UNIVERSITY
OF TRENTO - Italy

KNOW
DIVE

Managing diversity in Knowledge

Fausto Giunchiglia

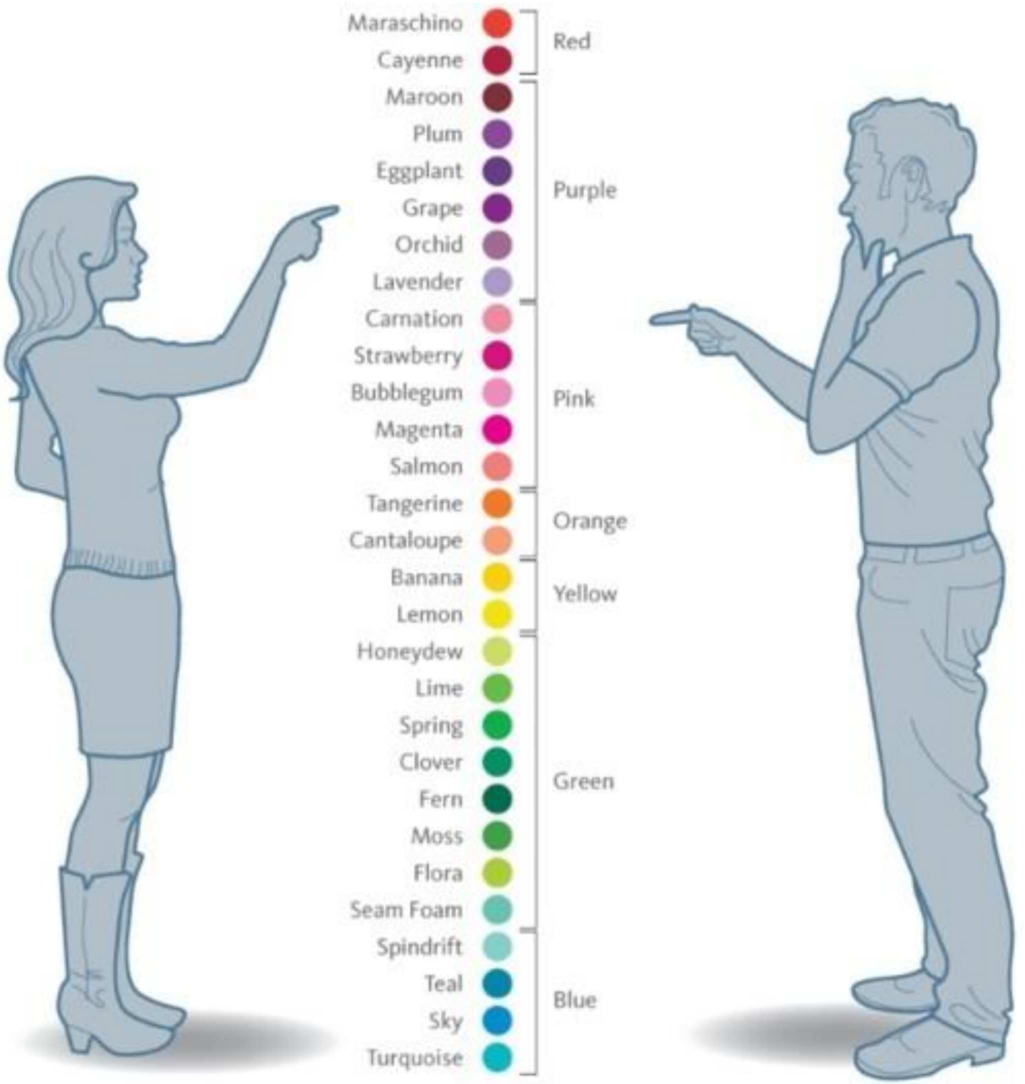
UKC
Universal Knowledge Core

Roadmap

- **Motivation and use-cases**
- **Existing approaches**
- **Diversity in Knowledge**
- **Entiy Centric representation of the world**
- **UKC**
- **Entitypedia**
- **The DERA methodology**

Motivation and use-cases

Diversity (was semantic heterogeneity)



The difficulty of establishing a certain level of connectivity between people, software agents or IT systems [Uschold & Gruninger, 2004] at the purpose of enabling each of the parties to appropriately *understand* the exchanged information [Pollock, 2002]

Use-cases

SEMANTIC SEARCH

SEARCH: automobile

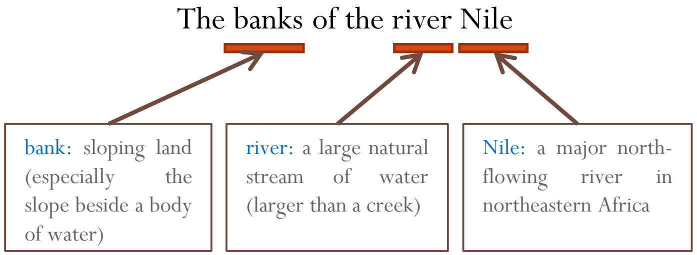
1957 Ferrari 625 TRC Spider

This two-of-a-kind classic Ferrari is lauded by historians as one of the prettiest Ferraris ever built. The 1957 Ferrari 625 TRC Spider is an absolutely stunning automobile, one as dashing in the garage as it is at 120 mph.

Back in the Saddle: Presenting our Porsche 911 (997) Carrera S Cabriolet

There's a reason the Porsche 911 is one of the most popular sports cars ever, and after a few minutes behind the wheel of one you'll understand why.

NLP



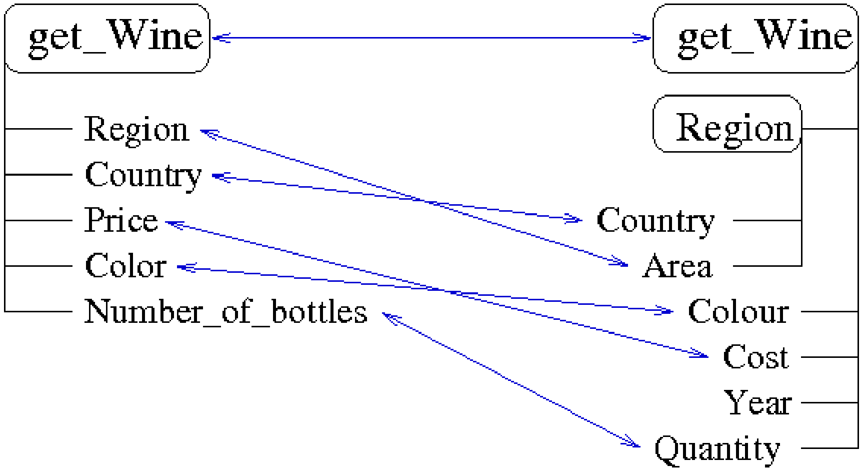
The banks of the river Nile

bank: sloping land (especially the slope beside a body of water)

river: a large natural stream of water (larger than a creek)

