

The conjecture of the universal ontology and its implications for conceptual modeling

Antoni Olivé
Universitat Politècnica de Catalunya

Three research questions:

What is/could be the universal ontology?

Can we prove the conjecture of the universal ontology?

If we can, then:

What are its implications for
conceptual modeling of information systems?

What do we mean by Universal Ontology?

- The set of **all shared named concepts** people use in all domains:
 - Entity types
 - Event types
 - Attributes
 - Associations
- Concepts may be base or derived
- The taxonomic and classification relationships between them
- The definition rules of the derived concepts
- Their constraints
- The language used to define rules and constraints.

What do we mean by Universal Ontology?

Each concept in the Universal Ontology has a definition:

- Single
 - Public
 - Includes the names given to the concept in each natural language
 - Includes the necessary and sufficient conditions for an object/relationship to be an instance of the concept
-
- Example: [Person](#)

OpenCyc (Current): [<http://sw.opencyc.org/concept/Mx4rvViAkpwpEbGdrcN5Y29ycA>]

OpenCyc (Versioned): [<http://sw.opencyc.org/2009/04/07/concept/Mx4rvViAkpwpEbGdrcN5Y29ycA>]

Search

OpenCyc Collection: person

Unique ID: [[Mx4rvViAkpwpEbGdrcN5Y29ycA](#)]

English ID: [[Person](#)]

English Aliases: [["human"](#), ["layman"](#), ["laymen"](#), ["manpower"](#), ["mortal"](#), ["mortals"](#), ["people"](#), ["persons"](#)]

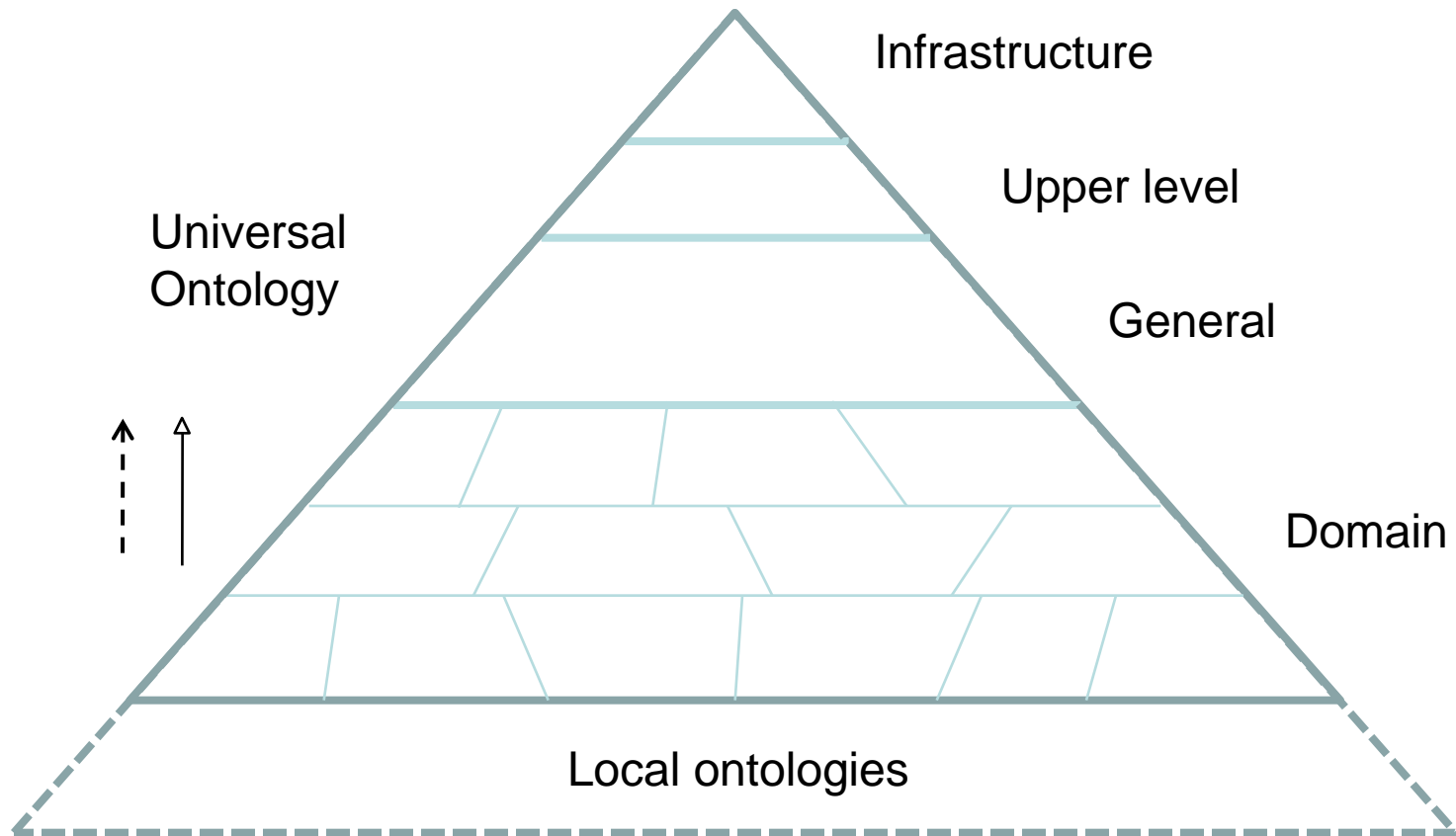
A specialization of [SocialBeing](#) and [IndividualAgent](#): the collection of all persons. Personhood is a vague, emotionally loaded yet extremely salient concept with respect to common-sense reasoning. Something is an instance of [Person](#) if it is an individual [IntelligentAgent](#) with perceptual sensibility, capable of complex social relationships, and possessing a certain moral sophistication and an intrinsic moral value, or -- if it lacks certain of these characteristics -- is a member of a distinct type of [SocialBeing](#) (usually a species) which generally possesses such characteristics and is therefore acknowledged by other members of that type as a person within their social systems. Most currently known instances of [Person](#) are instances of [HomoSapiens](#) but there is no reason why all need be (consider [Hobbits](#) in the fictional world of [LordOfTheRings_Trilogy](#)). They need not even be instances of [BiologicalLivingObject](#) (consider the possibility of a person-like AI). Also note that [Person](#) excludes non-human "legal persons", who are, however, included in the collection [LegalAgent](#).

