Semantic Web Systems

Ontologies

Jacques Fleuriot

School of Informatics

In the previous lecture

- ## What is the Semantic Web?

  - Web of machine-readable data

- ## Aims of the Semantic Web

  - Automated query-answering

  - Automated use of data

In this lecture

- ## What is an ontology?

- ## Frame-based approaches

- ## Folksonomy

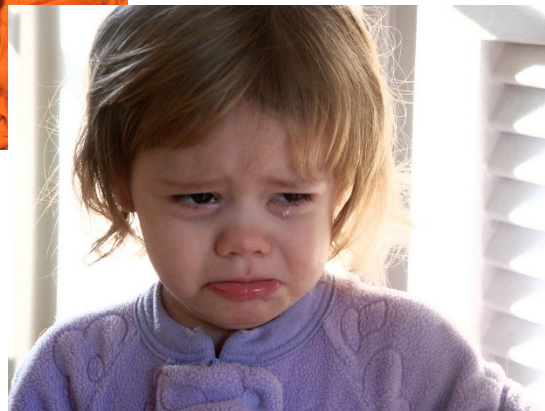# What is an ontology?

What is an ontology? (1)

Originally: a definitive account of what exists (derived from metaphysics).

Therefore, we can create a single ontology that describes the world – maybe dividing into smaller sub-ontologies as necessary.

But this is completely misconceived!

Same worldview?

A single ontology?

- A notion of relevant knowledge is highly subjective

  - Which parts of the world it is important to talk about

  - How to segregate and organise the world

  - What terms to use

- Ontologies are designed by individuals: central control is impossible and undesirable

…or several ontologies?

- But ontological differences are desirable and essential:
  - Freedom of expression
  - Ability to adapt to task
  - Changing environment

- Even direct contradictions can be desirable

- Is a tomato a fruit or a vegetable?

- The crucial task is managing these differences

Ontology definitions

"an explicit specification of a conceptualisation"

Gruber, 1993

"a formal, explicit specification of a shared conceptualisation"

Studer et al, 1998

"an explicit representation of a shared understanding of the important concepts in some domain of interest"

Kalfoglou, 2002

"a set of types, properties and relationships"
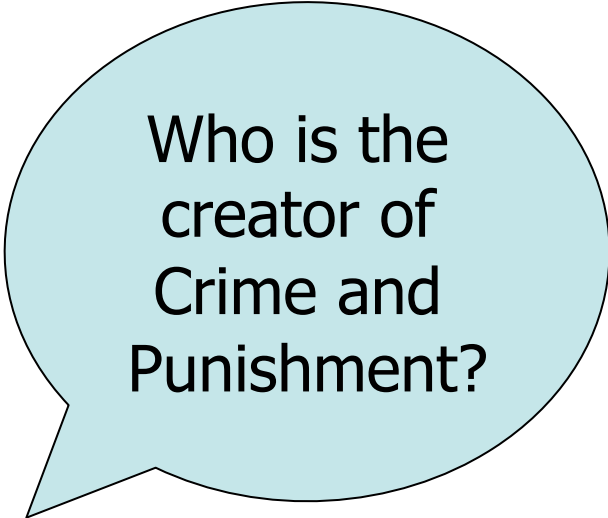
Wikipedia, 2010

But what does this mean?

What is an ontology? (2)

- Essentially: a way of encoding domain knowledge

- But there are many different choices as to how this is done.

- The word *ontology* is over-loaded: it means different things to different people.

- Ontologies are also sometimes referred to as **vocabularies**. A vocabulary is typically a light-weight ontology.
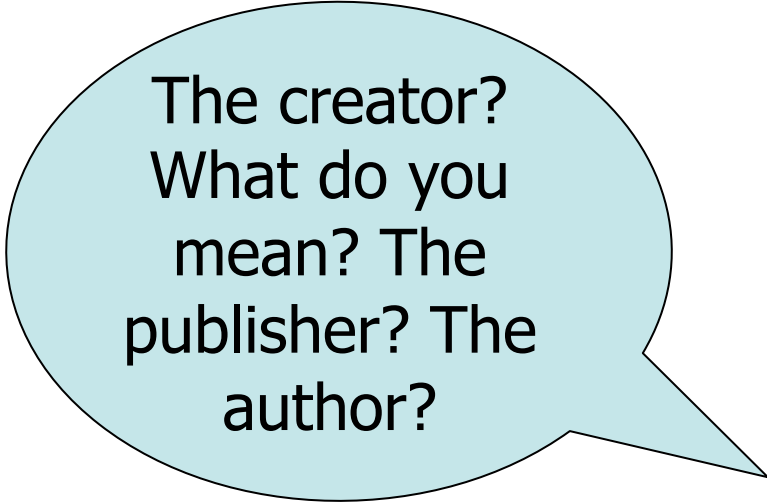
Why Semantic Web ontologies?

- Data integration

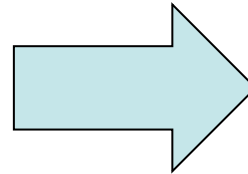Who is the creator of Crime and Punishment?

The creator? What do you mean? The publisher? The author?

Why Semantic Web ontologies?

- Data integration

- Inference

William Burnes is the
father of Robert Burns.

...

...

Father is a sublcass of
Parent.

...

...

William Burnes is the
parent of Robert Burns.

Why Semantic Web ontologies?