Nile: a major north-flowing river in northeastern Africa

SEMANTIC MATCHING



get_Wine

Region

Country

Price

Color

Number_of_bottles

get_Wine

Region

Country

Area

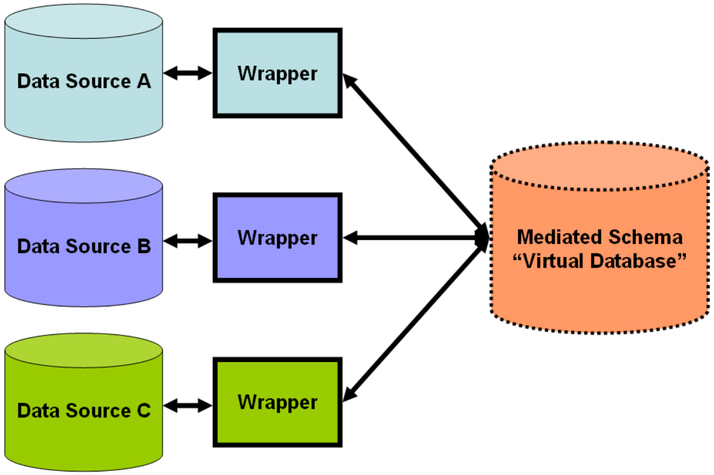
Colour

Cost

Year

Quantity

DATA INTEGRATION



Data Source A

Wrapper

Data Source B

Wrapper

Data Source C

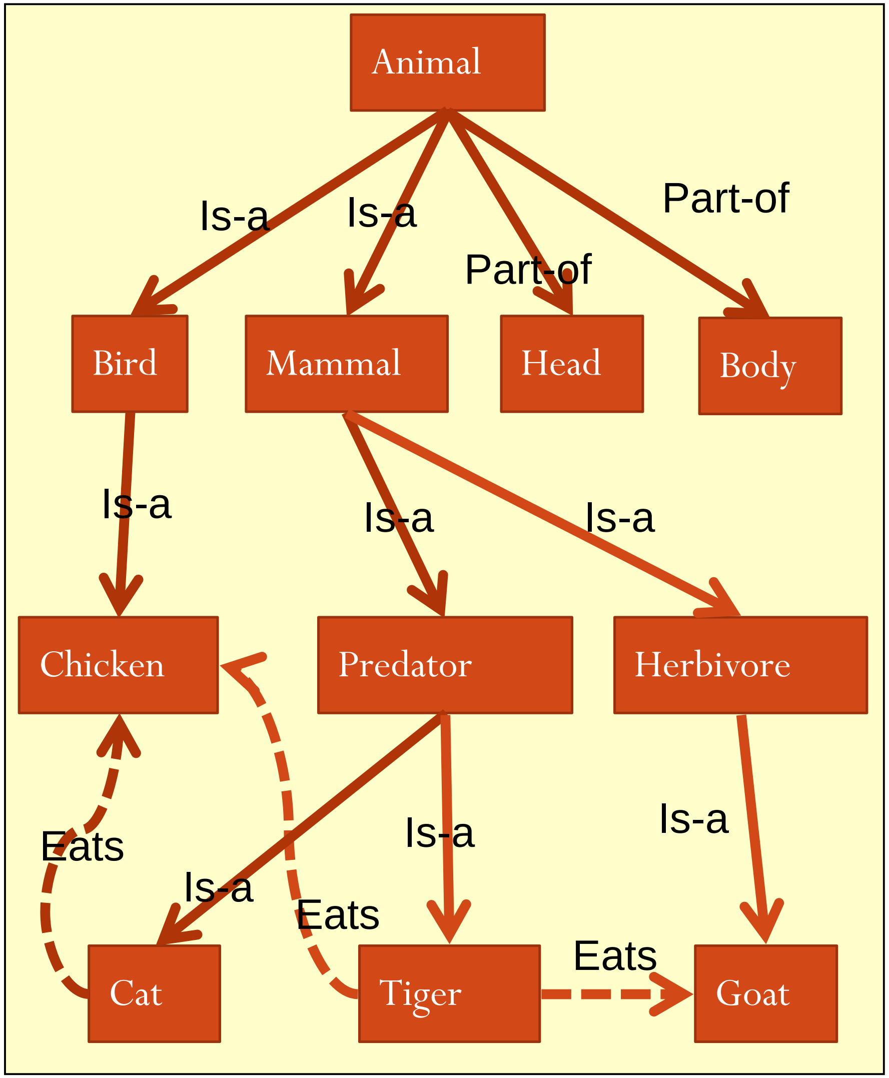
Wrapper

Mediated Schema
"Virtual Database"