A Type of: [human](#), [legal agent](#), [sentient](#), [social agent](#), the union of { [person](#), [animal](#) }, the union of { [person](#), [fictional people](#) }, the union of { [person](#), [groups of people](#) }, thing existing stably in time that is not an organisation

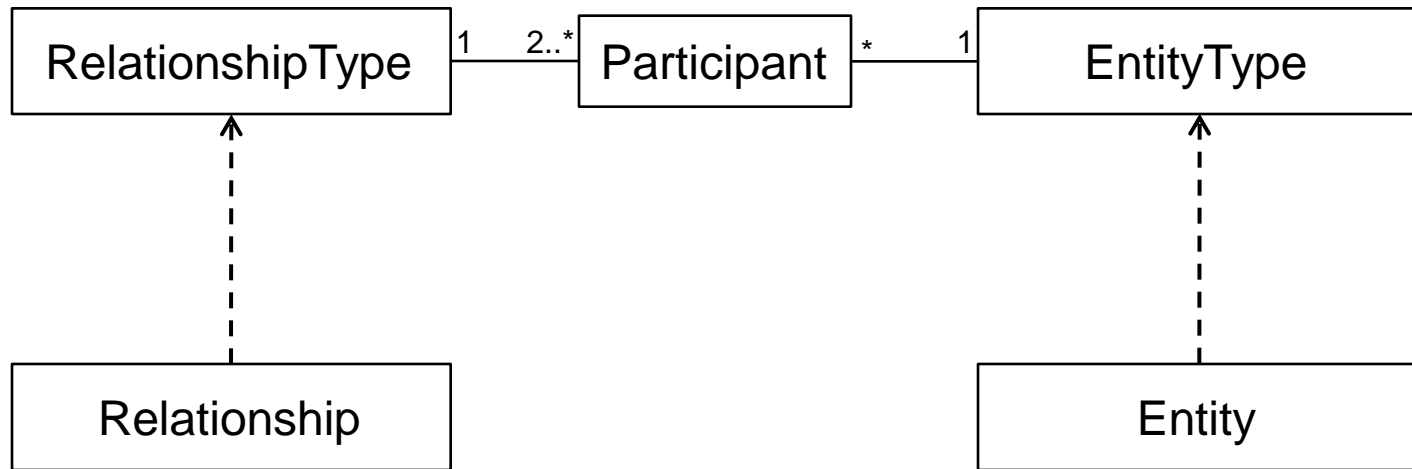
Instance of: [agents-topic](#), [calais ontology entity type](#), [candidate KB completeness node](#), [KE clarifying collection type](#), [Sample Instance Of Type For Program Fn type of _____ tangible thing the Cyc Analytic Environment](#), [terrorist attack target type type](#), the intelligence analyst reasoning domain, type of organism

Subtypes: [abolitionist](#), [acehnese](#), [activist](#), [actress](#), [adherent](#), [administrator](#), [administrator](#), [advocator](#), [Afghan](#), [Afghan](#)

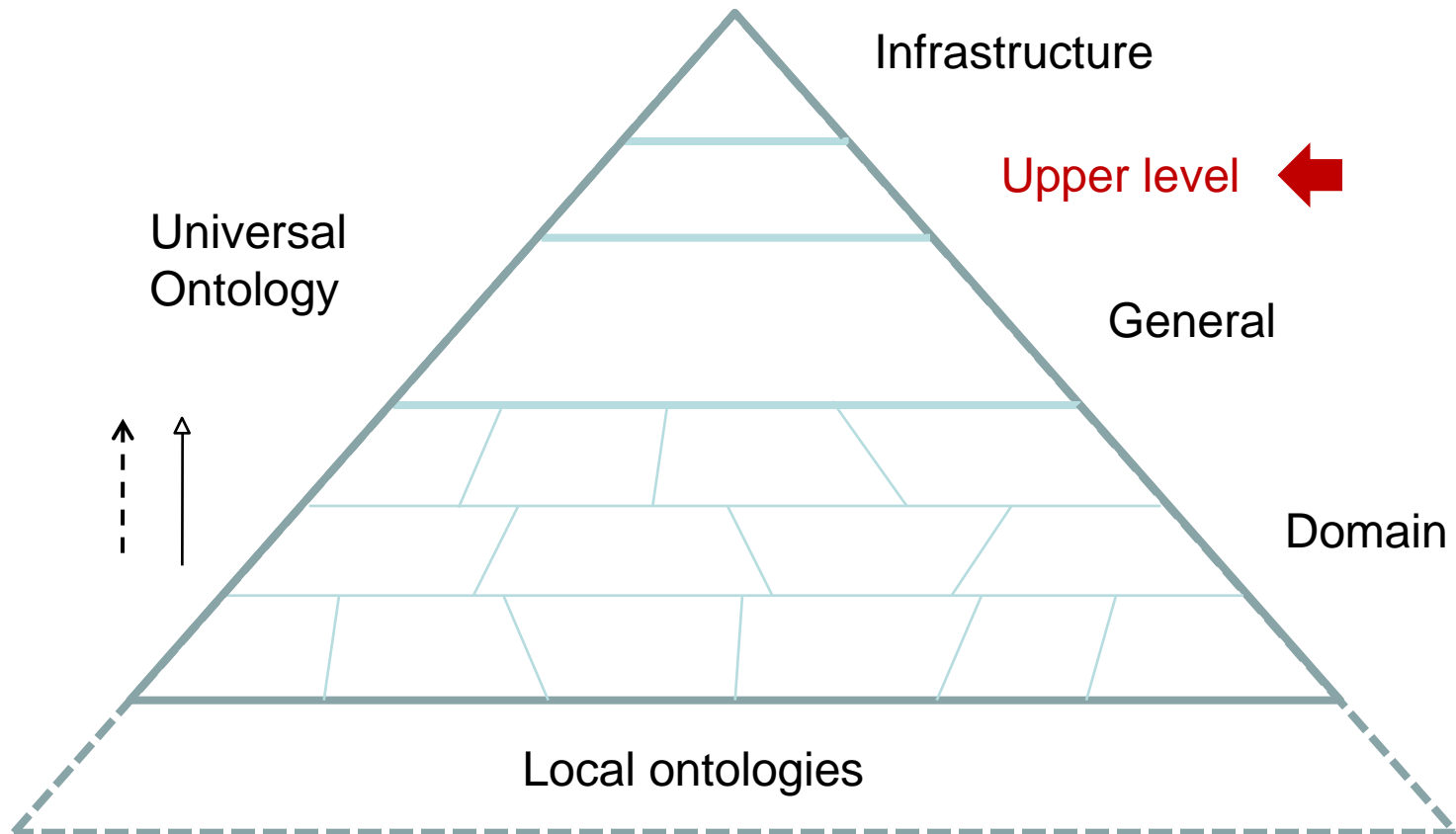
The pyramid of the Universal Ontology



Infrastructural UO Elements



The pyramid of the Universal Ontology



Upper level ontology

Abstract concepts useful for defining general knowledge

Choose one of:

- SUMO
- Cyc
- SOWA
- DOLCE
- Unified Foundational Ontology
- Another one?
- A combination of the above?