- Data integration

- Inference

- Automated query-answering

- Automated use of data

Example ontologies

	Language	Swoogle hits	Revised
Dublin Core	RDF	1,364,337	28 October 2006
FOAF	OWL DL	1,194,871	27 July 2005
TrackBack	RDF	502,401	
MetaVocab	RDF	441,790	16 February 2002
Basic Geo Vocabulary	RDF Schema	248,130	1 February 2006
BIO	RDF	220,228	5 March 2004
RSS 1.0	RDF Schema	201,786	6 December 2000
VCard RDF	RDF	181,962	22 February 2001
Creative Commons metadata	RDF Schema	112,216	
WOT	OWL DL	97,292	23 February 2004
SIOC	OWL DL	42,911	11 April 2008
GoodRelations	OWL DL	5,000	1 October 2011
DOAP	RDF Schema	1,442	5 November 2005
Programmes Ontology	OWL 2	943	7 September 2009
Music Ontology	OWL 2	646	14 February 2010
Pedagogical diagnosis	OWL DL	1	1 April 2012
DILIGENT Argumentation Ontology	OWL 2	1	13 September 2006
OpenGUID	RDF Schema	1	24 September 2008
Provenance Vocabulary	OWL DL	1	25 August 2009

http://semanticweb.org/wiki/Ontology

Representation and Reasoning

- Long history of attempts in Artificial Intelligence to develop knowledge-based systems.

  Given a knowledge base KB, is the sentence A true?

- We can't just look to see if A is contained in KB; typically need to do some **inference**.

- First-order logic can represent pretty much everything, but **intractable** inference seen as a major barrier.

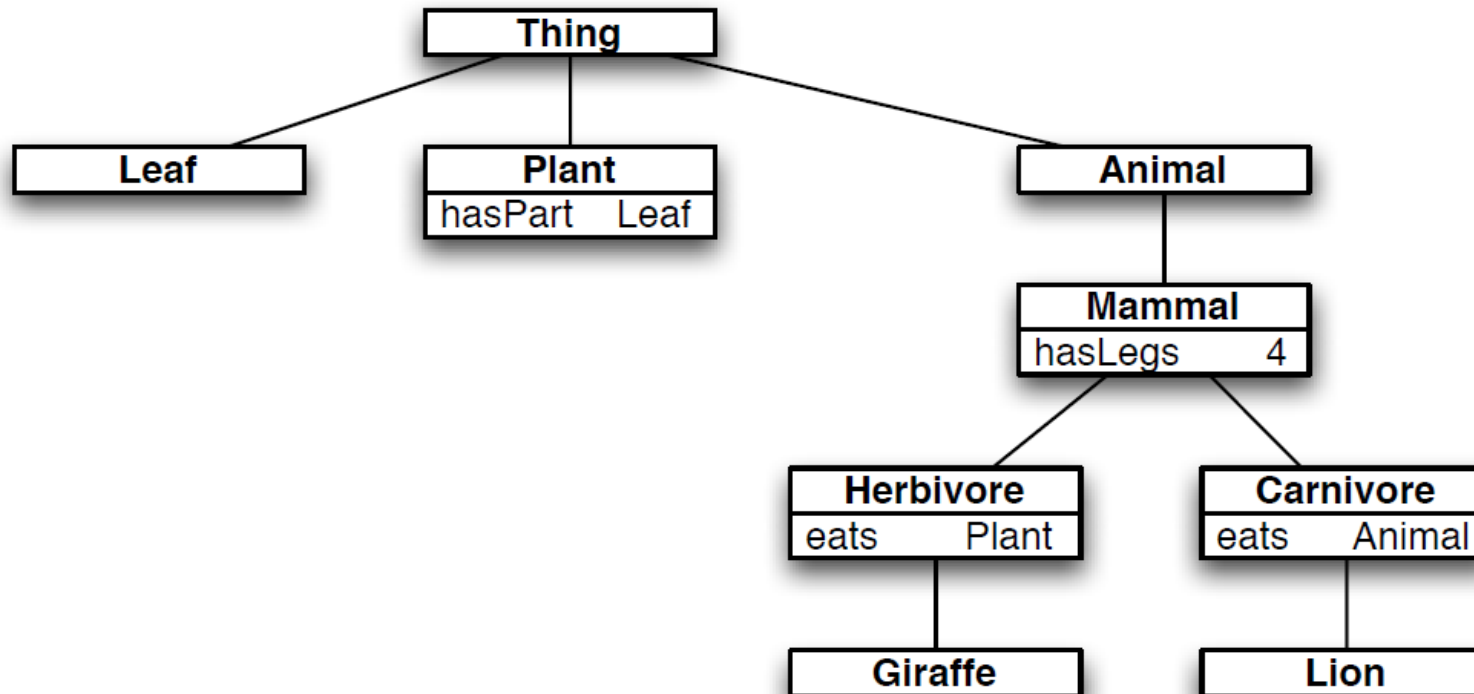 *In reality this depends on how complex the ontology is.*