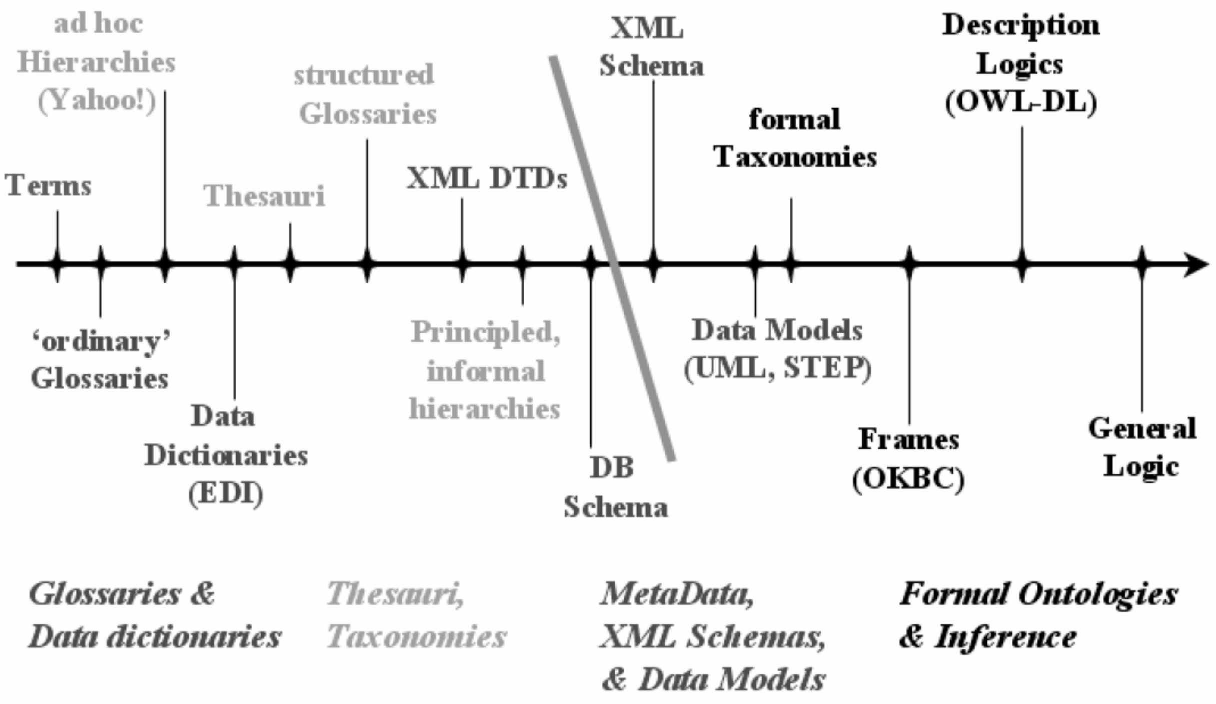
Existing approaches

Ontologies

- An ontology is an explicit specification of a shared conceptualization [Gruber, 1993]
- Ontologies are often thought of as directed graphs whose nodes represent concepts and whose edges represent relations between concepts
- By providing a common formal terminology and understanding of a given domain of interest, it allows for automation (logical inference), supports reuse and favor interoperability across applications and people.
- They differ according to the purpose and the semantics



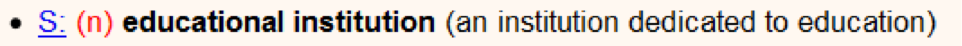
Kinds of ontologies



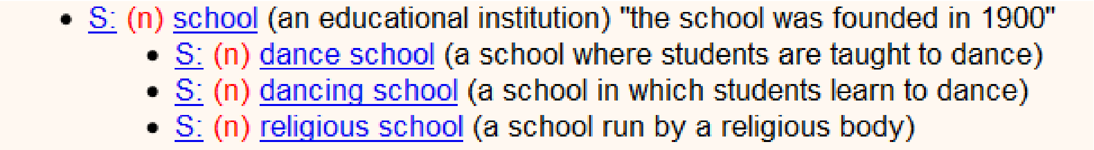
[Uschold and Gruninger, 2004]

- Informal representations
  - User classification
  - Web directories
  - Business catalogs
- Progressive formal
  - Enumerative (e.g. DDC)
  - Knowledge Organization systems
  - Faceted Classification systems
- Formal ontologies
  - Expressed into a formal logic language and represented using formal specifications, such as, OWL)

(Problems with) WordNet

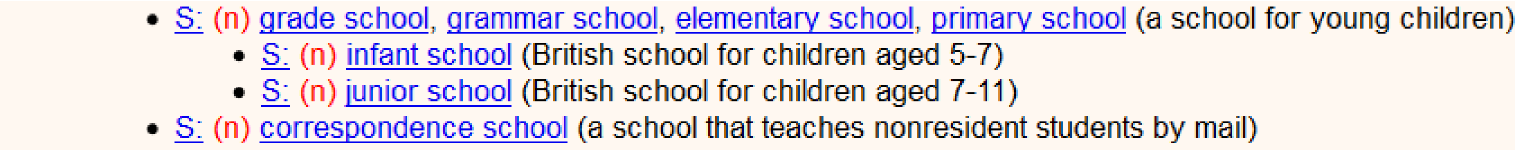


- S: (n) **educational institution** (an institution dedicated to education)



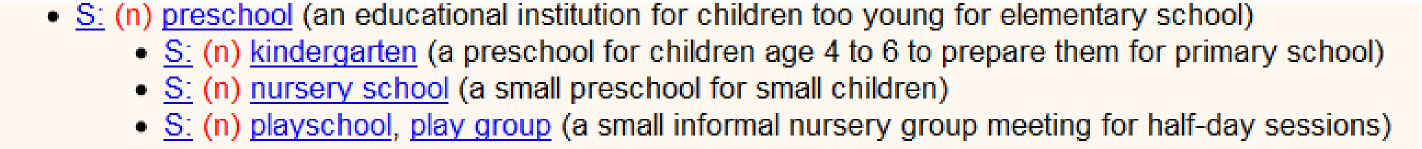
- S: (n) school (an educational institution) "the school was founded in 1900"
  - S: (n) dance school (a school where students are taught to dance)
  - S: (n) dancing school (a school in which students learn to dance)
  - S: (n) religious school (a school run by a religious body)

The position of nodes is driven by syntax



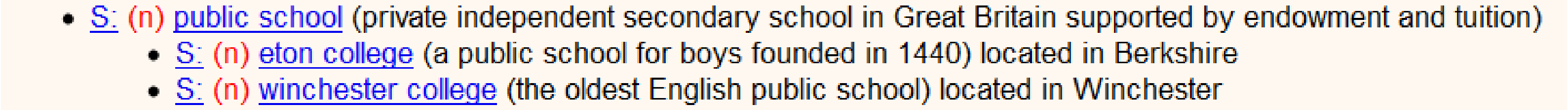
- S: (n) grade school, grammar school, elementary school, primary school (a school for young children)
  - S: (n) infant school (British school for children aged 5-7)
  - S: (n) junior school (British school for children aged 7-11)
- S: (n) correspondence school (a school that teaches nonresident students by mail)

Glosses exhibit space and time bias



- S: (n) preschool (an educational institution for children too young for elementary school)
  - S: (n) kindergarten (a preschool for children age 4 to 6 to prepare them for primary school)
  - S: (n) nursery school (a small preschool for small children)
  - S: (n) playschool, play group (a small informal nursery group meeting for half-day sessions)

Some concepts are too similar in meaning



- S: (n) public school (private independent secondary school in Great Britain supported by endowment and tuition)
  - S: (n) eton college (a public school for boys founded in 1440) located in Berkshire
  - S: (n) winchester college (the oldest English public school) located in Winchester

Some concepts are actually individuals

Diversity in knowledge

Diversity is pervasive in world descriptions

In language

- “watercourse” in English is same as “corso d’acqua” in Italian (*concepts*)
- There is no lemma in Italian for “biking” (*lexical GAP*)

In meaning (concepts)

- “bug as malfunction” vs. “bug as food” *(homonymy)*
- “stream” and “watercourse” have same meaning *(synonymy)*