Cyc Upper Ontology

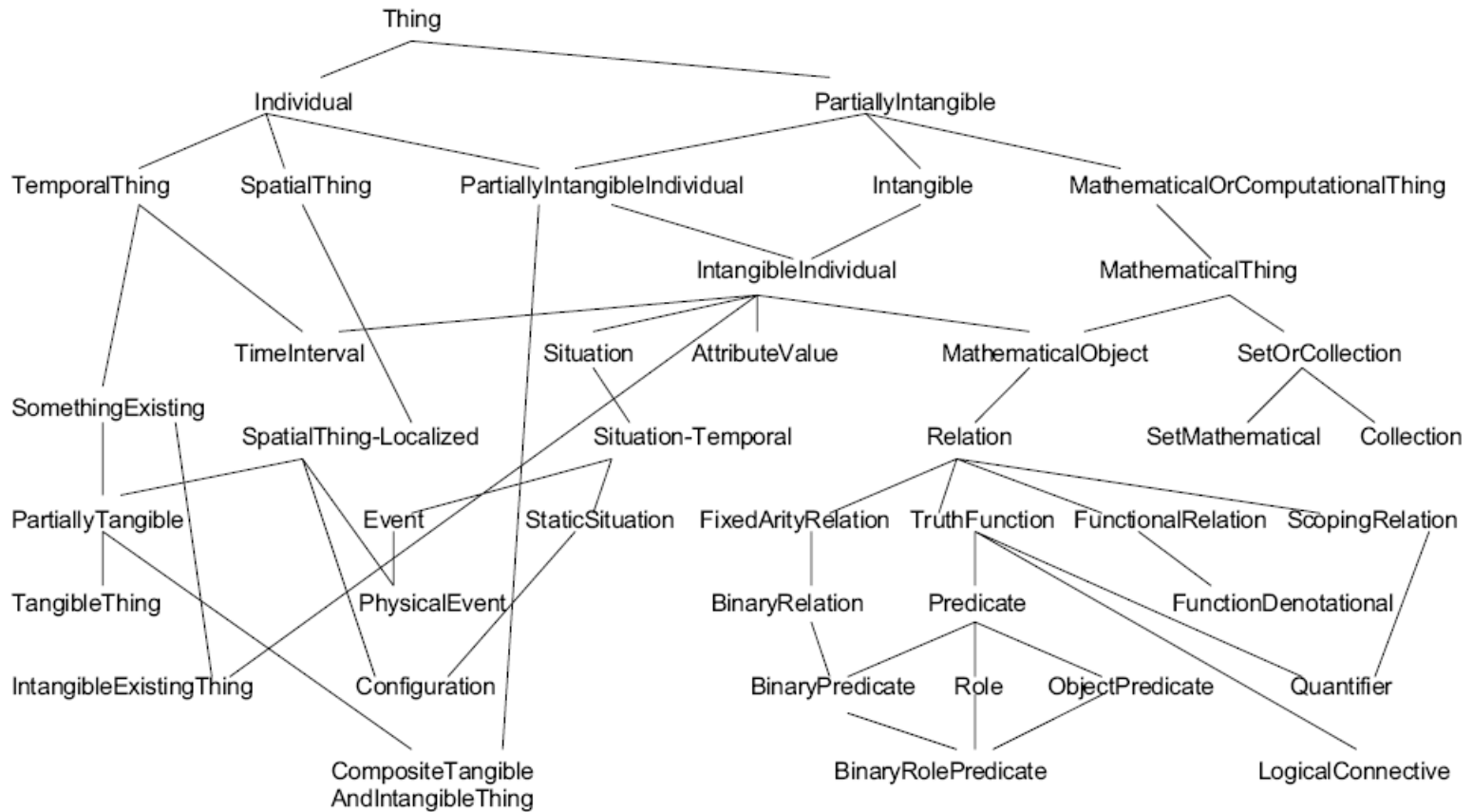


Figure 2: *The OpenCYC Upper Ontology.*

SUMO Ontology

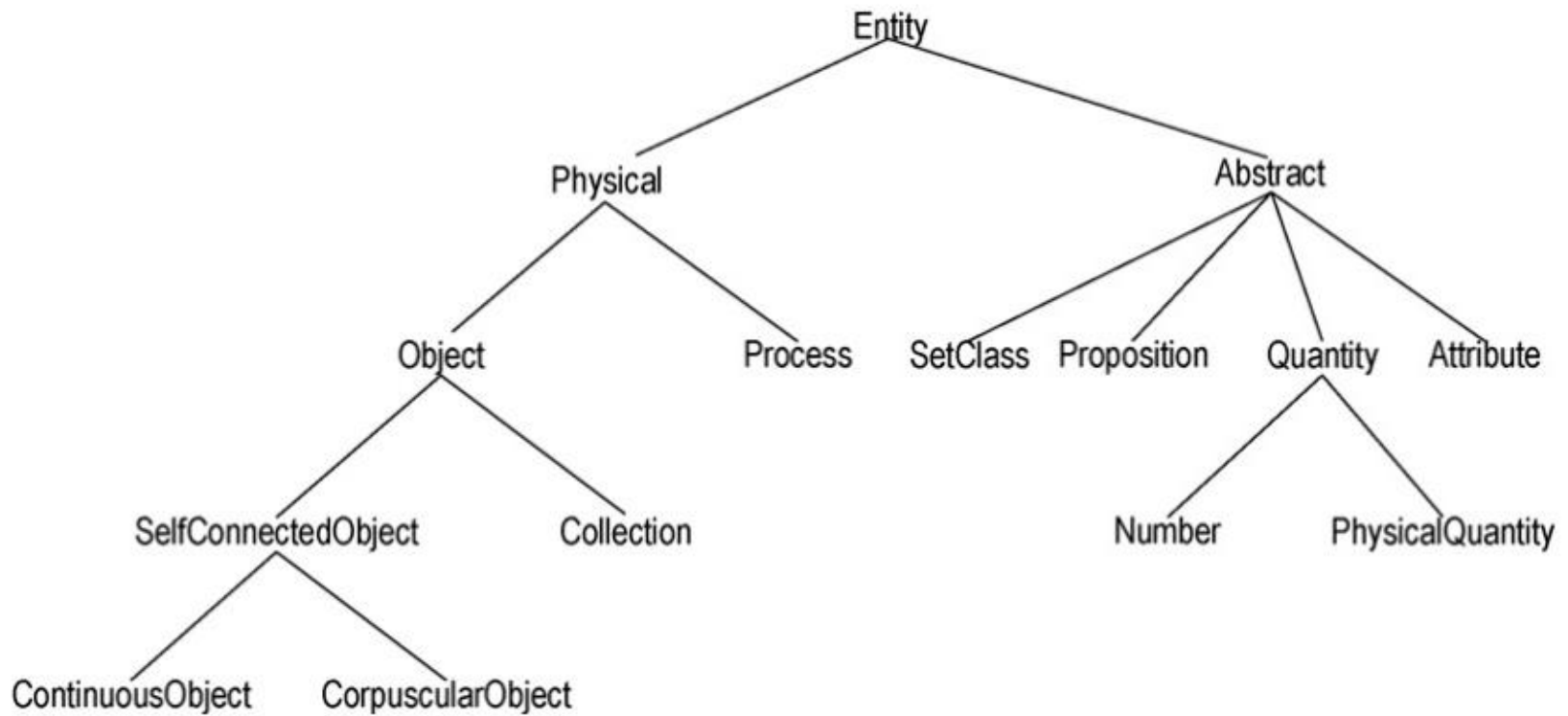


Fig. 1. SUMO taxonomy.