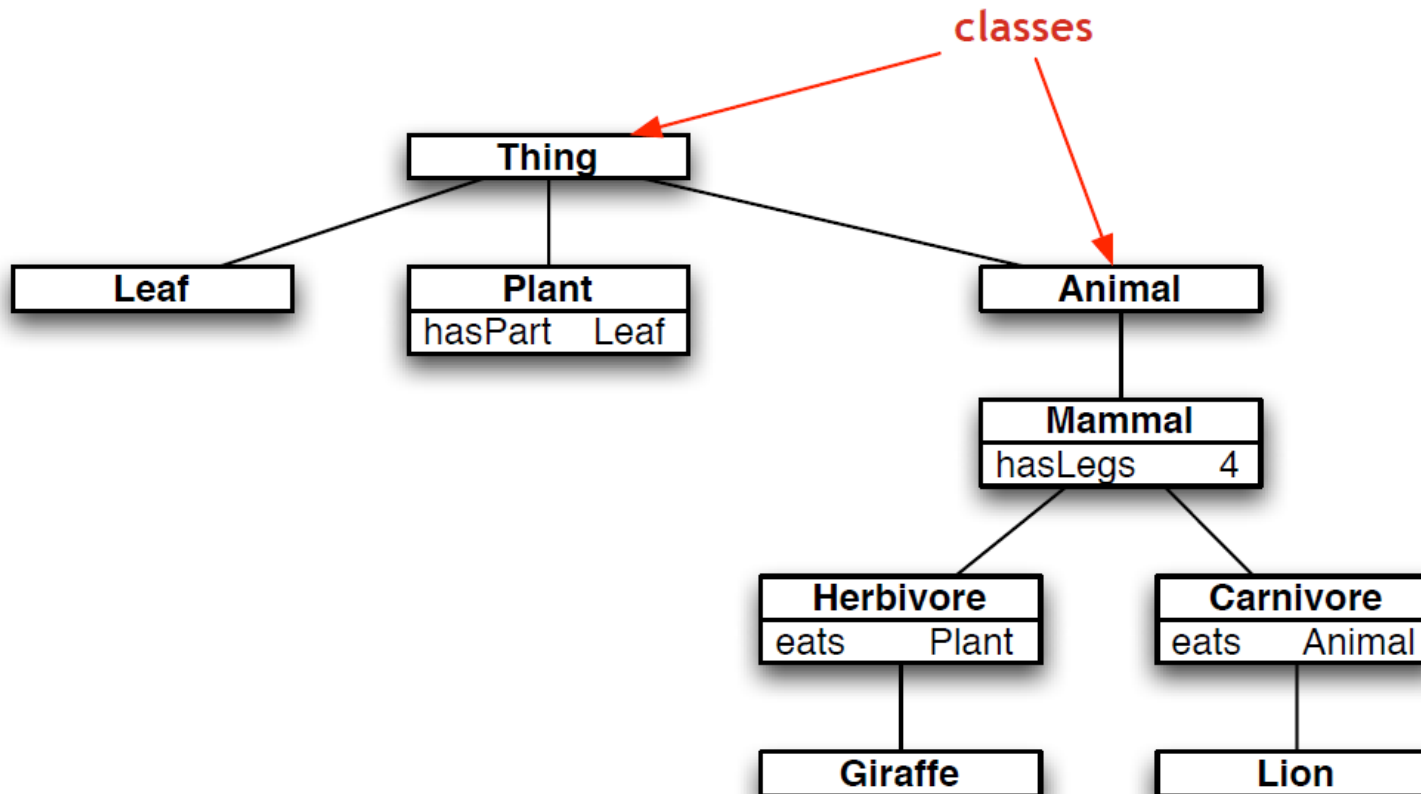- Much effort devoted to developing alternatives which were seen as more tractable.
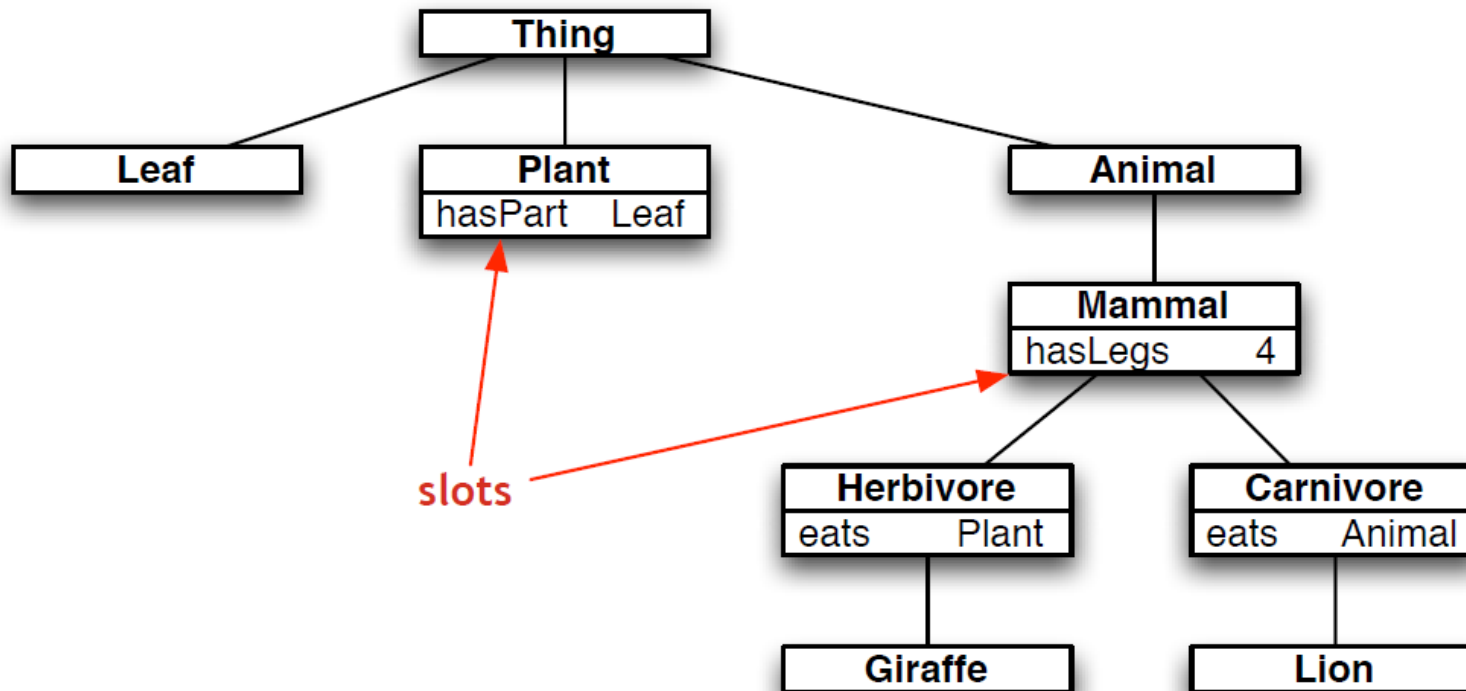
Frame-based approach

