G. Dallas and Margaret Hanna
Acknowledgements	courtesy CalPhotos
Copyright	© 2001 California Academy of Sciences



Metadata

Taxonomy

articles & notes
treehouses
collections
people
options

Explore Other Groups

other Caryophyllales
containing groups
subgroups
random page

http://tolweb.org/core_Caryophyllales/20965#titlefigcaption

Internet

start core Caryo... My Docum... Microsoft P... Jasc Paint ... EN 11:35 PM

Popular Groups on the Tree of Life

Click on the names below to go to the page for that group.

- Eubacteria
- **Eukaryotes**
  - **Animals**
    - Echinoderms (sea urchins, starfish, sea cucumbers, etc)
    - **Vertebrates (fish etc.)**
      - **Terrestrial Vertebrates**
        - Frogs
        - Salamanders
        - Turtles
        - **Dinosaurs**
          - Modern Birds
        - Mammals
      - Teleost fish
    - Cnidaria (jellyfish, anemones, corals, etc.)
    - Annelida (segmented worms)
    - Cephalopoda (octopods, squids, etc.)
    - **Arthropoda**
      - **Insects**
        - Dragonflies and Damselflies
        - Lice
        - True Bugs
        - Beetles
        - Wasps, Bees, and Ants
        - Flies
        - Butterflies and Moths
        - Crickets, Katydids, and Grasshoppers
      - **Arachnids**
        - Spiders
        - Mites
        - Scorpions
  - Fungi
  - **Green Plants**
    - Ferns
    - Flowering Plants

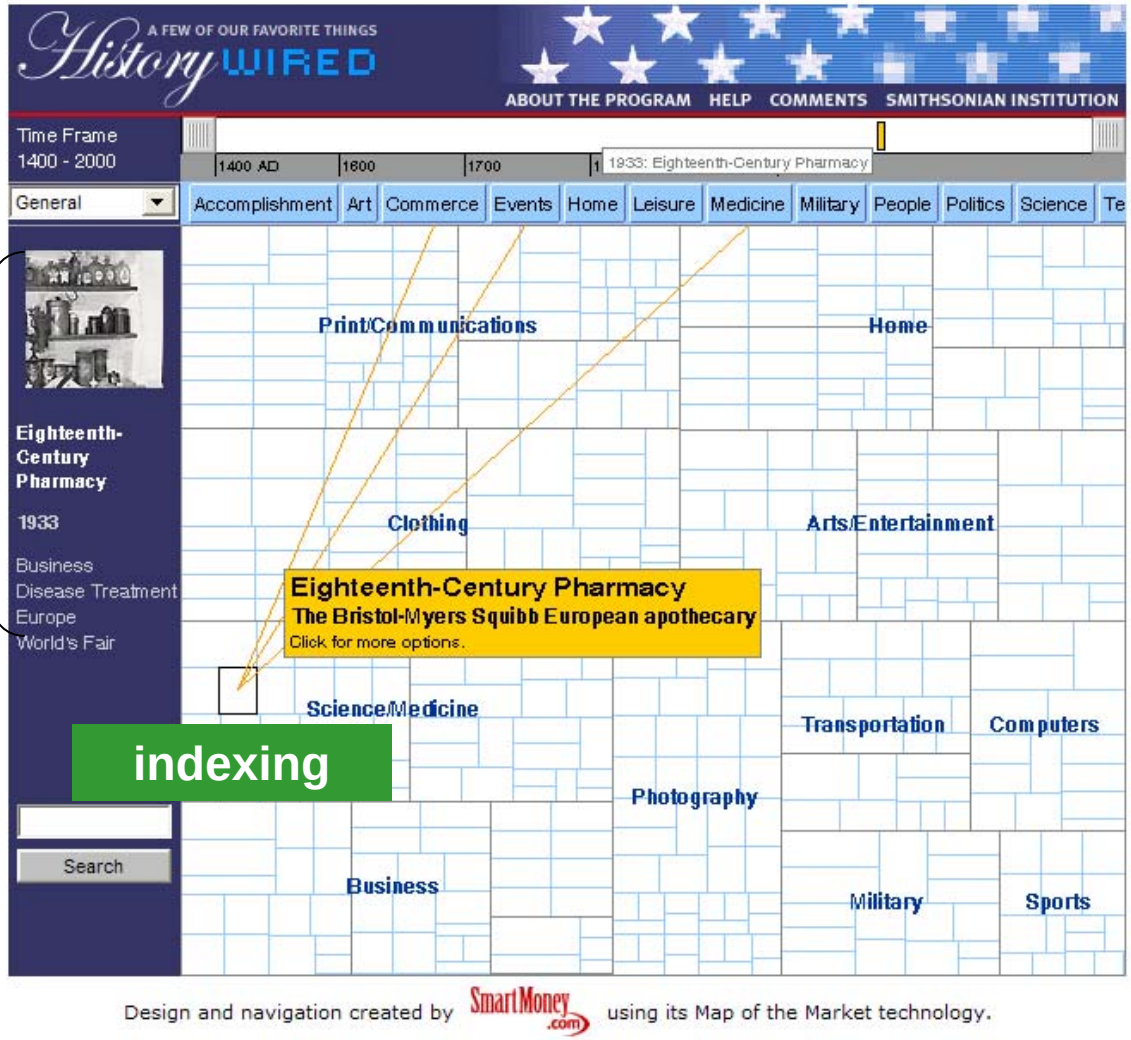
If the group you are looking for is not on this page, try doing a Search, or check out the page on Finding Things in the Tree of Life for a list of all the pages that can help you locate your favorite group.

Browse the Tree of Life

- Browse the ToL
- Root of the Tree
- Popular Pages
- Sample Pages
- Recent Additions
- Random Page
- Treehouses
- Biographies

Classification

History Wired: A few of our favorite things.”
http://historywired.si.edu/



time

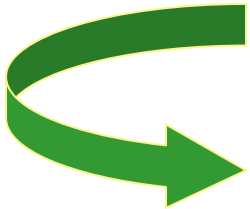
topic

metadata

indexing

category

bibliographic control



re-focuses on

resource discovery & delivering

Graphic based on http://www.oclc.org/membership/escan/future/implications.htm

3. Our Data:
From
Machine-Readable
to
Machine-Processable

New Requirements for KOS

- **Making KOS machine-processable (machine-understandable)**
  - **concern previously belonged to the domain of researchers in computer science and W3C pioneers**
  - **now in library and information sciences**

Recommendation of LC WG on Future of Bibliographic Control (Nov. 13, 2007)

- "Optimize LCSH for Use & Re-use"
  - "de-coupling [LCSH] subject strings"
  - "making data (including subject authority data) directed to Web services in order to make them machine-processable"
- All traditional KOS face such an issue which needs immediate action.