DOLCE Ontology

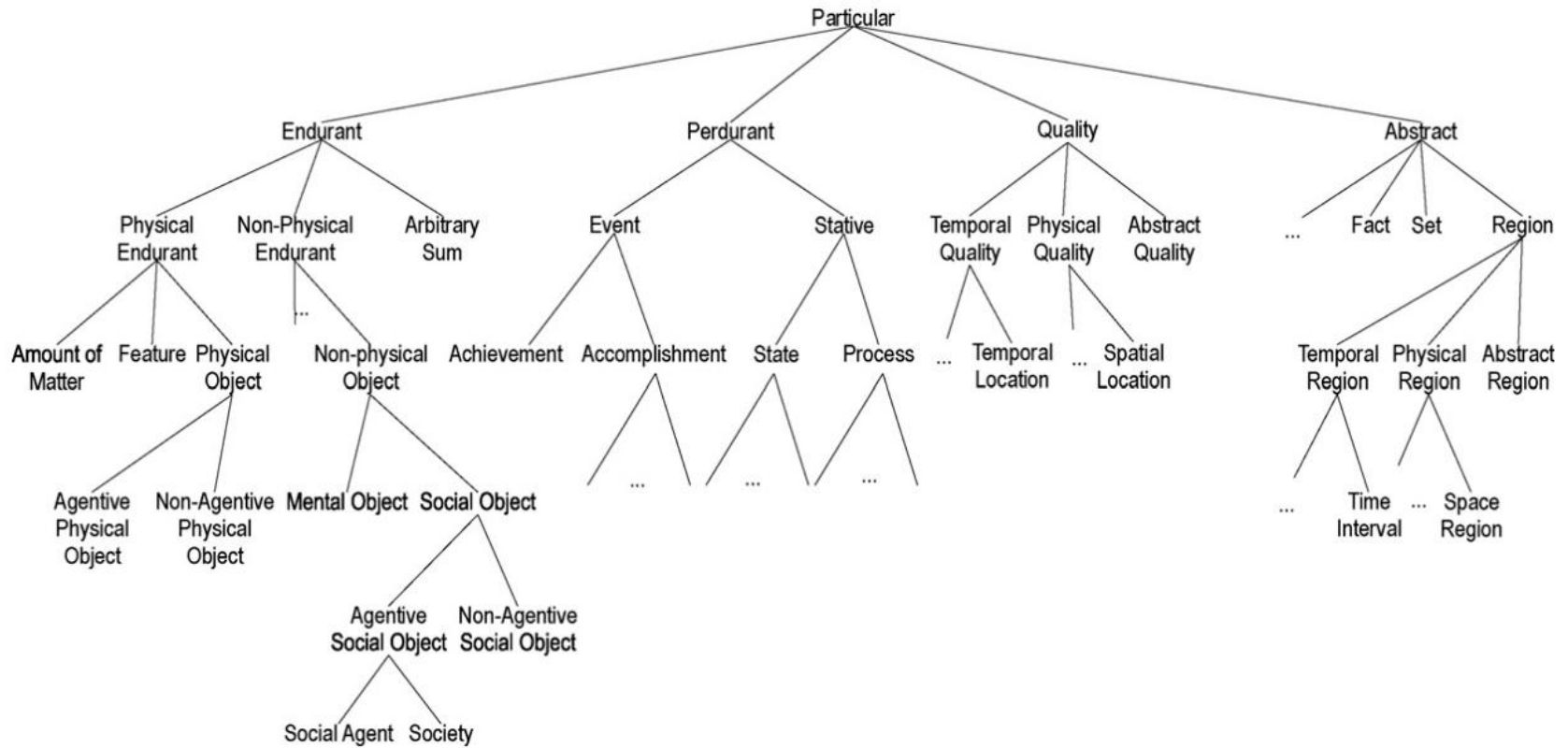
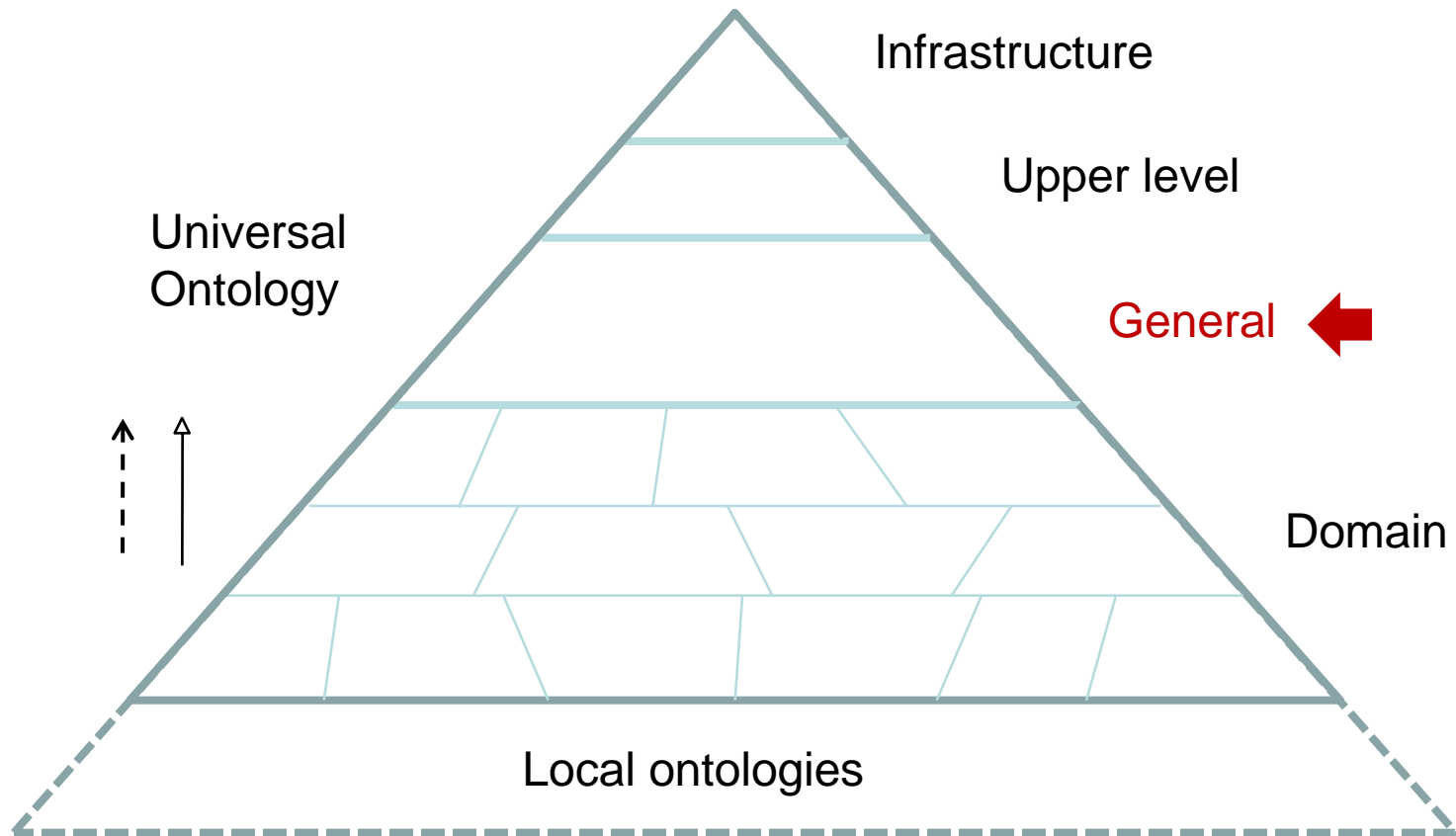