In (schematic) knowledge

- There are several types of bodies of water (*semantic relations*)
- Rivers have a length, lakes have a depth (*schematic knowledge*)

In (ground) knowledge (= data)

- The Adige river is 410 Km long; The Garda lake is 136 m deep

In opinions and viewpoints

- “Bugs are great food” vs. “how can you eat bugs?” *(the role of culture)*
- “Climate is/is not an important issue” *(the role of schools of thought)*

Diversity in language

Language	Number of native speakers
Mandarin Chinese	~ 880 M
Spanish	~ 325 M
English	~315-380 M
Arab	~205-425 M
Hindi	~185 M
Portuguese	~180 M
Bengali	~ 175 M
Russian	~145 M
Japanese	~130 M
German	~95 M

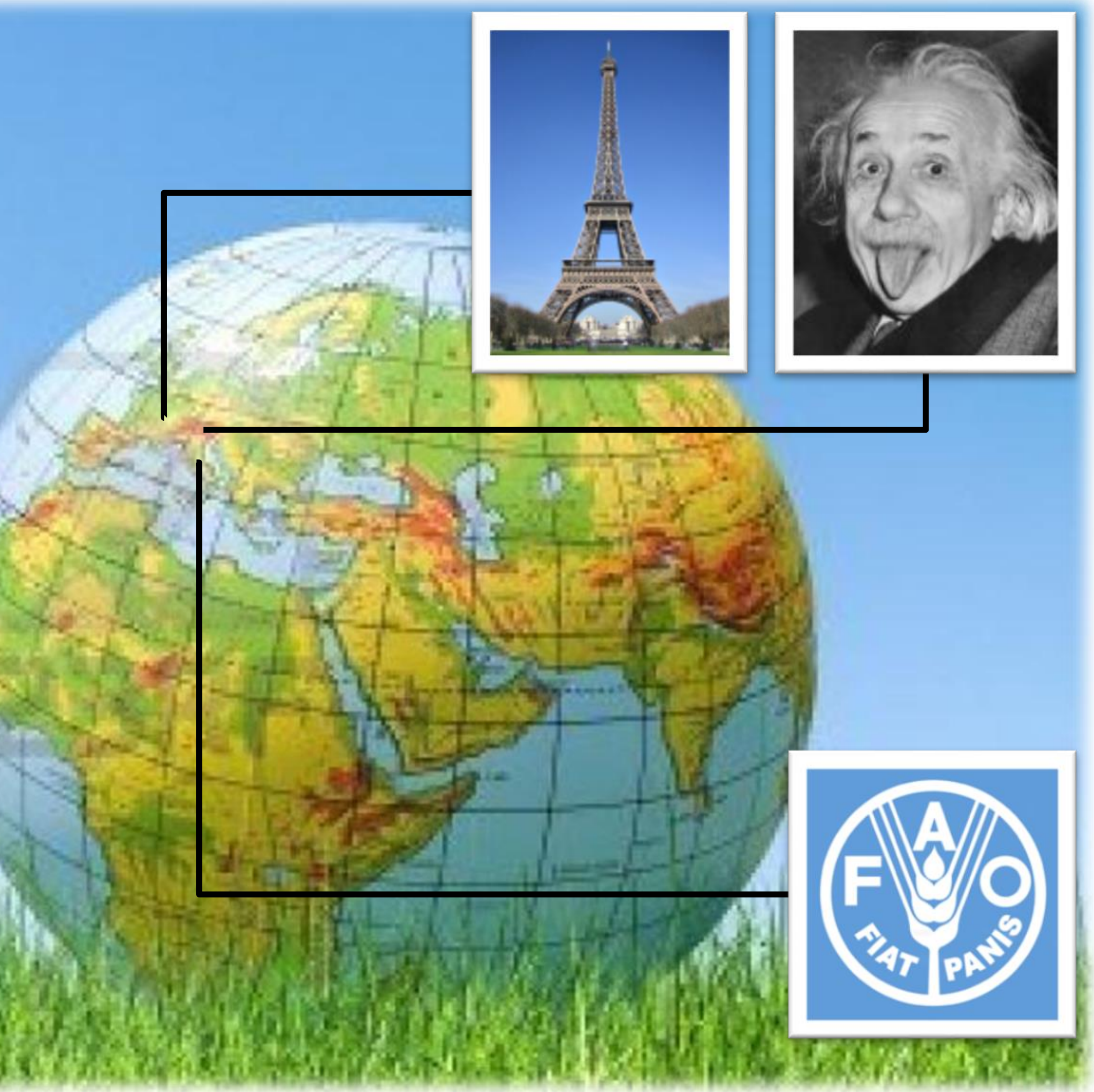
- **Around 200 countries**
- **More than 6800 spoken languages**
- **94% of the languages is spoken by 6% of the world population**
- **234 languages in Europe**

ENGLISH dictionaries

- **More than 170.000 words**
- **More than 110.000 different meanings**

Diversity in Knowledge

- Billions of locations
- Billions of people
- Millions of organizations
- … and events, artifacts, creative works, …



Entity centric representation of Diversity

AN ENTITY-CENTRIC VISION OF THE WORLD

What is an entity?

- **Entities** are objects which are so important in our everyday life to be referred with a proper name
- Each entity is described by its own attributes (e.g. latitude, longitude, height…)
- Each entity is described in relation with other entities (e.g. Eiffel Tower is located in Paris, France)
- Each entity as a reference class (e.g. monument) which determines its entity type (e.g. location)