"...making data (including subject authority data) directed to Web services in order to make them machine-processable"

Requires

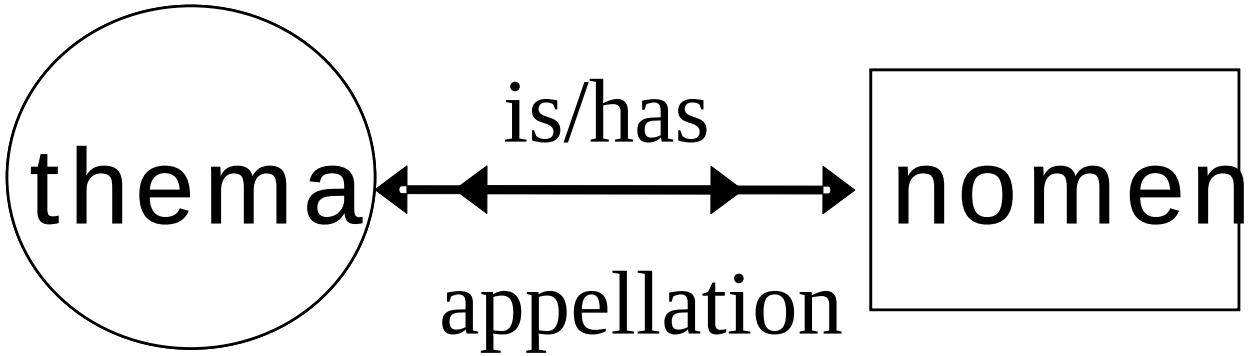
- **Expressive data structures**
- **Vocabularies encoded for the Web**
- **Access mechanisms for search and retrieval**
- **Content are accessible via URI**
- **Open protocols and standards**

Based on Vizine-Goetz, 2008

"...making data (including subject authority data) directed to Web services in order to make them machine-processable"

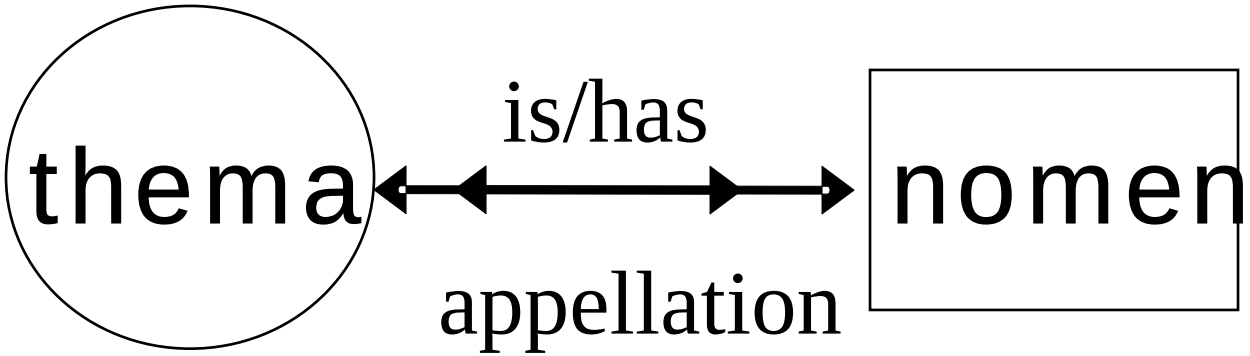
- **Two critical components**
  - **A conceptual model**
  - **An encoding mechanism**

- **FRSAR Model**



- **A key point here is to separate a [stuff] from what it is called, referred to, or addressed as.**

Understanding the Conceptual model of aboutness (2)



Thema: anything that may be the subject of a work

Nomen: any alpha numeric, sound, visual etc., symbol or combination of symbols by which a thema is known, referred to, or addressed as.

- [records examples:
  - Art and Architecture Thesaurus
  - Getty Thesaurus of Georagraphic Names (TGN)
  - Union List of Artist Names

ID: 7009861 → nomen type="ID"

Record Type: administrative

Stockholm (inhabited place) → place as thema

Coordinates:
Lat: 59 20 00 N *degrees minutes* Lat: 59.3333 *decimal degrees* → nomen form="coordinators"
Long: 018 05 00 E *degrees minutes* Long: 18.0833 *decimal degrees*

Note: Built upon numerous islands as well as on the mainland, primarily by the Swedish ruler Birger Jarl in 13th cen.; rapidly developed as member of Hanseatic League; extensively rebuilt in 1634; destroyed by fires & rebuilt in 18th cen., expanded in 19th cen.

nomen –nomen relations

Names:
Stockholm (**preferred**, C ,V,N)
Estocolmo (H ,V,N) → nomen status="historical"
→ nomen script ="vernacular"

Hierarchical Position:
World (facet)
.... Europe (continent)
..... Sweden (nation)
..... Stockholm (county)
..... Stockholm (inhabited place)

thema –thema relations

Place Types:
inhabited place (**preferred**, C) ..... construction here was documented by 1252

city (C)
national capital (C) ..... established by Queen Christina in 1634
county seat (C)
port (C)
market center (C) ..... grew rapidly as the result of trade agreements with German city of Lübeck

cultural center (C)
industrial center (C)
educational center (C)

→ **thema – type (place-specific)**

Encoding the vocabularies

- **Machine-readable:**
  - **System-specific format**
    - **MARC authority format**
  - **General Web format**
    - **MARCXML**

- **Machine-understandable, processable:**
  - **For the Semantic Web**
    - **SKOS (for KOS)**
    - **OWL (for ontologies)**