Fig. 2. DOLCE taxonomy.

The pyramid of the Universal Ontology



General ontology

Concepts for general and common sense knowledge

Choose one of:

- Cyc
- Wordnet
- SUMO
- Another one?
- A combination of the above?

General ontology: Cyc?

The latest version of OpenCyc, 2.0, was released in July 2009.

OpenCyc 1.0 includes the entire Cyc ontology containing hundreds of thousands of terms, along with millions of assertions relating the terms to each other.

The knowledge base contains 47,000 concepts and 306,000 facts and can be browsed on the OpenCyc website. (<http://sw.opencyc.org/>)

It provides:

- a foundation of concepts that can be immediately used and easily extended; and
- CycL, an expressive language that supports the OpenCyc ontology.

General ontology: Wordnet?

WordNet® is a large lexical database of English.

Nouns, verbs, adjectives and adverbs are grouped into sets of cognitive synonyms (synsets), each expressing a distinct concept.

Synsets are interlinked by means of conceptual-semantic and lexical relations.

	Synsets
Noun	82115
Verb	13767
Adjective	18156
Adverb	3621
Totals	117659

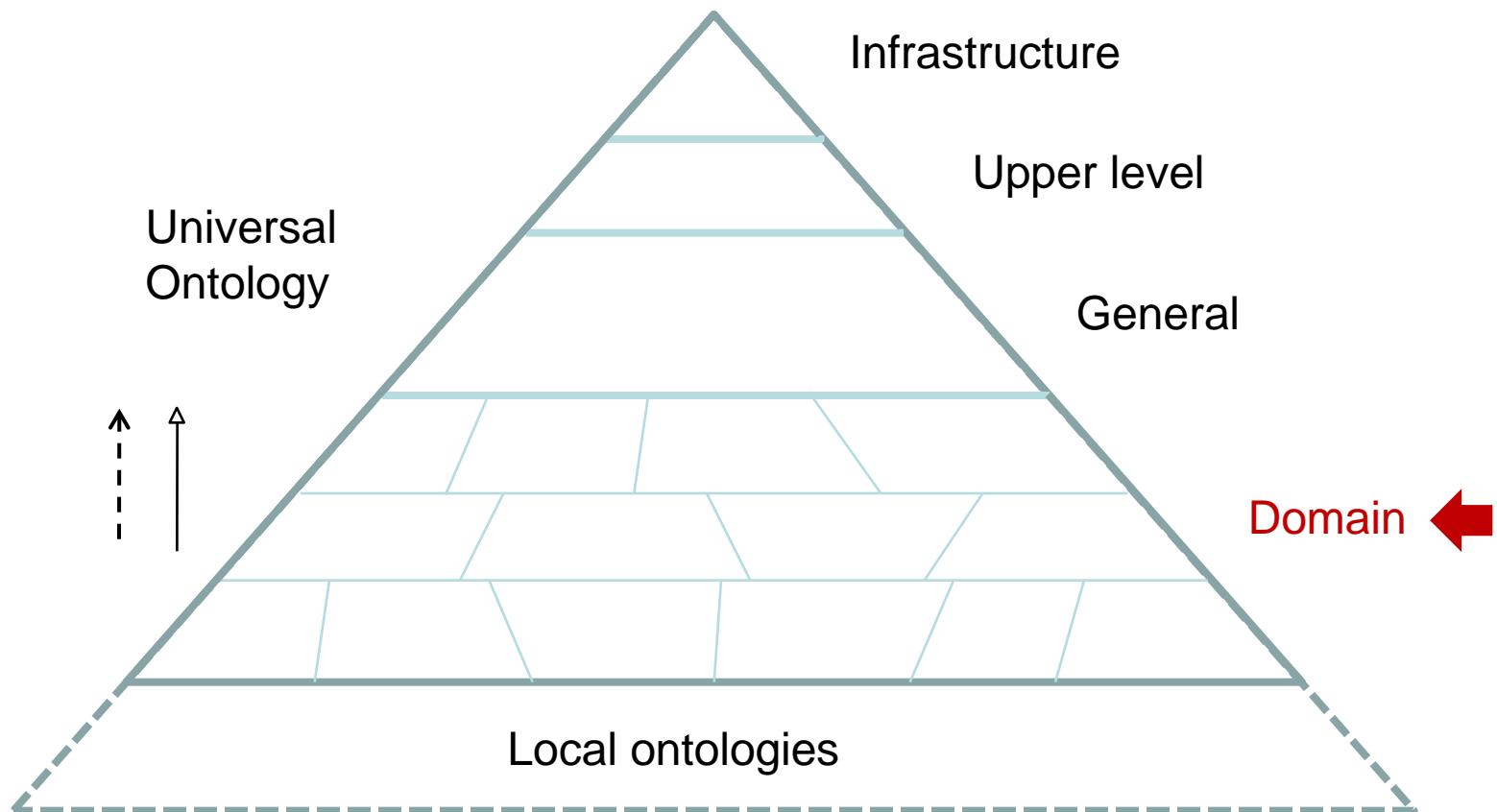
General ontology: SUMO?

The Suggested Upper Merged Ontology (SUMO) and its domain ontologies form the largest formal **public** ontology in existence today (26652 terms).

SUMO is the only formal ontology that has been mapped to all of the WordNet lexicon.