Hierarchies and Frames

Hierarchies and Frames
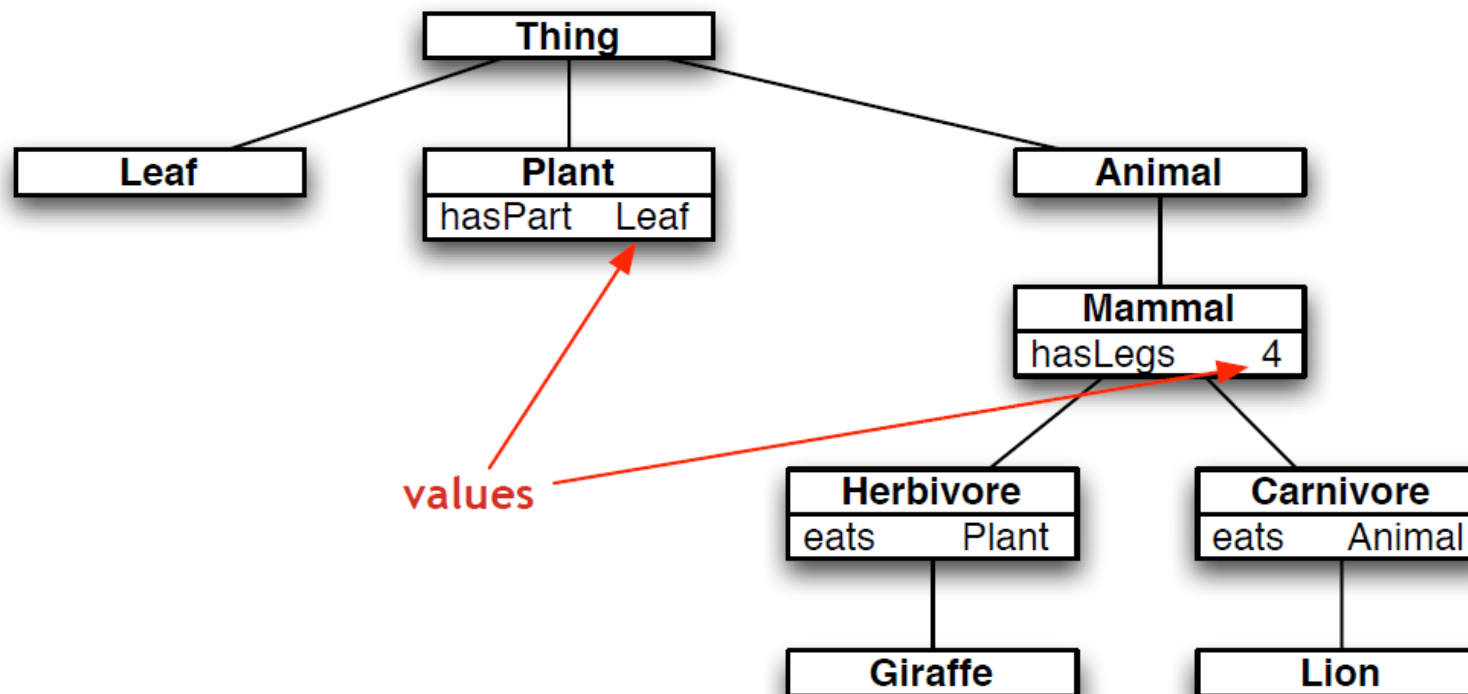
Hierarchies and Frames
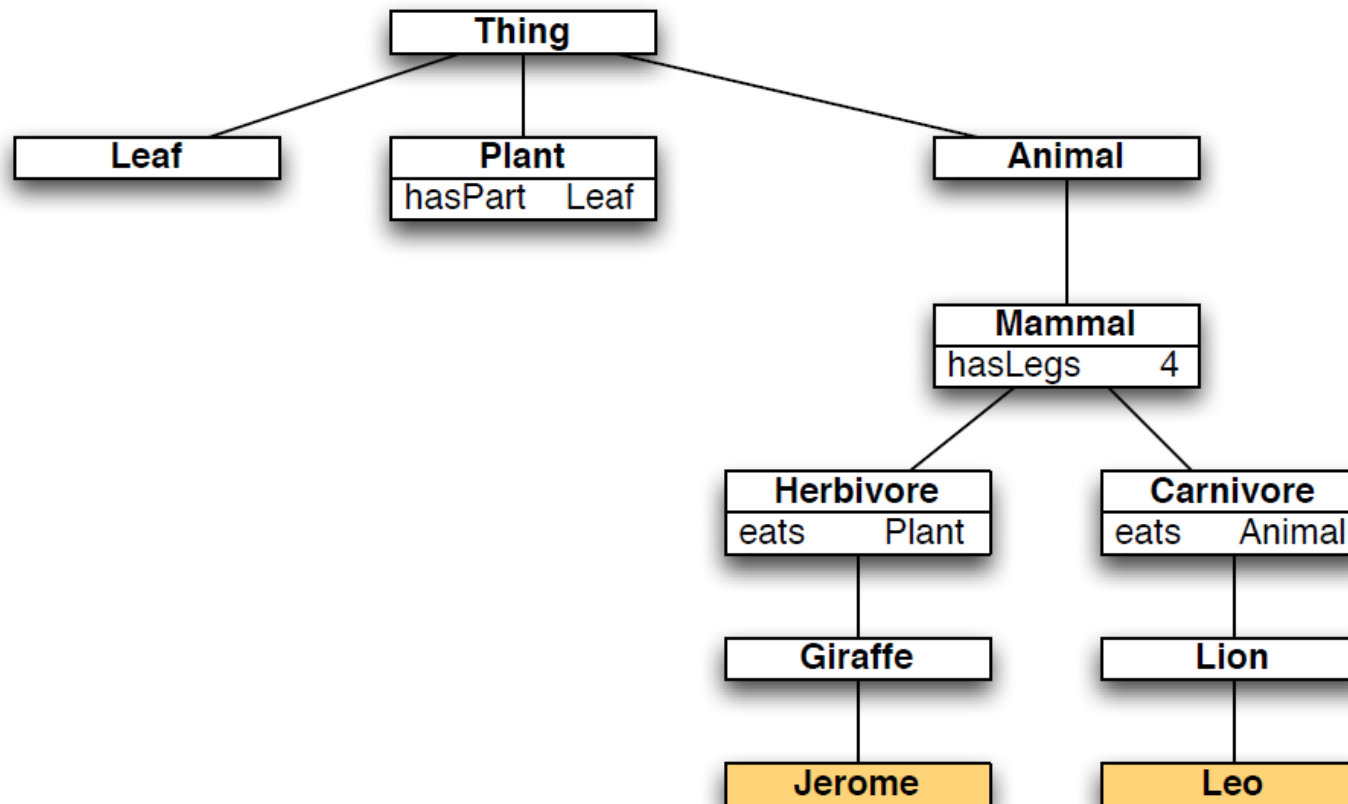
Hierarchies and Frames