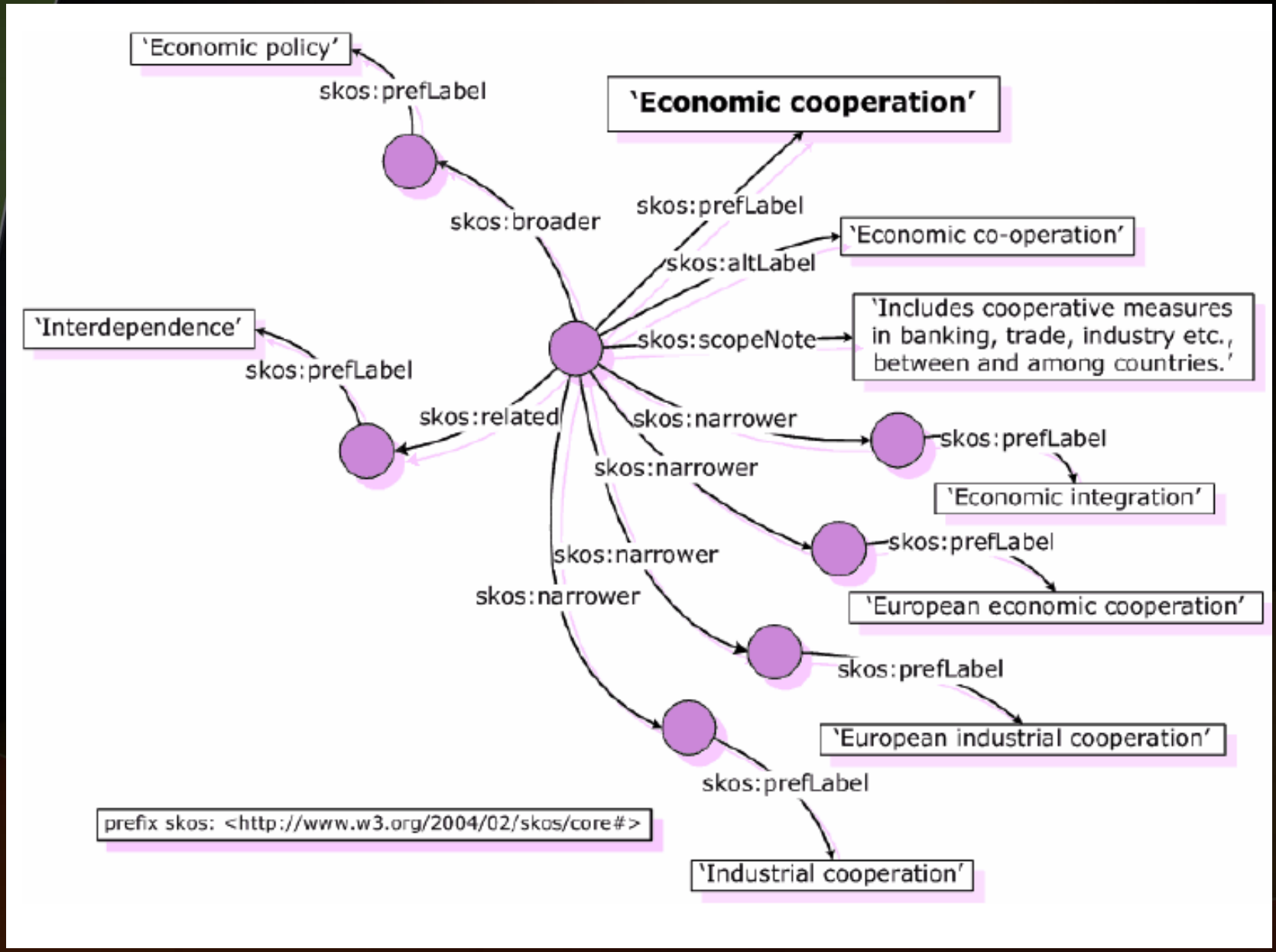
SKOS

SKOS = Simple Knowledge Organisation System

- **a common data model for sharing and linking knowledge organization systems (KOS) via the Semantic Web.**
  - **KOS examples: thesauri, taxonomies, classification schemes, subject heading systems … …**

SKOS Reference, http://www.w3.org/TR/skos-reference/

The example is expressed as an RDF graph using SKOS:



Source: Miles' presentation at UK ISKO Meeting, July 21, 2008, London

An RDF/XML serialization of the RDF description of the 'Economic cooperation' concept:

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:skos="http://www.w3.org/2004/02/skos/core#">

    <skos:Concept rdf:about="http://www.ukat.org.uk/thesaurus/concept/1750">
        <skos:prefLabel>Economic cooperation</skos:prefLabel>
        <skos:altLabel>Economic co-operation</skos:altLabel>
        <skos:scopeNote>Includes cooperative measures in banking, trade, industry etc.,
          between and among countries.</skos:scopeNote>
        <skos:broader rdf:resource="http://www.ukat.org.uk/thesaurus/concept/4382"/>
        <skos:narrower rdf:resource="http://www.ukat.org.uk/thesaurus/concept/2108"/>
        <skos:narrower rdf:resource="http://www.ukat.org.uk/thesaurus/concept/9505"/>
        <skos:narrower rdf:resource="http://www.ukat.org.uk/thesaurus/concept/15053"/>
        <skos:narrower rdf:resource="http://www.ukat.org.uk/thesaurus/concept/18987"/>
        <skos:related rdf:resource="http://www.ukat.org.uk/thesaurus/concept/3250"/>
        <skos:inScheme rdf:resource="http://www.ukat.org.uk/thesaurus"/>
    </skos:Concept>