SUMO is written in the SUO-KIF language.
(<http://sigma.ontologyportal.org:4010/sigma/Browse.jsp?kb=SUMO>)

The pyramid of the Universal Ontology



Domain ontologies

- A domain ontology is a set of concepts relevant to a particular domain.
- A domain ontology is based on a set of explicitly imported ontologies:
 - Infrastructure, Upper level, General
 - Other domain ontologies
- A domain ontology is defined by a domain authority.
- The domain authority ensures that:
 - Each concept is correctly defined
 - A concept in the domain ontology is not redundant with:
 - Another concept in the domain ontology
 - Another concept in the imported ontologies
 - The defined constraints are valid

Domain ontologies



Unified Medical Language System® (UMLS®)

Home > Biomedical Research & Informatics > UMLS > Knowl

Statistics - 2011AA Release

Official Counts:

Release version: 2011AA

Release format: RRF

Concepts: 2,404,937

Number of concept names (AUIs): 10,655,002

Domain ontologies



▶ eCl@ss – a summary

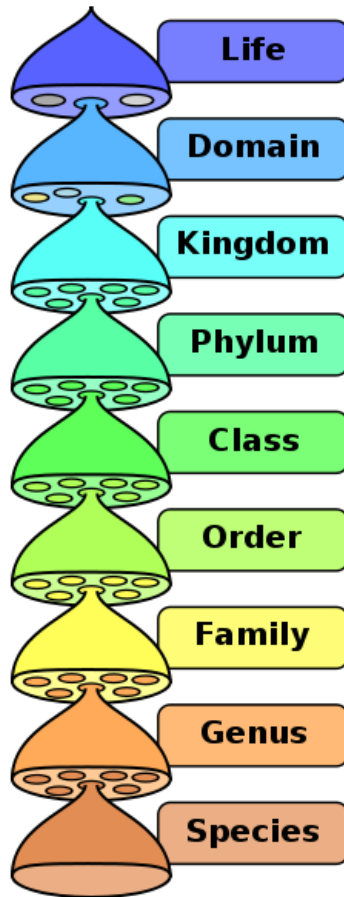
eCl@ss is THE cross-industry product data standard for classification and clear description of products and services. eCl@ss has established itself as the only ISO/IEC compliant industry standard nationally and internationally.

Besides its classical applications such as procurement, controlling and distribution, eCl@ss release 7.0 also shows particular strength regarding company-wide process data management as well as engineering.

▶ eCl@ss – the established standard

Since 2000, eCl@ss holds an established position in industry, commerce, crafts, food, services and much more. With its 38,000 product classes and 16,000 properties, eCl@ss covers the majority of traded goods and services. Many industry standards (e.g. from the electronic industry, medical technology, construction engineering, paper industry / office technology) are looking for interoperability to realize the potentials of a cross-industry standard.

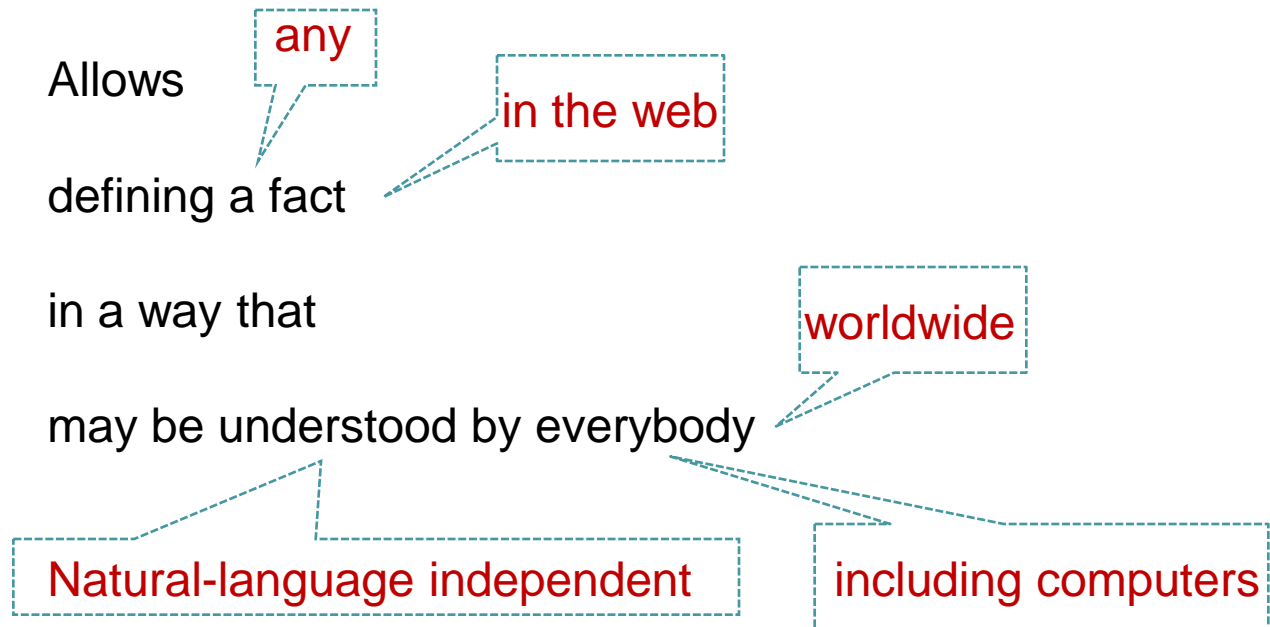
Domain ontologies



How many species are known to currently exist in the world.

Category	Species	Totals
Vertebrate Animals		
Mammals	5,490	
Birds	9,998	
Reptiles	9,084	
Amphibians	6,433	
Fishes	31,300	
Total Vertebrates		62,305
Invertebrate Animals		
Insects	1,000,000	
Spiders and scorpions	102,248	
Molluscs	85,000	
Crustaceans	47,000	
Corals	2,175	
Others	68,827	
Total Invertebrates		1,305,250
Plants		
Flowering plants (angiosperms)	281,821	
Conifers (gymnosperms)	1,021	
Ferns and horsetails	12,000	
Mosses	16,236	
Red and green algae	10,134	
Total Plants		321,212
Others		