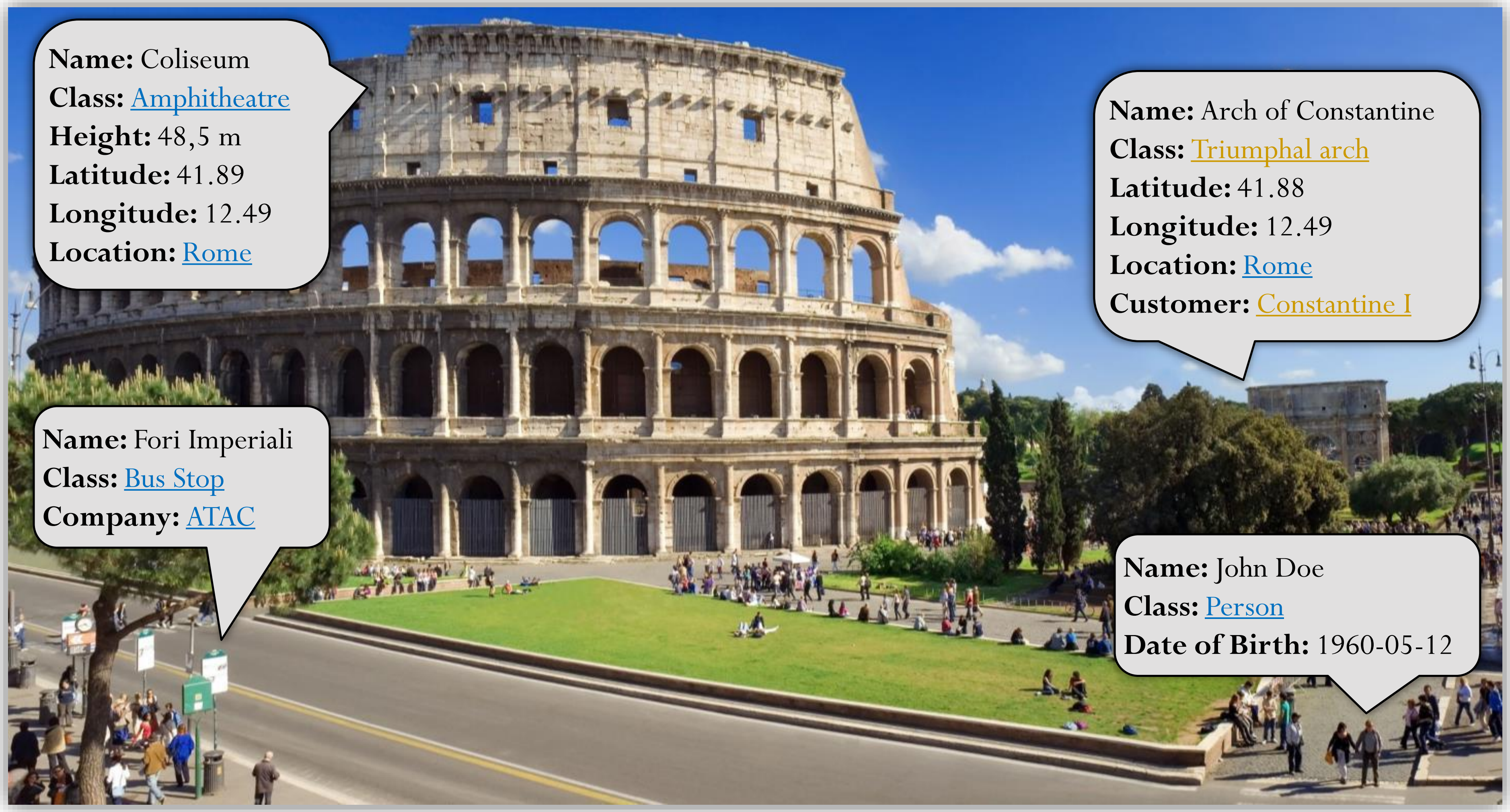
Eiffel Tower

How to represent an entity?

Entity Class	Class:	Monument
Attributes	Name:	Eiffel Tower
	Latitude:	48.86
	Longitude:	2.29
	Height:	324 m
Relations	Part-of:	Paris (France)

Eiffel Tower

What do we aim to? How to achieve that?



Name: Coliseum
Class: Amphitheatre
Height: 48,5 m
Latitude: 41.89
Longitude: 12.49
Location: Rome

Name: Arch of Constantine
Class: Triumphal arch
Latitude: 41.88
Longitude: 12.49
Location: Rome
Customer: Constantine I

Name: Fori Imperiali
Class: Bus Stop
Company: ATAC

Name: John Doe
Class: Person
Date of Birth: 1960-05-12

UKC (Universal Knowledge Core)

Roadmap

- **Motivation and use-cases**
- **Existing approaches**
- **Diversity in Knowledge**
- **Entiy Centric representation of the world**
- **UKC**
- **Entitypedia**
- **The DERA methodology**

Diversity is pervasive in world descriptions

In language

- “watercourse” in English is same as “corso d’acqua” in Italian (*concepts*)
- There is no lemma in Italian for “biking” (*lexical GAP*)

In meaning (concepts)

- “bug as malfunction” vs. “bug as food” *(homonymy)*
- “stream” and “watercourse” have same meaning *(synonymy)*

In (schematic) knowledge

- There are several types of bodies of water (*semantic relations*)
- Rivers have a length, lakes have a depth (*schematic knowledge*)

In (ground) knowledge (= data)

- The Adige river is 410 Km long; The Garda lake is 136 m deep

In opinions and viewpoints

- “Bugs are great food” vs. “how can you eat bugs?” *(the role of culture)*
- “Climate is/is not an important issue” *(the role of schools of thought)*

Codifying language: the UKC

- **The formal language:** a core of concepts (200 k+)
- **The natural language:** vocabulary of words for each language (200 x 200k+)
- **Schematic Knowledge:** a schema describing the structure of entities (hundreds)
- **Domain knowledge:** terminology organized into domains (hundreds)