</rdf:RDF>
```

http://www.w3.org/TR/2005/WD-swbp-thesaurus-pubguide-20050517/

SKOS current version

- **SKOS Reference (W3C Working Draft 9 June 2008)**
  - **Replaces SKOS Core Guide (i.e. *SKOS 2005*)**
- **SKOS Primer (W3C Working Draft 21 February 2008)**
- **SKOS RDF Schema (June 3rd 2008 Edition)**
- **SKOS eXtension for Labels (XL) RDF Schemas (June 3rd 2008 Edition)**

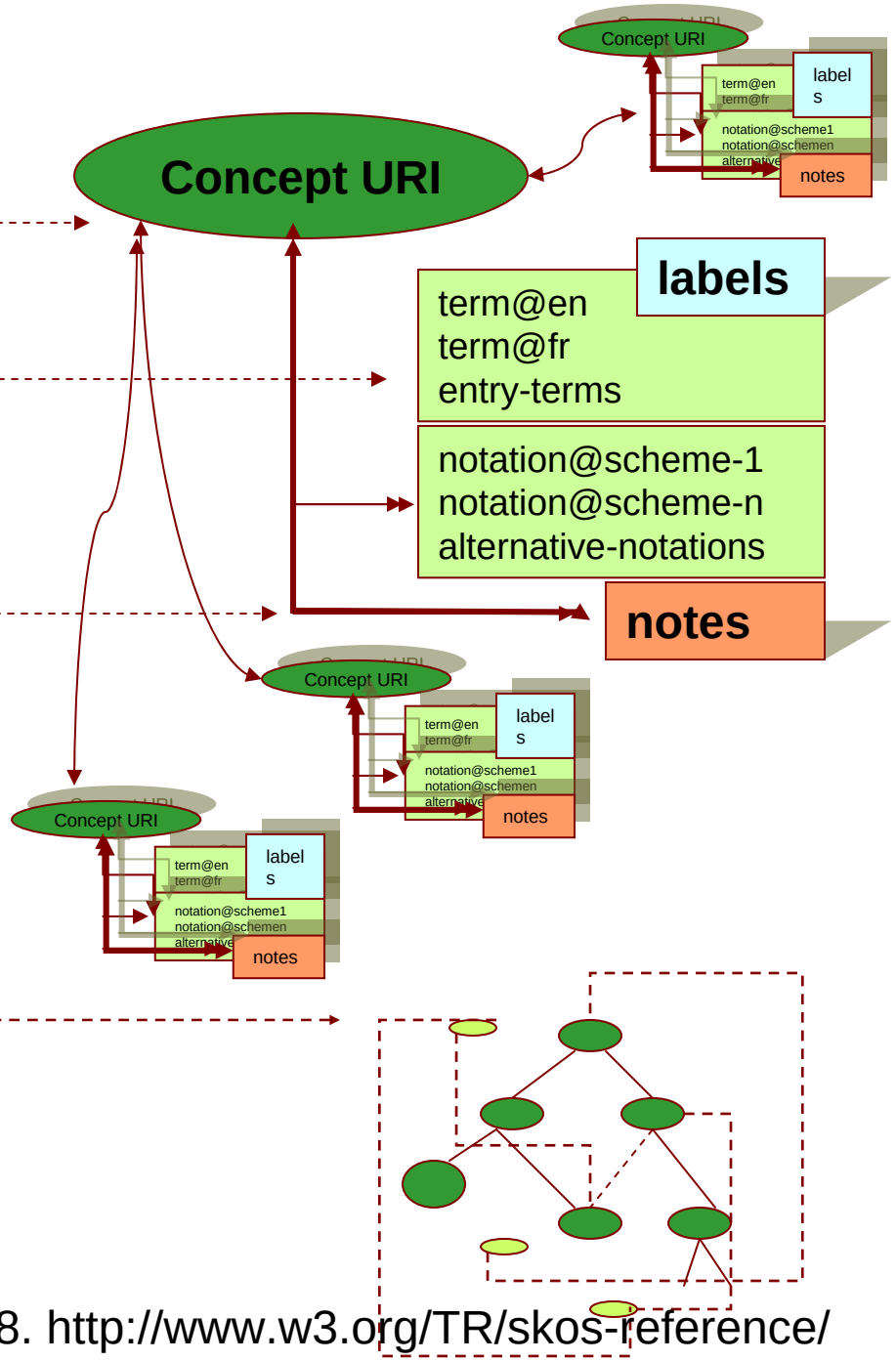
SKOS Synopsis (1)

Using SKOS,

- concepts **can be identified using URIs**
- labeled **with lexical strings in one or more natural languages, and assigned** notations **(lexical codes)**
- documented **with various types of note**
- linked **to other concepts**

  and

  organized into informal hierarchies and association networks
- **(to be continued →)**



Drawn based on *SKOS Reference*, June 2008. http://www.w3.org/TR/skos-reference/

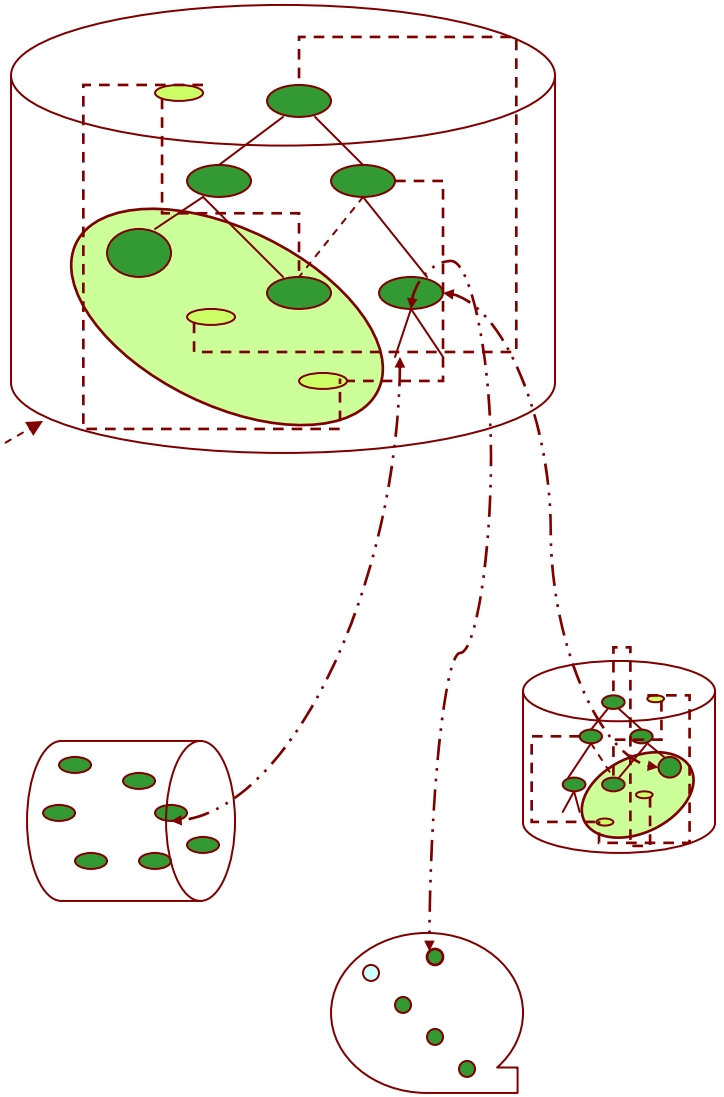
SKOS Synopsis (2)

(continued from previous page)

•aggregated into **concept schemes**,

• [grouped into labeled and/or ordered **collections**,]

• and **mapped** to concepts in other schemes.

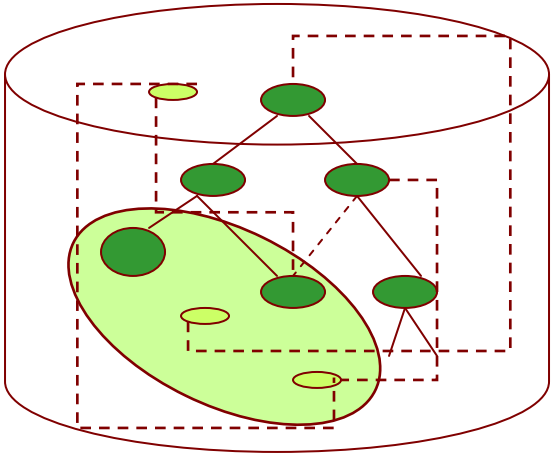


Drawn based on *SKOS Reference*, June 2008. http://www.w3.org/TR/skos-reference/

SKOS: Standard set of…

- **Resource types (Classes)**
- **Link types (Properties)**

Classes	Properties	
•ConceptScheme •Concept	• semantic relations (broader, narrower, related, etc.)	•Lexical properties (prefLabel, altLabel) • Document notes (scopeNote, definition, etc.)
•Collection •OrderedCollection	• membership relations (member, memberList)	



SKOS eXtension for Labels (XL)

- **an OPTIONAL extension of SKOS**
- **provides additional support for identifying, describing and linking lexical entities**
- **Defines:**

Class
xl:Label

relation property
xl:labelRelation

labeling properties
xl:prefLabel
xl:altLabel
xl:hiddenLabel

2008 June version changes:
- A new section on notations has been included.
- A new appendix "SKOS eXtension for Labels" has been included.
- The section on label relations has been dropped. This is replaced by the xl:labelRelation property in the eXtension for Labels appendix.