Hierarchies and Frames

- Frames are a way of describing classes or concepts or types

- Usual to think of classes in terms of sets of individuals

- Frames contain **slots** with **values**

- Values can be restricted in various ways:

  - Integer, boolean or literal values

  - Enumerated values

  - Instances of a specified class

Classes and Individuals

Classes and Individuals

- Ambiguity about nature of the edge in the graph. Reflected in English:

  - A lion is a carnivore

  - Jerome is a giraffe

- Two different relations/labels:

  - **ISA**: taxonomic — a lion is a kind of carnivore

    - Lion ⊆ Carnivore $[\forall x.\ Lion(x) \rightarrow Carnivore(x)]$

  - **IO**: instance-of/membership — Jerome is a member of the class of giraffes

    - Jerome ∈ Giraffe $[Giraffe(Jerome)]$

Classes and Individuals
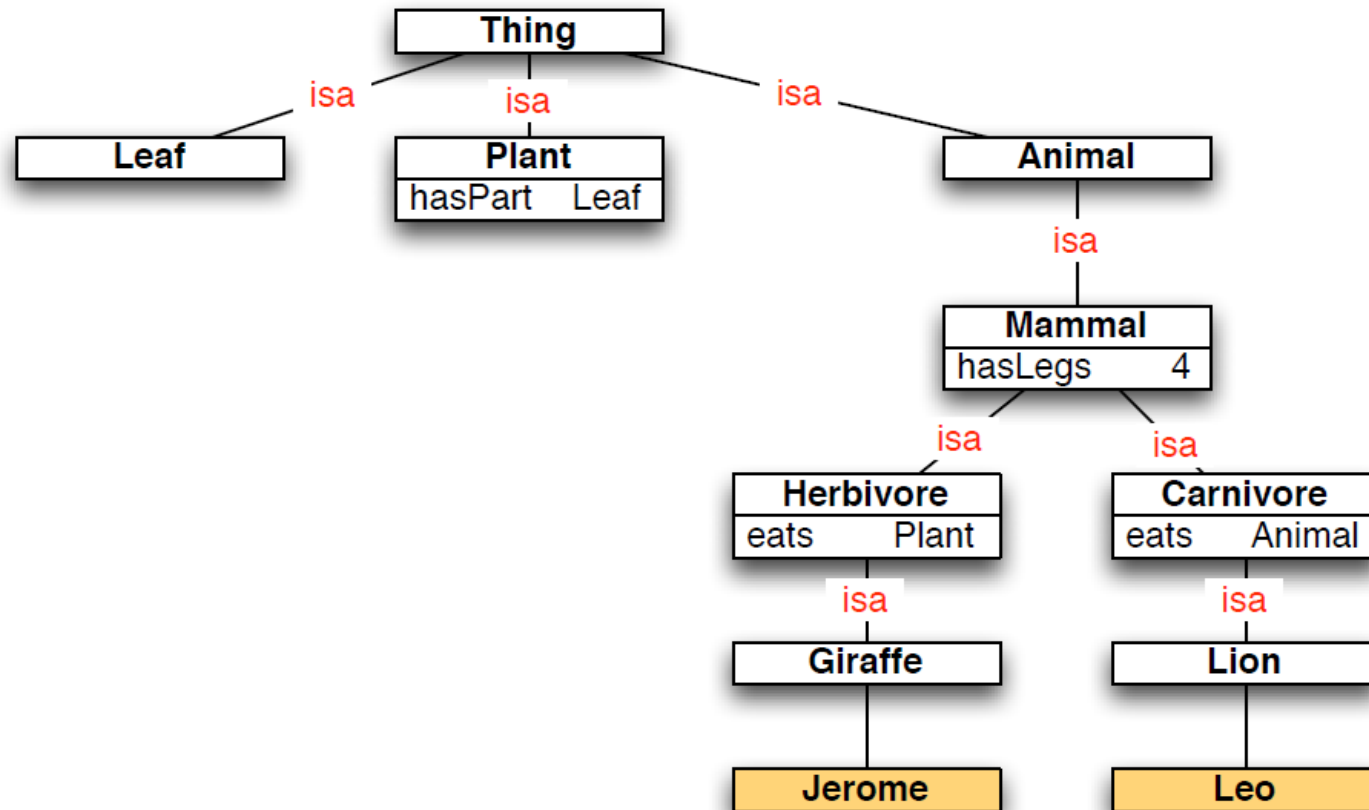
Inheritance

How many legs does Jerome have?   4

⇒ Jerome is an instance of Giraffe

⇒ Every instance of Giraffe is an instance of Herbivore

⇒ Every instance of Herbivore is an instance of Mammal

⇒ Mammals have 4 legs

So the attribute of having 4 legs is inherited by Giraffe from Mammal

Assertion vs Terminology

- Assertions - simple facts about the world:
  - Joe is married to Sue
  - Bill has a brother with no children
  - Harry's friends are Bill's cousins

- Terminology:
  - *ancestor* is the transitive closure of *parent*
  - *brother* is *sibling* restricted to males
  - *favourite-cousin* is a special type of cousin

- The KRYPTON system (Brachman, Fikes Levesque, 1983) proposed dividing KR system into two main components:
  - **ABox** (assertions)
  - **TBox** (terminological structure)

26

Folksonomy

Folksonomy

Folksonomy ⇐ Folk + Taxonomy

- Folksonomy emerged from growing practise of ad hoc tagging and labelling

    - e.g. Delicious, Flickr

    - tagging seemed to help the discovery of related resources
      "tagging that works"

- Unlike most formal ontologies, collaborative tagging is not hierarchical or centrally controlled.

- Folksonomy brings agents back into process of constructing meaning.

Tags on Flickr

Hot tags

16/11/2015

In the last 24 hours
feb23, whitenight, whitenightmelbourne, farligt, avstand, facedowntuesday, encuentros, whenigrowup, dp3merrill, whitenightmelb, fdt, northplatterealestate, fencefriday, benchmonday, genomskinlig, myvalentine, tva, dp3, happyvalentines, leicammonochrom

Over the last week
focuspocus, location4, sotu, wildgoosechase, locspring2013, whitenightmelbourne, ds106photoblitz, rosenmontagszug, myattic, kl112, londonicesculptingfestival, stroll1302, bemyflickrvalentine, week8, australiaday2013, week5theme, wolfmoon, whitenight, a4p, giornodellamemoria

Jump to: [_____] [GO]

All time most popular tags

animals architecture art asia australia autumn baby band barcelona beach berlin bike bird birds birthday black blackandwhite blue bw california canada canon car cat chicago china christmas church city clouds color concert dance day de dog england europe fall family fashion festival film florida flower flowers food football france friends fun garden geotagged germany girl graffiti green halloween hawaii holiday house india instagramapp iphone iphoneography island italia italy japan kids la lake landscape light live london love macro me mexico model museum music nature new newyork newyorkcity night nikon nyc ocean old paris park party people photo photography photos portrait raw red river rock san sanfrancisco scotland sea seattle show sky snow spain spring square squareformat street summer sun sunset taiwan texas thailand tokyo travel tree trees trip uk unitedstates urban usa vacation vintage washington water wedding white winter woman yellow zoo

29

Tags on stackoverflow

stack**overflow**

Questions | Tags | Users | Badges | Unanswered | Ask Question

Tags

popular   name   new

A tag is a keyword or label that categorizes your question with other, similar questions. Using the right tags makes it easier for others to find and answer your question.