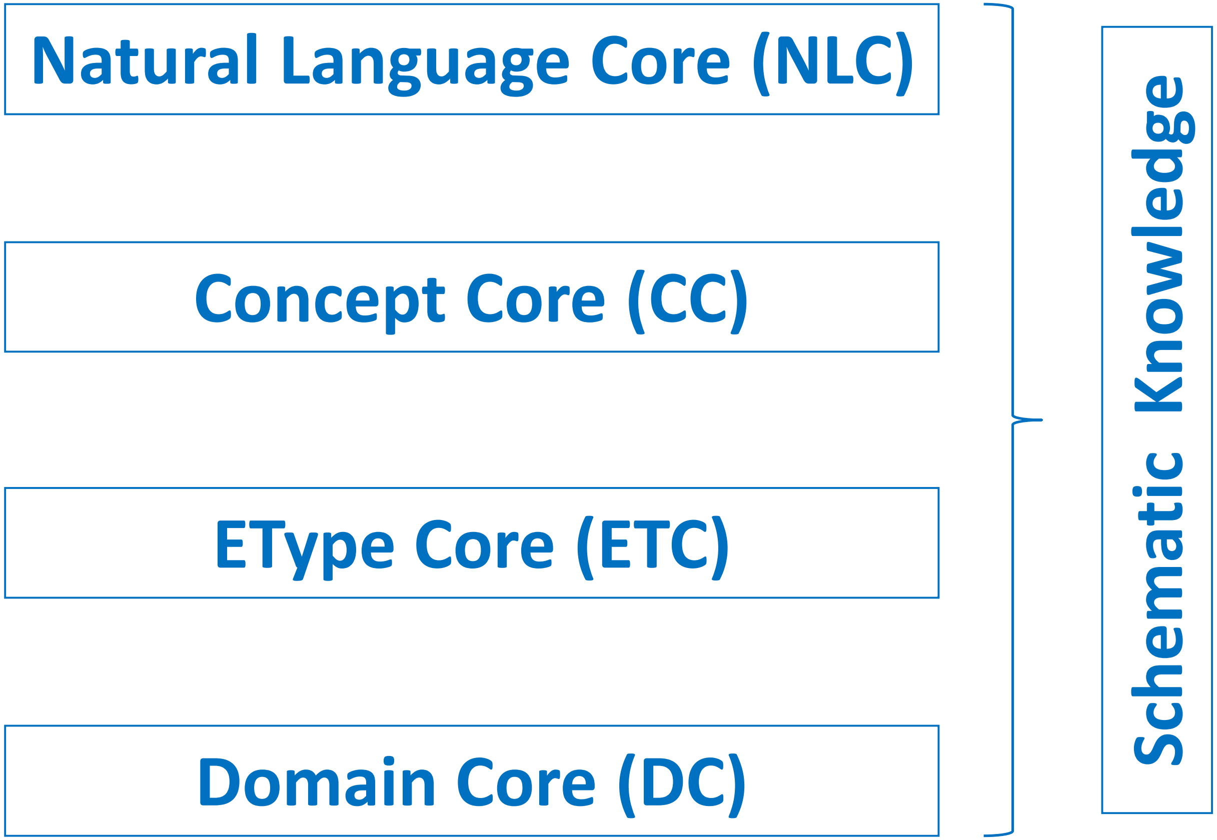
The UKC components

The natural language:
our vocabulary in multiple languages

The fomal language:
our graph of language-independent notions

Schematic knowledge:
Our schema of basic entity types

Domain knowledge:
Domain-specific partition of the language above



Natural Language Core (NLC)

Concept Core (CC)

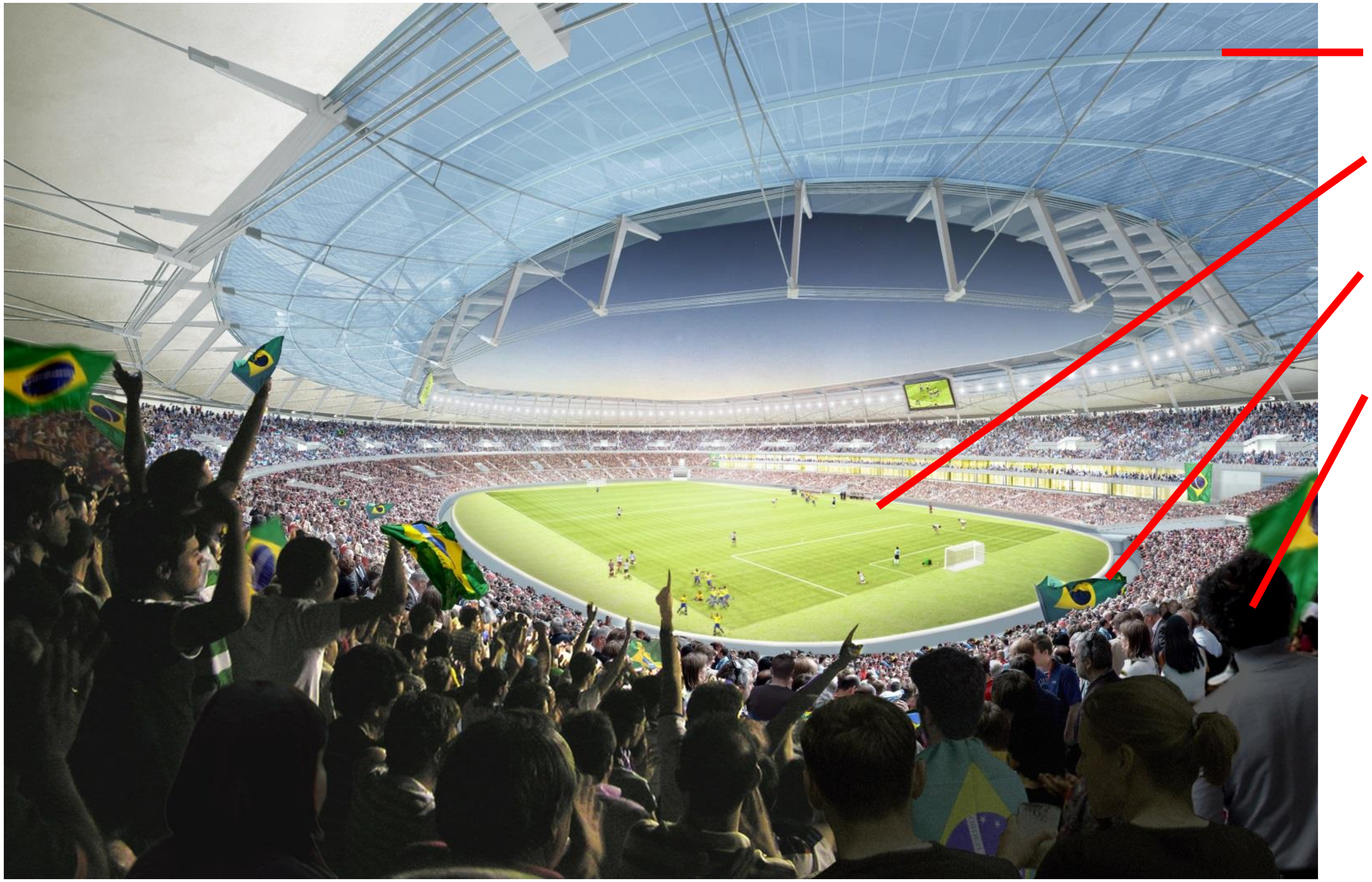
EType Core (ETC)

Domain Core (DC)

Schematic Knowledge

UKC: Entity types

Entities are of different types



location

event

organization

person

…

Entities are not all the same; they have different metadata according to the type of entity

The Etype Core (ETC)

ETPYE Each etype is a sort of template for the entities of that kind

ATTRIBUTES Each etype describes a set of mandatory and optional attributes with corresponding data type

RELATIONS Relations are special attributes that connect entities *(e.g. author connects a person with a document)*

- Controlling the terminology: each attribute name and attribute value is mapped to a concept in the Concept Core
- Inheritance of properties: etypes are arranged into a lattice