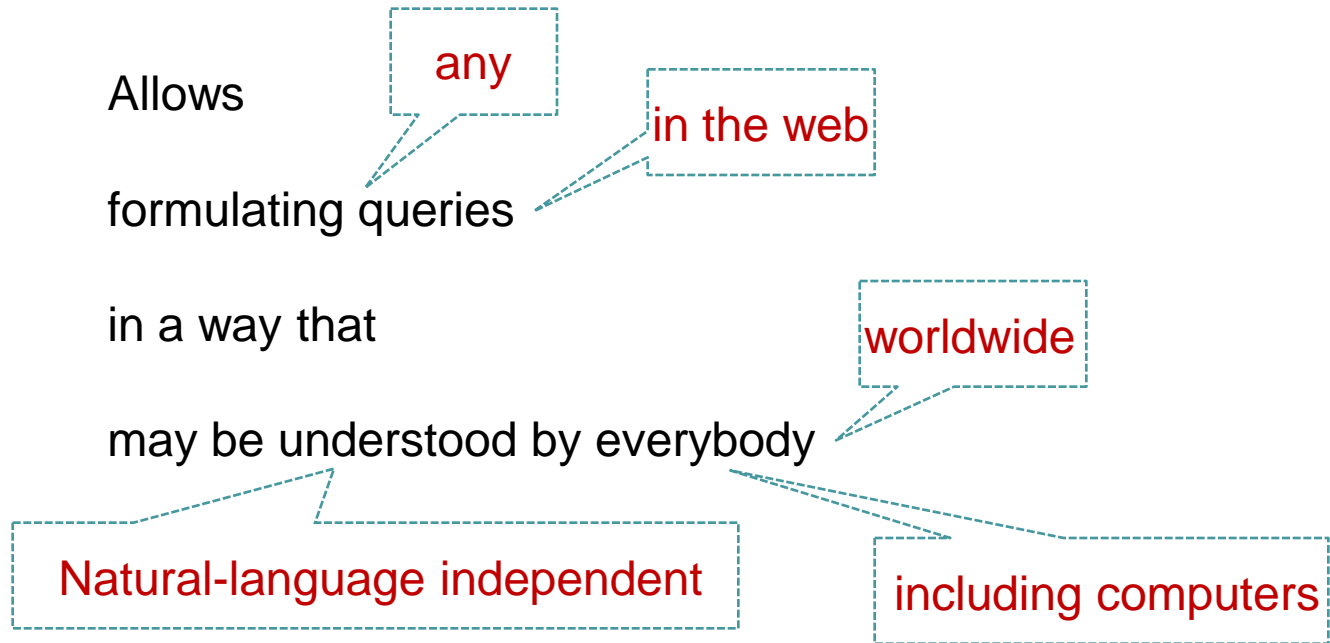
Uses of the Universal Ontology



The fact(s) is an instance of:

- a UO concept or
- a concept derived from the UO concepts

Uses of the Universal Ontology



What do we mean by the conjecture of the Universal Ontology?

The Universal Ontology will be a reality
in the foreseeable future

What people say...

Semantic Integration: A Survey Of Ontology-Based Approaches

(December 2004)

Natalya F. Noy

It is of course unrealistic to hope that there will be an agreement on one or even a small set of ontologies.

What people say...

Contextualizing ontologies[☆]

Paolo Bouquet^{a,b,*}, Fausto Giunchiglia^{a,b}, Frank van Harmelen^c,
Luciano Serafini^a, Heiner Stuckenschmidt^c

The need for terminology integration has been widely recognized in the medical area leading to a number of efforts for defining standardized terminologies. It is, however, also acknowledged by the literature, that the creation of a single universal terminology for the medical domain is neither possible nor beneficial, because different tasks and viewpoints require different, often incompatible conceptual choices [9]. As a result,

(2004)

SW Ontologies \neq a central, big ontology!

- The “ethos” of the Semantic Web is on sharing, ie, sharing ontologies (small or large)
- **A huge, central ontology would be unmanageable**
- The practice:
 - SW applications using ontologies always mix large number of ontologies and vocabularies (FOAF, DC, and others)
 - the real advantage comes from this mix: that is also how new relationships may be discovered

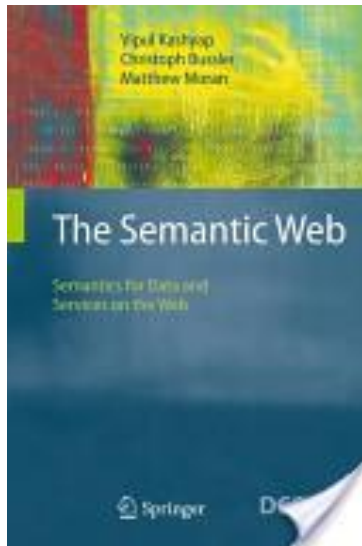
How could we prove the conjecture of the Universal Ontology?

- Define the minimal contents of the UO and its main uses.
- **Analyze feasibility:**
 - Can we build the UO?:
 - technically (ontology integration)
 - socially
 - Can we evolve the UO?:
 - technically
 - socially
- **Analyze viability:**
 - Ensure that the main uses of the UO have enough interest
 - Check that there is someone that has the resources needed to build and evolve the UO

What people say...

Why Should I share my knowledge?

Socio-economic factors dictate reality. **Who will develop the common ontologies and will they invest the effort and then allow them to be used for free?**



Introducing schema.org: Search engines come together for a richer web

6/02/2011 10:00:00 AM

(Cross-posted on the [Inside Search Blog](#))

Today we're announcing schema.org, a new initiative from Google, Bing and Yahoo! to create and support a common vocabulary for structured data markup on web pages. With schema.org, site owners and developers can learn about structured data and improve how their sites appear in major search engines. The site aims to be a one stop resource for webmasters looking to add markup to their pages.

On June 2nd we announced a collaboration between Bing, Google and Yahoo to create and support a standard set of schemas for structured data markup on web pages. Although our companies compete in many ways, it was evident to us that collaboration in this space would be good for each search engine individually and for the industry as a whole.

We want to continue making the open web richer and more useful. We know that it takes time and effort for webmasters to add this markup to their pages, and adding markup is much harder if every search engine asks for data in a different way. That's why we've come together with other search engines to support a common set of schemas, just as we came together to support a common standard for sitemaps in 2006. With schema.org, site owners can improve how their sites appear in search results not only on Google, but on Bing, Yahoo! and potentially other search engines as well in the future.

Organization of Schemas

The schemas are a set of 'types', each associated with a set of properties. The types are arranged in a hierarchy.

Browse the full hierarchy:

- [One page per type](#)
- [Full list of types, shown on one page](#)

Or you can jump directly to a commonly used type:

- Creative works: [CreativeWork](#), [Book](#), [Movie](#), [MusicRecording](#), [Recipe](#), [TVSeries](#) ...
- Embedded non-text objects: [AudioObject](#), [ImageObject](#), [VideoObject](#)
- [Event](#)
- [Organization](#)
- [Person](#)
- [Place](#), [LocalBusiness](#), [Restaurant](#) ...
- [Product](#), [Offer](#), [AggregateOffer](#)
- [Review](#), [AggregateRating](#)

We also have a small set of [primitive data types](#) for numbers, text, etc. More details about the data model, etc. are available [here](#).

Thing > Organization

An organization such as a school, NGO, corporation, club, etc.

Property	Expected Type	Description
Properties from Thing		
description	Text	A short description of the item.
image	URL	URL of an image of the item.
name	Text	The name of the item.
url	URL	URL of the item.
Properties from Organization		
address	PostalAddress	Physical address of the item.
aggregateRating	AggregateRating	The overall rating, based on a collection of reviews or ratings, of the item.
contactPoints	ContactPoint	A contact point for a person or organization.
email	Text	Email address.
employees	Person	People working for this organization.
events	Event	Upcoming or past events associated with this place or organization.
faxNumber	Text	The fax number.
founders	Person	A person who founded this organization.
foundingDate	Date	The date that this organization was founded.
interactionCount	Text	A count of a specific user interactions with this item—for example, 20 UserLikes , 5 UserComments , or 300 UserDownloads . The user interaction type should be one of the sub types of UserInteraction .
location	Place or PostalAddress	The location of the event or organization.
members	Person or Organization	A member of this organization.
reviews	Review	Review of the item.
telephone	Text	The telephone number.