If the thesaurus becomes machine-processable, why do we still need an ontology?

An RDF/XML serialization of the RDF description of the 'Economic cooperation' concept

http://www.w3.org/TR/2005/WD-swbp-thesaurus-pubguide-20050517/

What is an ontology?

- **Ontology: a way of declaring**
  - **what** *types of things* **exist, and**
  - **what** *types of relationships* **they have with each other.**

Based on: Alistair Miles, Taxonomies and the Semantic Web, CISTRANA Workshop 02/05

"An ontology is a formal,

explicit specification of

a shared conceptualization"

machine-processable

concepts, properties relations, functions, constraints, and axioms are explicitly defined.

consensual knowledge

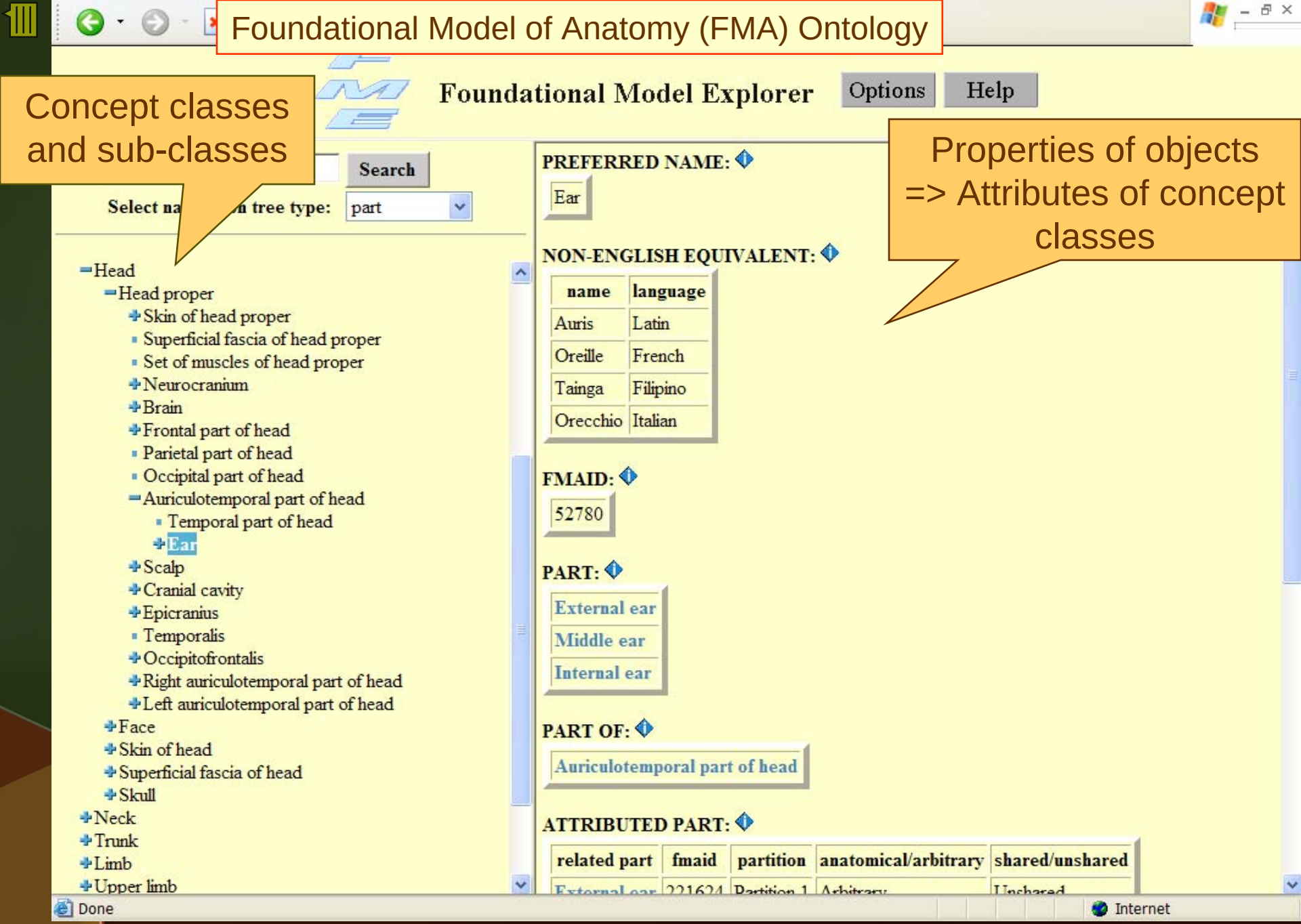
abstract model of some phenomenon in the world

Based on Studer, R., Benjamins, and Fensel, D. Knowledge engineering: Principles and methods, *Data and Knowledge Engineering*, 25(1998): 161-197.

Foundational Model of Anatomy (FMA) Ontology

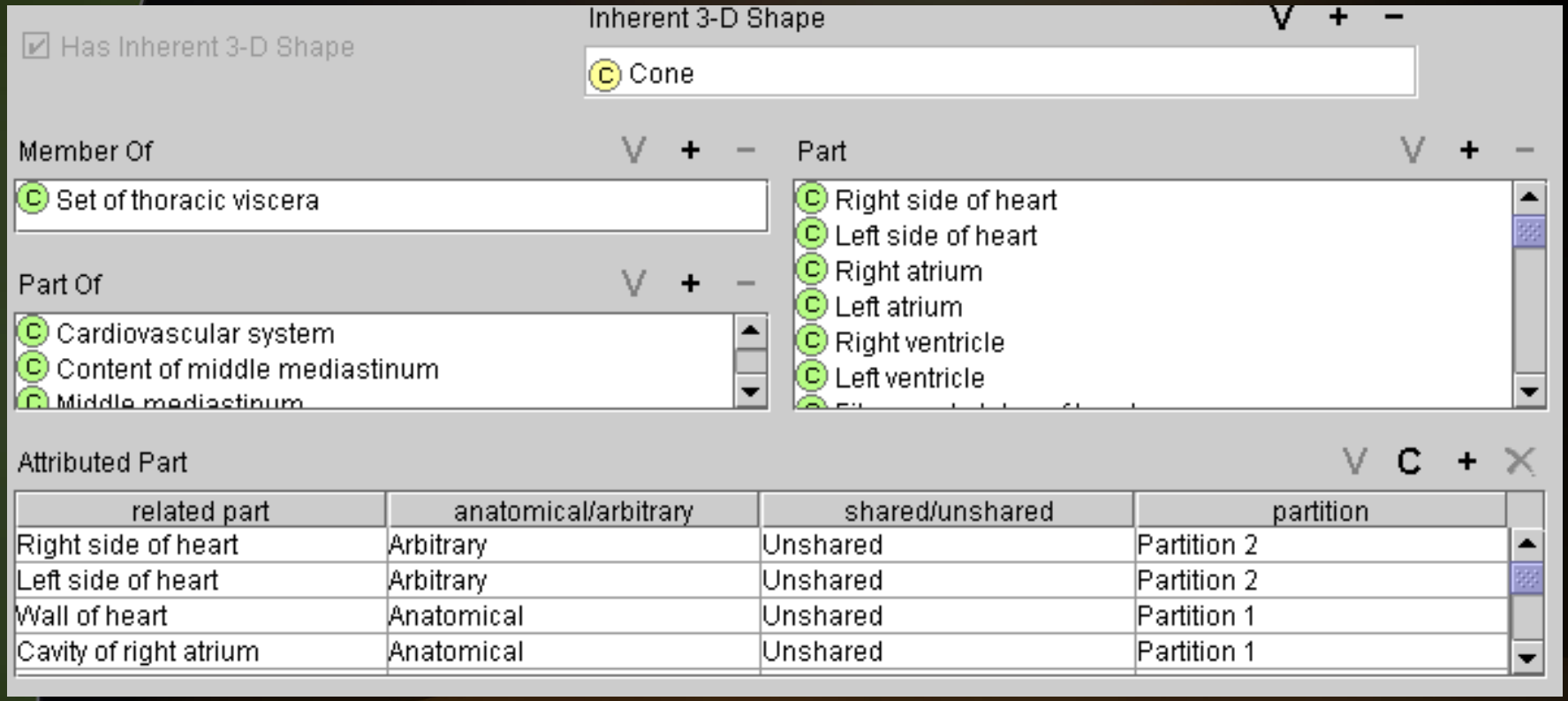
Concept classes and sub-classes

Properties of objects => Attributes of concept classes



http://sig.biostr.washington.edu/projects/fm/AboutFM.html

Foundational Model of Anatomy (FMA) Ontology



related part	anatomical/arbitrary	shared/unshared	partition
Right side of heart	Arbitrary	Unshared	Partition 2
Left side of heart	Arbitrary	Unshared	Partition 2
Wall of heart	Anatomical	Unshared	Partition 1
Cavity of right atrium	Anatomical	Unshared	Partition 1

Attributes for class *Heart* (from the classes-tab)
Source: FME 2007, http://sig.biostr.washington.edu/projects/fm/FAQs.html

Concepts, properties, relations, functions, constraints, and axioms are explicitly defined.

Foundational Model of Anatomy (FMA) Ontology viewing from BioPortal

Attributes & values

Tree View

Tree view constructed based on *hasSubclass* hierarchy

- Anatomical entity
  - Non-physical anatomical entity
  - Physical anatomical entity
    - Immaterial physical anatomical entity
    - Material anatomical entity
      - Anatomical set