Examples of etypes

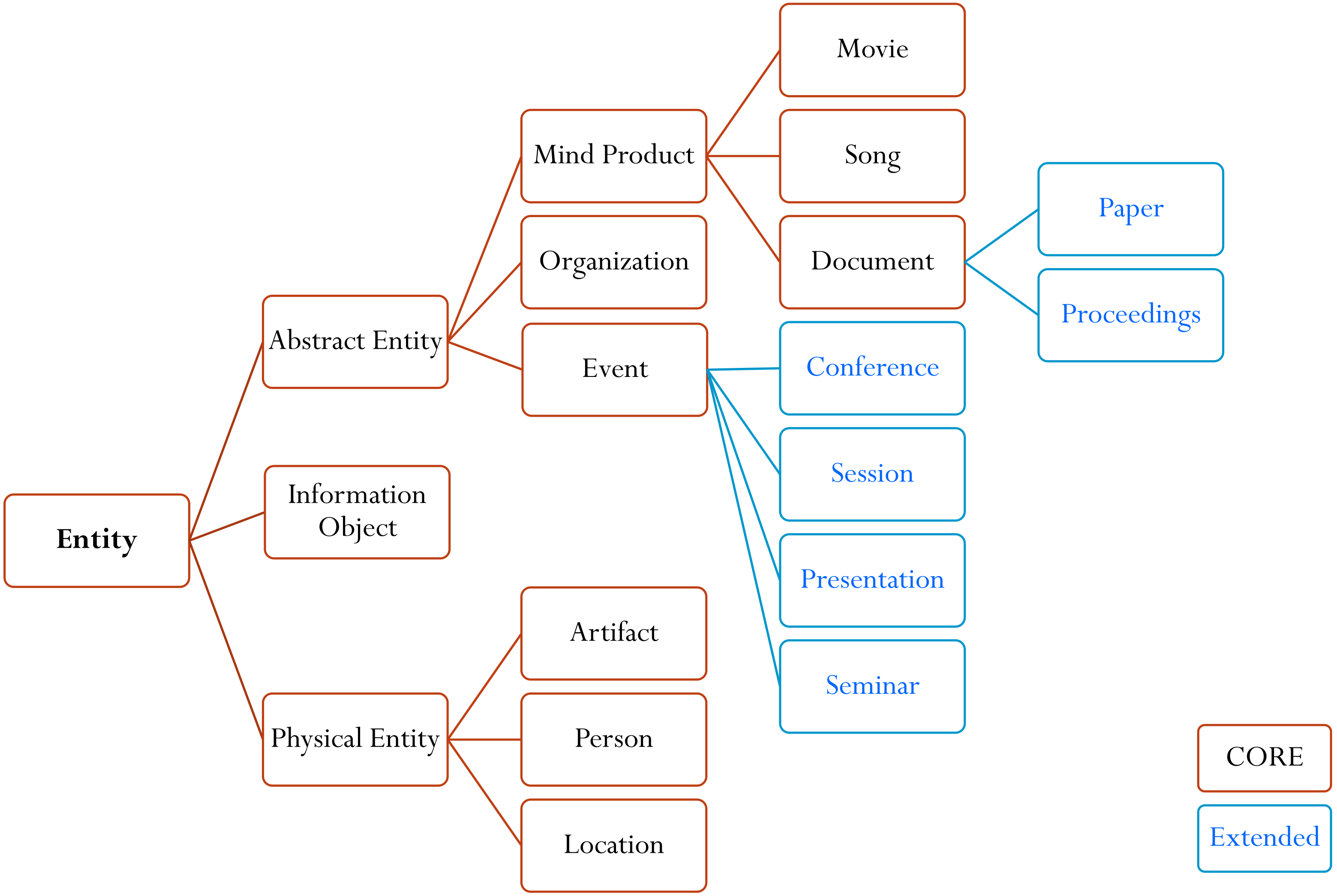
ENTITY

Name:	String [ ]
Class:	Concept [ ]
Description:	SString [ ]
Part-Of:	<Entity>
Start:	Date
End:	Date
Duration:	Duration

LOCATION extends ENTITY

Latitude:	float
Longitude:	float
Altitude:	float
…	

Etype lattice (exemplified)



Entity
Abstract Entity
Information Object
Physical Entity
Mind Product
Organization
Event
Artifact
Person
Location
Movie
Song
Document
Conference
Session
Presentation
Seminar
Paper
Proceedings