Type to find tags: |

---

javascript × 984082
JavaScript (not to be confused with Java) is a dynamic, weakly-typed language typically used for client-side scripting. Use
972 asked today, 6023 this week

java × 961525
Java (not to be confused with JavaScript) is a general-purpose programming language designed to be used in
837 asked today, 5244 this week

c# × 872559
C# — a multi-paradigm, managed, object-oriented programming language. Questions should include code examples,
582 asked today, 3782 this week

php × 831954
a general-purpose programming language that is especially suited for web development.
690 asked today, 4031 this week

---

android × 751501
Google's OS for digital devices [Phone, Tablet, Auto, TV, Wear, Glass]. Please use Android-specific tags such as android-
661 asked today, 4035 this week

jquery × 675941
a popular cross-browser JavaScript library that facilitates DOM (Document Object Model - HTML Structure) traversal, event
445 asked today, 2665 this week

python × 495968
a dynamic and strongly typed programming language that is designed to emphasize usability. Two similar but
585 asked today, 3583 this week

html × 474642
the standard markup language used for structuring web pages and formatting content. HTML describes the structure of a
458 asked today, 2769 this week

---

c++ × 407804
a general-purpose programming language based on C. Use this tag for questions about code (to be) compiled with a C++
290 asked today, 2069 this week

ios × 385813
a mobile operating system developed by Apple. It was originally named "iPhone OS", but was renamed to "iOS" in June,
382 asked today, 2215 this week

mysql × 358115
a freely available open source Relational Database Management System (RDBMS) that uses Structured Query Language
292 asked today, 1718 this week

css × 347749
a style sheet language used for describing the look and formatting of HTML (Hyper Text Markup Language) and XML
312 asked today, 1917 this week

---

sql × 302537
a language for querying databases. Questions should include code examples, table structure, sample data, [SQL Fiddle]
198 asked today, 1184 this week

asp.net × 273126
a Microsoft web application development framework that allows programmers to build dynamic web sites, web applications
125 asked today, 839 this week

objective-c × 245118
should be used only on questions that are about Objective-C features or depend on code in the language. The tags [cocoa]
120 asked today, 620 this week

ruby-on-rails × 224075
an open source full-stack web application framework written in Ruby. It follows the popular MVC framework model and is
143 asked today, 792 this week

.net × 220972
a software framework designed mainly for the Microsoft Windows operating system. It includes an implementation of the Base
78 asked today, 574 this week

iphone × 208686
unless you are addressing Apple's iPhone and/or iPod touch specifically. For questions not dependent on hardware, use
43 asked today, 218 this week

c × 199107
a general-purpose computer programming language used for operating systems, libraries, games and other high
154 asked today, 1033 this week

arrays × 156155
an ordered data structure consisting of a collection of elements (values or variables), each identified by one (single
147 asked today, 1131 this week

30

Folksonomy – Vander Wal (2007)

- Result of personal free tagging of information and objects for one's own retrieval

- Done in a social environment (usually open and shared)

- Value is derived from people using their own vocabulary and adding explicit meaning

- Not so much categorizing, as providing a means to connect items

Folksonomy vs Formal ontology

- Vander Wal: folksonomy is not categorization

- Shirky: folksonomy is a more robust and scalable approach to categorization than formal ontology

Folksonomy vs Formal ontology

	Formal Ontology	Tagging
Domain	Small corpus	Large corpus
	Formal categories	No formal categories
	Stable entities	Unstable entities
	Restricted entities	Unrestricted entities
	Clear edges	No clear edges
Participants	Expert catalogers	Naïve catalogers
	Authoritative sources of judgement	No authority
	Coordinated users	Uncoordinated users
	Expert users	Amateur users

Shirky (2005) 'favourable characteristics'

Categorisation vs Classification – Jacob (2004)

Categorisation: division of world of experience into groups that share some perceptible similarity in a given context; context dependence provides categorisation with its power and flexibility

Classification: orderly assignment of each entity to one and only one class within a system of mutually exclusive and non-overlapping classes

- Distinction is not the same as common usage

- But formal ontologies aspire to classification, in the above sense.