      - Anatomical structure
        - Acellular anatomical structure
        - Anatomical cluster
          - Anatomical compartment
          - Body part subdivision cluster
            - Cell cluster
          - Cell part cluster
          - Compartment subdivis
          - Heterogeneous cluste
          - Organ cluster
            - Cluster of meninges
            - Erector spinae muscle group
            - Iliopsoas
            - Intertransverse muscle group
            - Lacrimal apparatus
            - Pericardium
            - Pia-arachnoid
            - Rotatores muscle group
            - Transversospinales muscle group
            - Wall of intestine
          - Organ part cluster
          - Tissue cluster

Each concept has an ID (URI)

Class/Type Details

General

Class/Type Name	Organ cluster
Id	32406

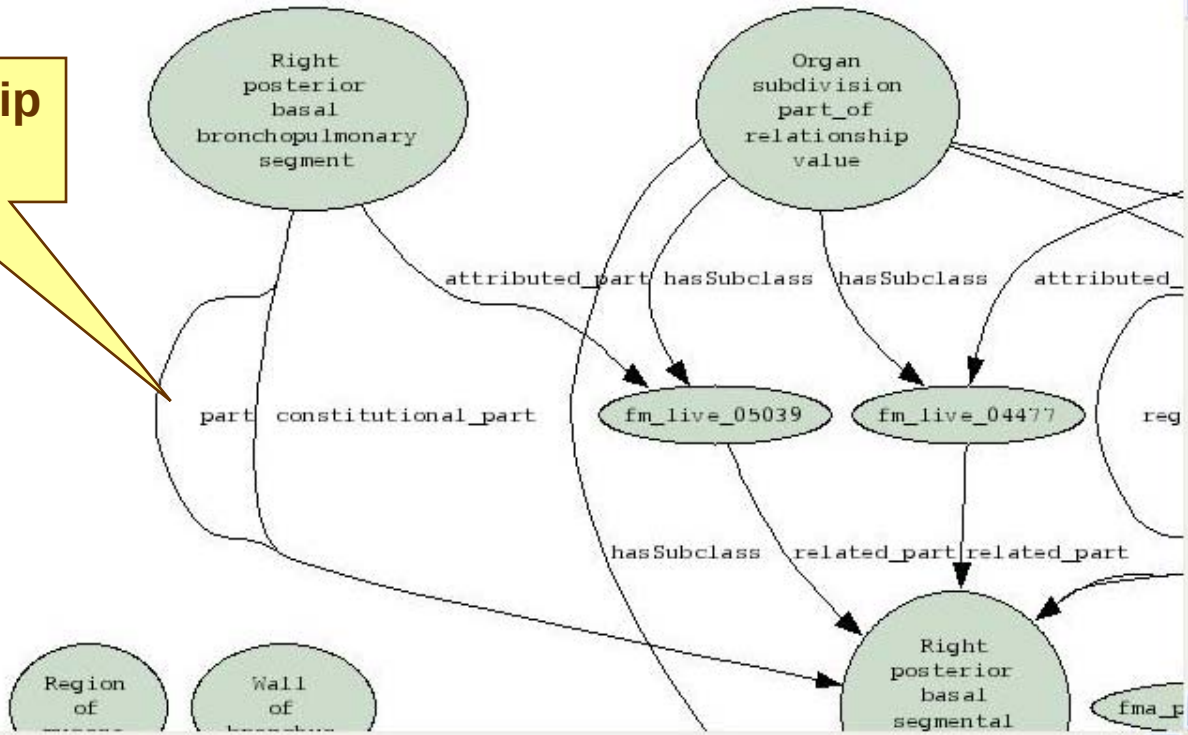
Attributes

has_boundary	true
Definition	Anatomical cluster which consists of a
has_inherent_3-D_shape	true
dimension	3-dimension
has_dimension	true
has_mass	true

Graph View

Graph Type: Hierarchy To Root (All Relationships)

Relationship types



Internet

Where are the differences? (4)

Terms and relationships of classification schemes, taxonomies, thesauri ...

become

Web resources

Terms and relationships of KOS vocabularies should be available as Web resources

Registries and repositories for metadata and KOS vocabularies, powered by semantic technologies such as RDF, SKOS, OWL, have emerged in recent years.

Registries = authoritative, centrally controlled store of information

– *W3C Web Services Glossary*

Primary functions of registries

- **Registering**
- **Publishing**
- **Managing**

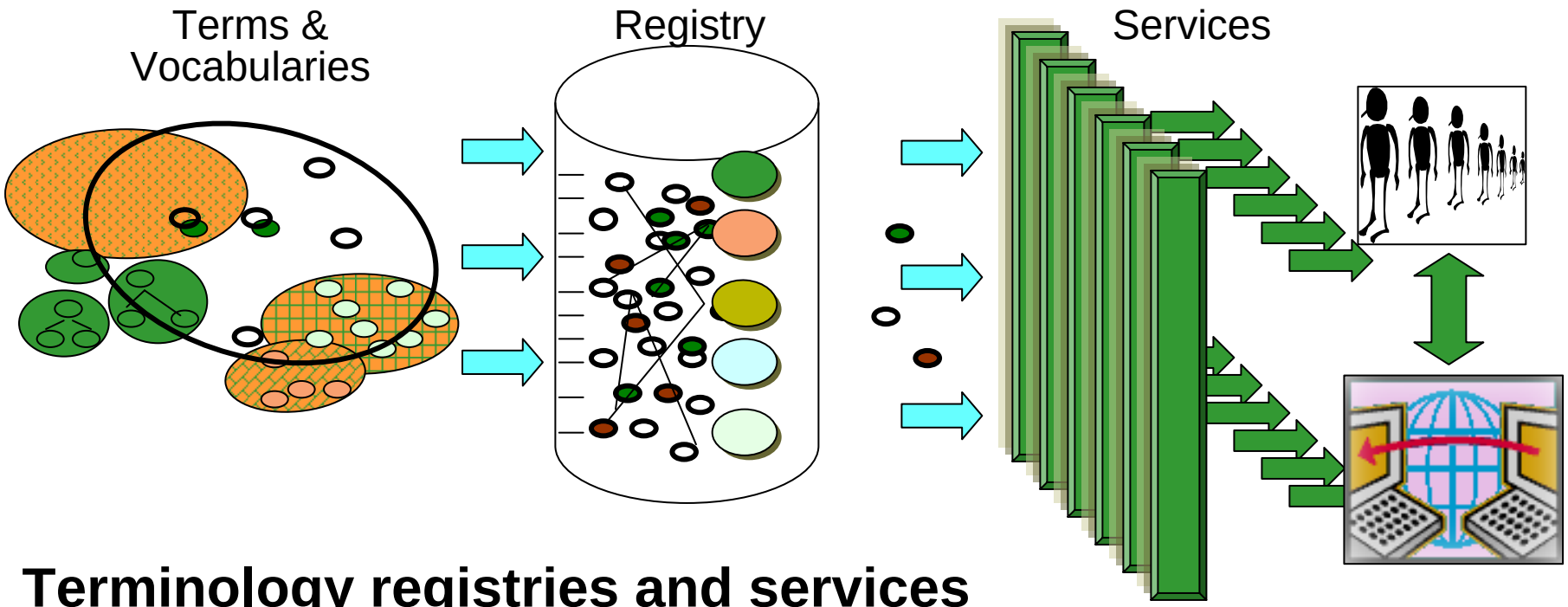
- **Data Storage**
- **User Services via a Web Interface**
  - **Navigation**
  - **Searching**
  - **Browsing**
  - **M2M Services**
    - **Querying**
    - **Using an API (Application Programming Interface) to programmatically create, view, and modify contents**
  - **Crosslinking**
  - **Crosswalking**
  - **Schema translation**

And many more … …

Terminology-based Web Services