More specific types

- [Corporation](#)
- [EducationalOrganization](#)
- [GovernmentOrganization](#)
- [LocalBusiness](#)
- [NGO](#)
- [PerformingGroup](#)
- [SportsTeam](#)

Introducing Schema.org

Google, Bing, and Yahoo! have announced the formation of schema.org. The current schema.org type hierarchy can be found on the organization's website [here](#). OpenCyc concepts corresponding to this hierarchy may be explored at sw.opencyc.org. See for example the entry for [movies](#).

- [Google's announcement](#)
- [Bing's announcement](#)
- [Yahoo!'s announcement](#)

WEDNESDAY, SEPTEMBER 28, 2011

Schema.org Workshop Wrap-Up

On September 21st, we held the first ever schema.org workshop in Mountain View, California. There were 75 attendees from web markup & standards groups (including W3C, Microformats, RDFa, Creative Commons), as well as other search engines (Ask, Yandex, Baidu), and top content publishers and tools vendors (including NYTimes, Disney, Foursquare, Shopping.com, OpenTable, Drupal, Sharepoint). The objective of this workshop was to evolve the schema.org specification, solicit feedback from the standards community, and build momentum for web publishers.

We felt the event was a success -- there was lots of enthusiasm around creating extensions to the vocabulary as well as great feedback on how to evolve the syntax. We will be working through this feedback in our working group over the coming months, so stay tuned for more developments, and thanks in advance for all your feedback.

At the workshop, we also announced a couple new schema.org developments. First, we announced that we have adopted IPTCs rNews specification into schema.org (see [earlier blog post](#)). In addition to being a fantastic vocabulary extension for news articles, we believe this a great template for future industry-specific vocabulary extensions, and there are a couple ongoing discussions for new extensions in education (working with LRMI) and for job postings (working with Whitehouse CTO). We expect

The type hierarchy presented on this site is not intended to be a 'global ontology' of the world. It only covers the types of entities for which we (Microsoft, Yahoo! and Google), think we can provide some special treatment for, through our search engine, in the near future.



[Everything Google](#) › [Corporate information](#) › Company

Overview

[Technology](#)

Google's mission is to organize the world's information and make it universally accessible and useful.

Assuming for the moment that
the conjecture of the Universal Ontology is true...

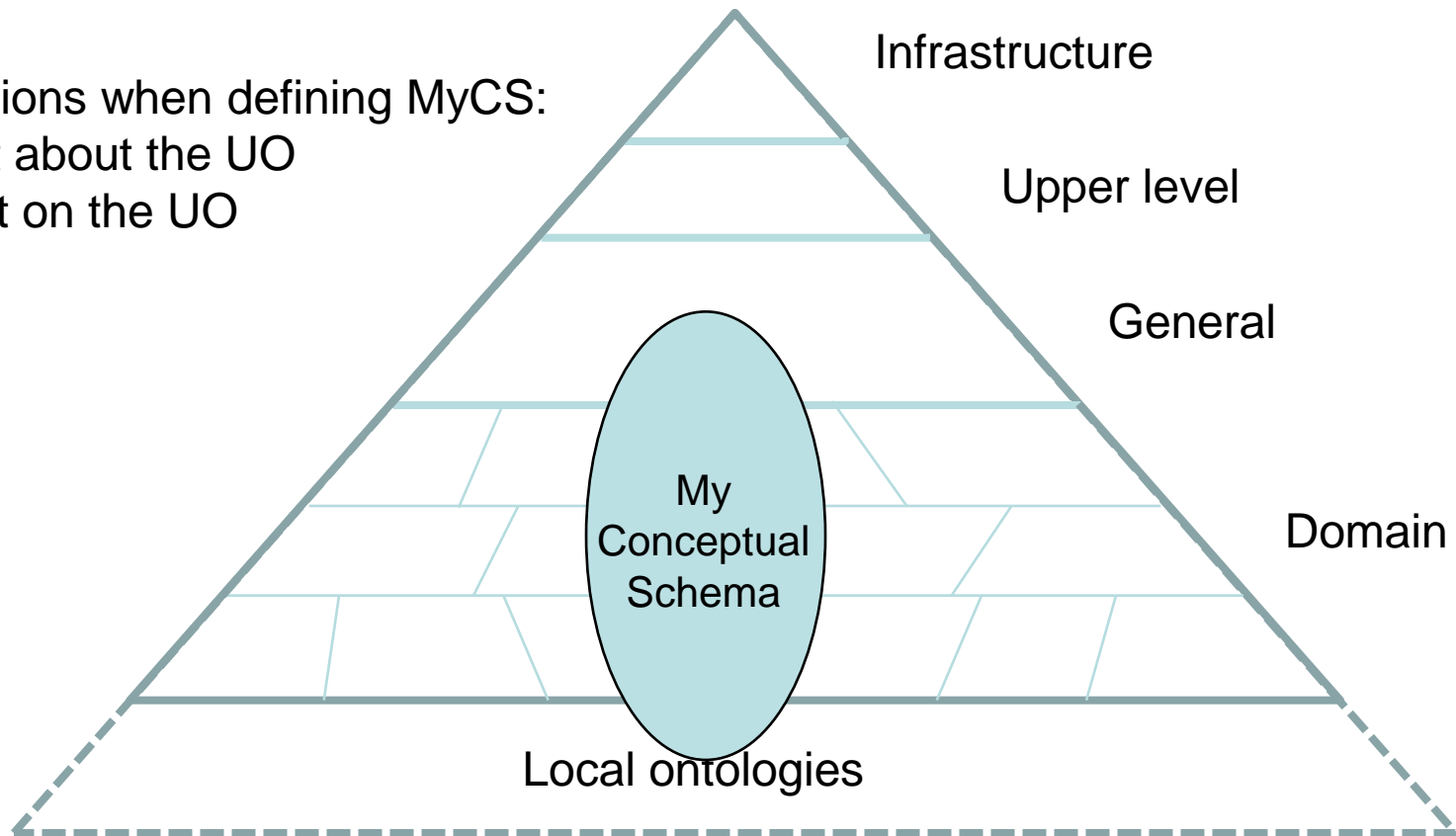
What are its implications for conceptual modeling?

Impact of the UO on Conceptual Modeling

Like it or not ...

Two options when defining MyCS:

- Forget about the UO
- Base it on the UO



Basing My Conceptual Schema on the UO

A concept of MyCS is:

- A UO concept
- A subtype of a UO concept
- A redefinition of a relationship type
- A derived Entity type
- A derived Relationship type
- A reification of a UO relationship type

It should be easy to:

- Find the relevant UO concepts
- Define MyCS from them

Basing My Conceptual Schema on the UO

Advantages:

- Eases universal understanding of MyCS
- Reuses existing concept definitions
- Eases system interoperability
- Eases service specification

Drawbacks:

- May be difficult to do

Three research questions:

What is/could be the universal ontology?

Can we prove the conjecture of the universal ontology?

If we can, then:

What are its implications for
conceptual modeling of information systems?