Graph Structure of Tagging Systems
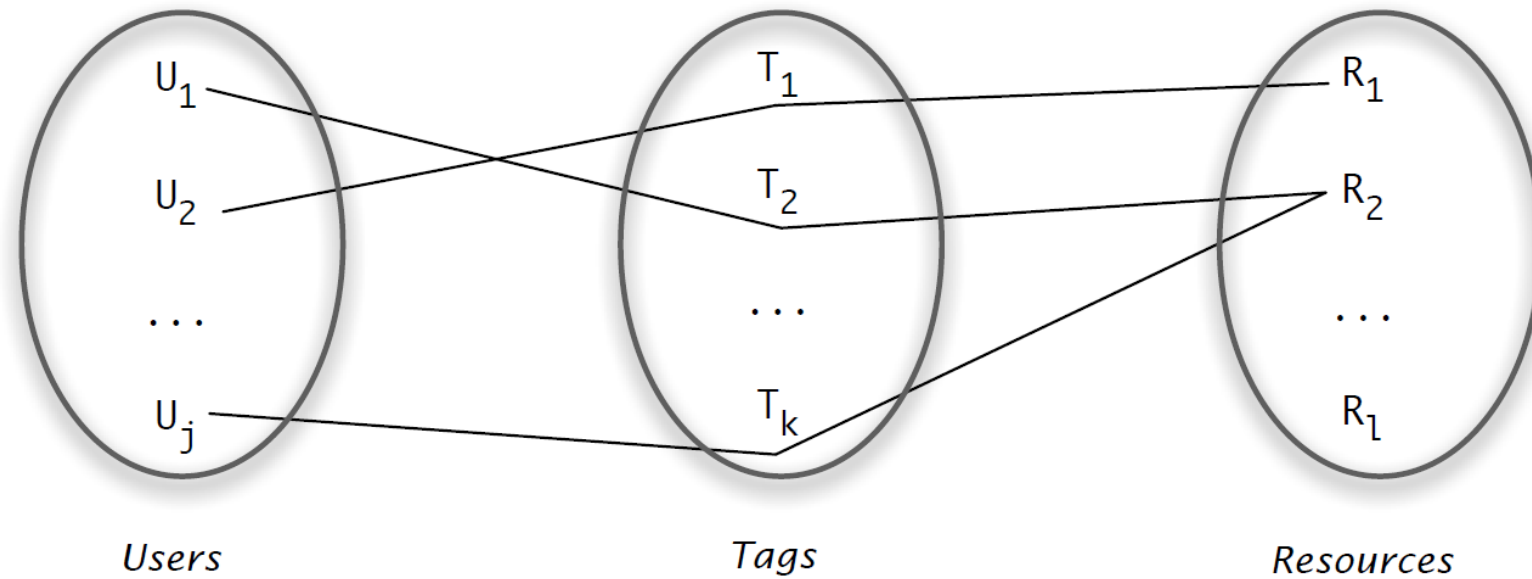
A tagging instance is a triple (user, tag, resource)
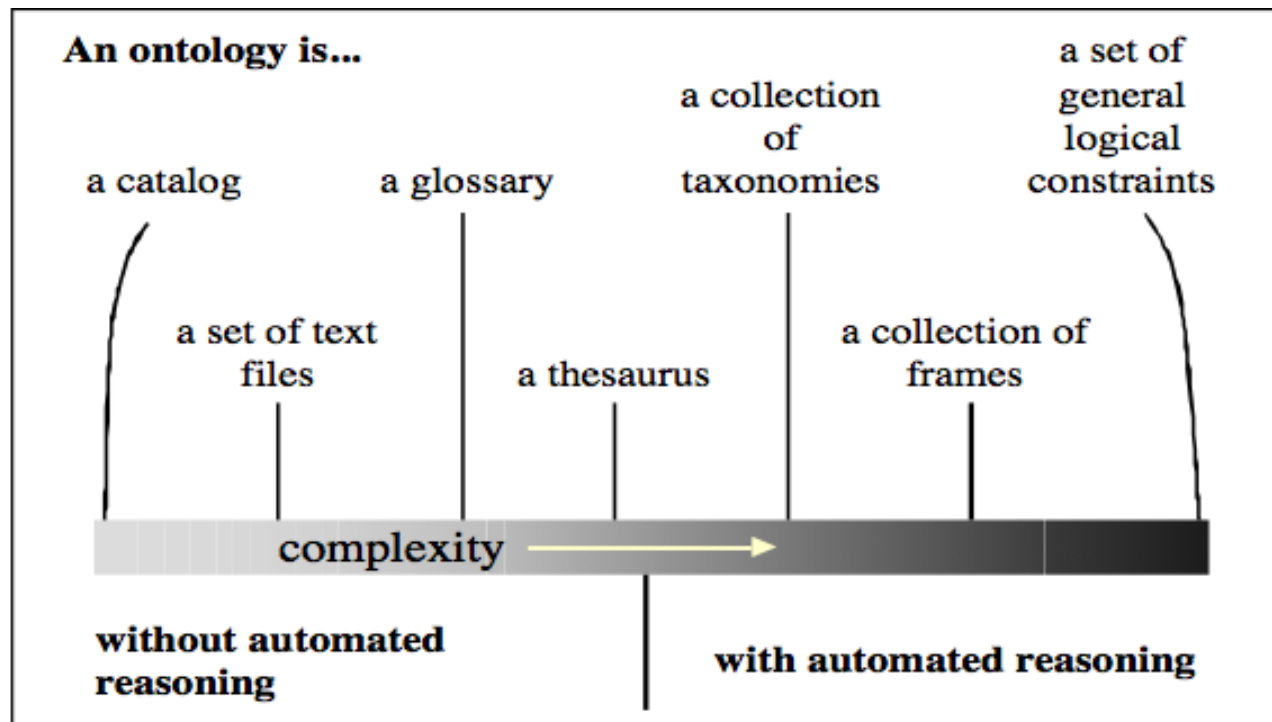
Tag distribution – Halpin et al (2007)

- What is the distribution of tags used to categorise a specific resource (e.g. a Delicious bookmark)?

- Observation: tagging distribution is stable in the sense that a small proportion of tags are consistently used to label the resource; and

- New users tend to reinforce tags in the same frequency as the stable distribution.

- Can be viewed as a 'collective categorization scheme'; i.e. ontology can emerge from collaborative tagging.

Emergent Semantics – Mika (2005)

- **Emergent semantics**: interaction of large number of agents leads to global semantic effects.

- Ontology arises from activity within network as opposed to a fixed, limited contract.

- Goal: more scalable and easily maintainable Semantic Web, incorporating social context.

The bigger picture

An ontology is...

a set of general logical constraints

a collection of taxonomies

a catalog

a glossary

a set of text files

a thesaurus

a collection of frames

complexity →

without automated reasoning

with automated reasoning

Summary

- Ontology
  - No consensus on definition
  - A formal, explicit specification of a shared conceptualisation
  - Helps define concepts, relationships and constraints (more on this next week!)

- Frames
  - A way of describing classes or concepts or types
  - Contain slots with values
  - ISA vs IO relations

- Folksonomy
  - Tagging that works
  - A scalable approach

Task

- Look at the Flickr tags (slide 29)

- Choose a subsection and try to mould it into a small formal ontology

Reading

- Chapter 2 in SWWO

- W3C page on Ontology:
  http://www.w3.org/standards/semanticweb/ontology

- Papers in MASWS Wiki on Folksonomy:
  http://sites.google.com/site/masws09/folksonomy