- **Related to the terminology registries are** services**.**
- **These services, based on terminology, are used for automatic classification, term expansion, disambiguation, translation, and semantic reasoning.**

A simplified illustration



Terminology registries and services

- registering machine-accessible KOS
- mapping among concepts/terms
- making KOS content available in different kinds of tools via terminology (web) services

Registry Types and Examples

- **Metadata Schemas Registries**
  - Elements and refinements, application profiles, schemas in different bindings …
    - e,g., UKOLN CORES Registry
- **Terminology Registries / Repositories**
  - Registries for schemes (metadata) only
  - Registries of the entries of vocabularies (usually accompanied by scheme's metadata)
    - e.g., OCLC Terminologies Service; BioPortal ontology repository
- **Service Registries**
  - Terminology services may be listed in a terminology registry or separately hosted in a service registry
- **Data Standards Registries (integrated)**
  - Registries/repositories of data standards (e.g., data dictionaries, data models, schemas, and code sets)

Conclusions *of the three perspectives:*

1. The Web:

   From the Web of Documents

   to the Web of Data

2. Our Work:

   From [bibliographic and authority] control

   to knowledge organization and resource Discovery & delivery

3. Our Data:

   From machine-readable

   to machine-processable and become Web resources

1. The Web:

- Current Web:
  - Web of documents
  - the hyperlinks define a relationship between the current page and the target.
  - the link used by a human on the (traditional) Web is not named; and its role is deduced by the human reader.
  - The language: html
    - HTML allows human-understandable content to be distributed across the web and 'hyperlinked'.

- The Semantic Web
  - Web of data (resources)
  - relationships can be established between *any* two resources; there is no notion of "current" page.
  - the relationship (i.e, the link) itself is *named*
    - The definition of those relations allow for a better and automatic interchange of data.
  - The language: RDF
    - RDF allows machine-understandable content to be distributed across the web and 'hyperlinked'.

Based on: Alistair Miles, Taxonomies and the Semantic Web, CISTRANA Workshop 02/05

Conclusions *of the three perspectives:*

1. The Web:

   From the Web of Documents
   to the Web of Data

2. Our Work:

   From [bibliographic and
   authority] control
   to knowledge organization and
   resource discovery & delivery

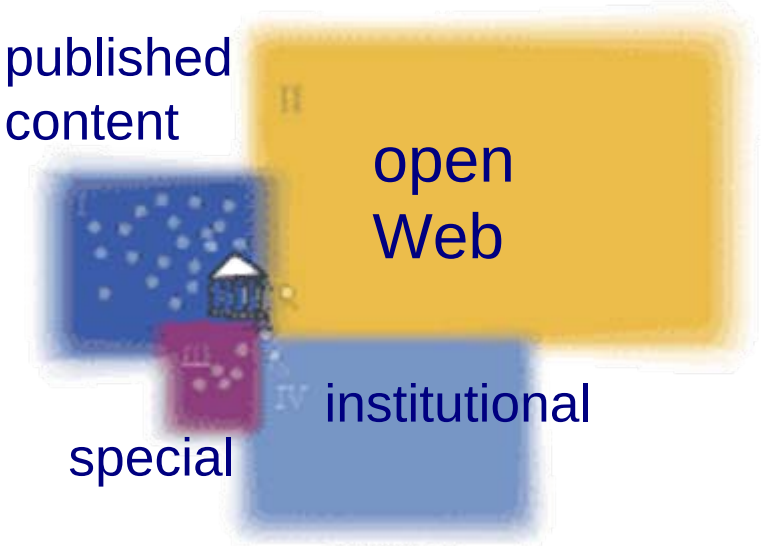
3. Our Data:

   From machine-readable
   to machine-processable and
   become Web resources

2. Our Work

- **authority data → discovery tools**
- **systems used for *bibliographic control* and *collection management* → semantic tools for Web resources**
- **bibliographic control →resource discovery & delivering**

Libraries' space within the user's "infosphere"



bibliographic control

re-focuses on

resource discovery & delivering

Potential role of libraries in the "infosphere"

2. Our Work

Graphic based on http://www.oclc.org/membership/escan/future/implications.htm

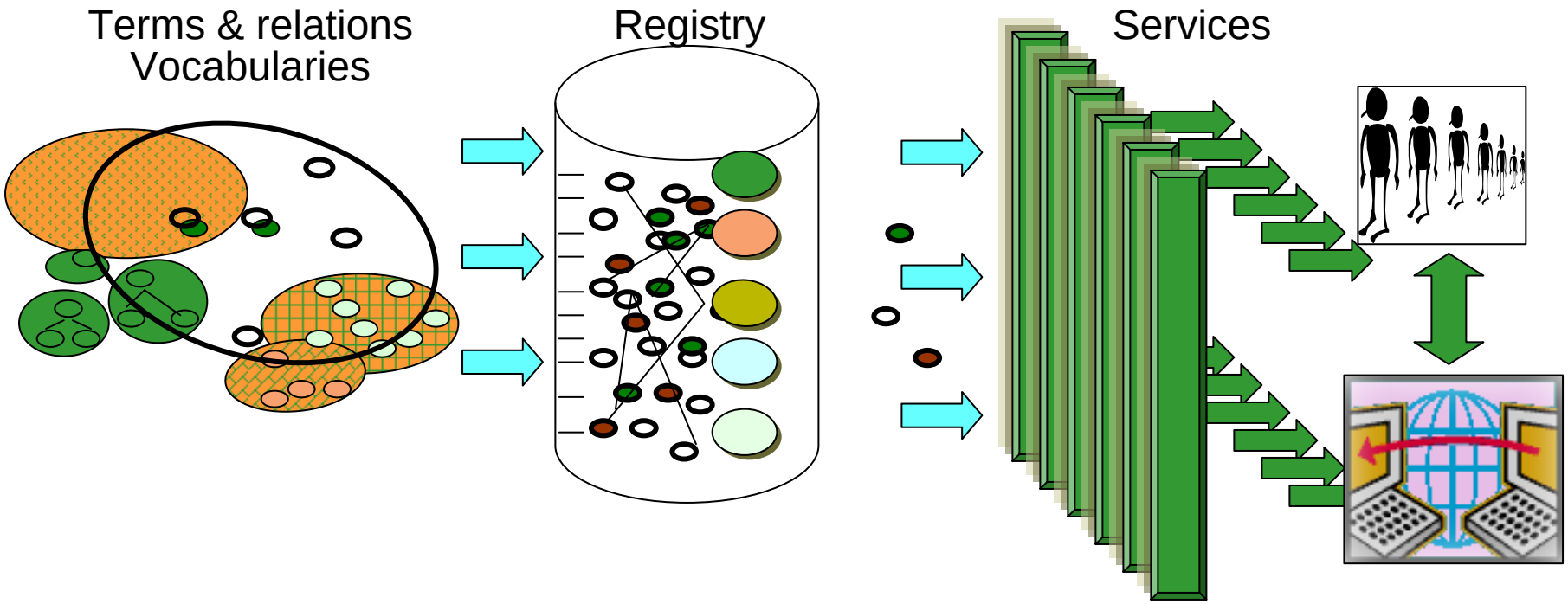
Conclusions of the three perspectives:

1. The Web:
   From the Web of Documents
   to the Web of Data

2. Our Work:
   From [bibliographic and
   authority] control
   to knowledge organization and
   resource discovery & delivery

3. Our Data:
   From machine-readable
   to machine-processable, and
   become Web resources

3. Our Data:



machine-readable ➔ machine understandable / processable

terms and relationships of classification schemes, taxonomies, thesauri, etc. ➔ Web resources

Remember:

- **The semantic Web is a web of data/resources (i.e., machine understandable content.)**

The information and knowledge professionals contribute to the creation, managing, description, accessing, discovery, and delivery of data / resources / machine-understandable content.

The Semantic Web and the Role of Information & Knowledge Professionals

Thank you!

Marcia Lei Zeng
Kent State University

mzeng@kent.edu
http://metadataetc.org