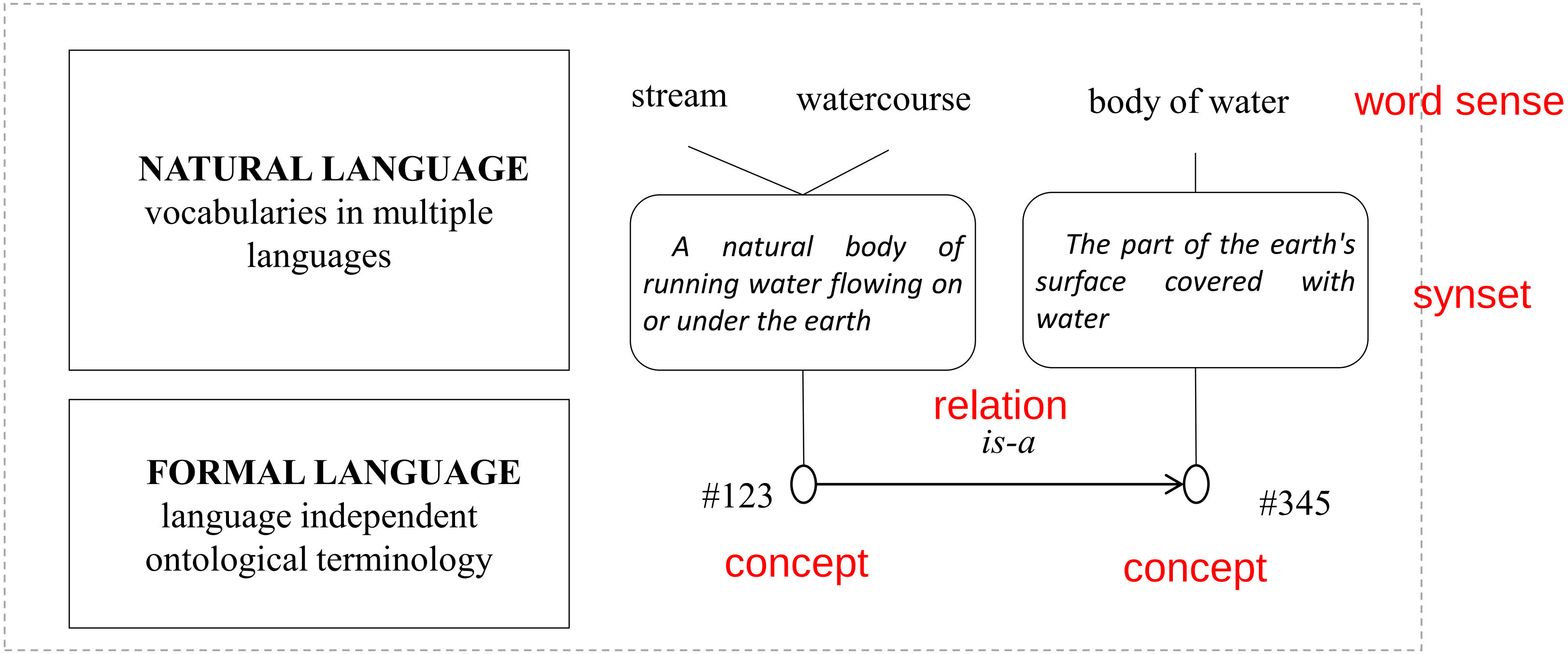
CORE
Extended

UKC: Natural and formal language

formal/natural language (thanks Wordnet!)

Natural language: terms, synonyms, synsets, lexical relations in multiple languages

Formal language: concepts and semantic relations between them


NATURAL LANGUAGE
vocabularies in multiple languages
FORMAL LANGUAGE
language independent ontological terminology
stream
watercourse
body of water
word sense
A natural body of running water flowing on or under the earth
The part of the earth's surface covered with water
synset
relation
is-a
#123
#345
concept
concept

Terminology

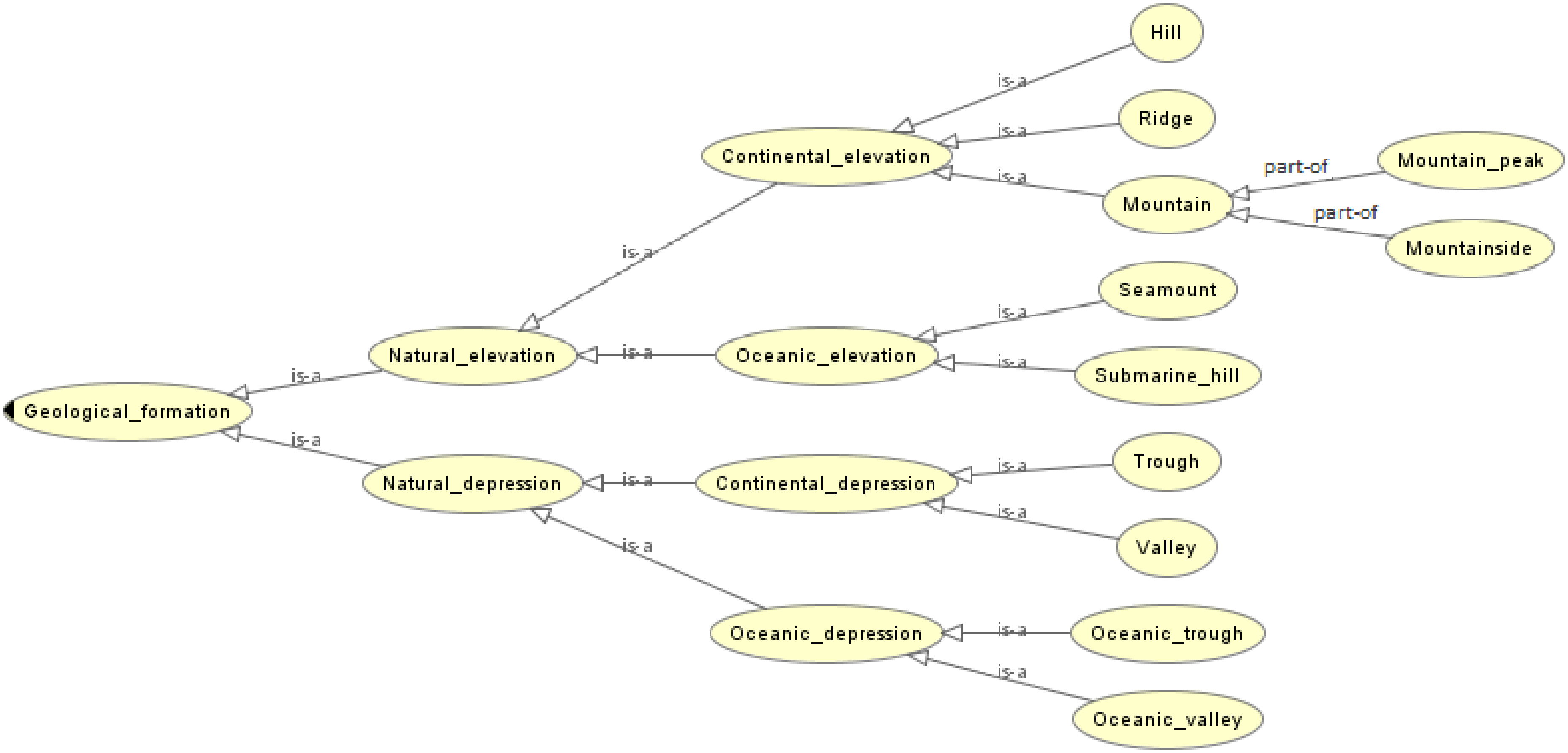
CONCEPT	A concept is a language independent representation of a group of things of the same kind
SEMANTIC RELATION	The relationship between concepts is established with semantic relation
WORD	A basic lexical unit that has a meaning on its own in a given language
SYNSET	A set of words that share the same meaning in a given context
LEXICAL RELATION	It establishes connection between linguistic elements

Concept

Geological formation
Natural depression
Oceanic depression
Oceanic valley
Oceanic trough
Continental depression
Trough
Valley

Natural elevation
Oceanic elevation
Seamount
Submarine hill
Continental elevation
Hill
Mountain
Ridge

Concept Hierarchy



The UKC is multilingual

Language	Synset	Gloss
en	Ridge	A long narrow natural elevation or stiration
mn	нуруу	урт нарийхан байгалийн өндөрлөг эсвэл ховил
bn	সেতুবন্ধ	দীর্ঘ সরু প্রাকৃতিক উঁচু ভূখণ্ড অথবা কোনো গিরিখাদের দুইপাশের উঁচু অংশ

Language	Synset	Gloss
en	Rivulet	A small stream
mn	GAP	

Language	Synset	Gloss
en	Oxbow lake	A crescent-shaped lake
bn	GAP	

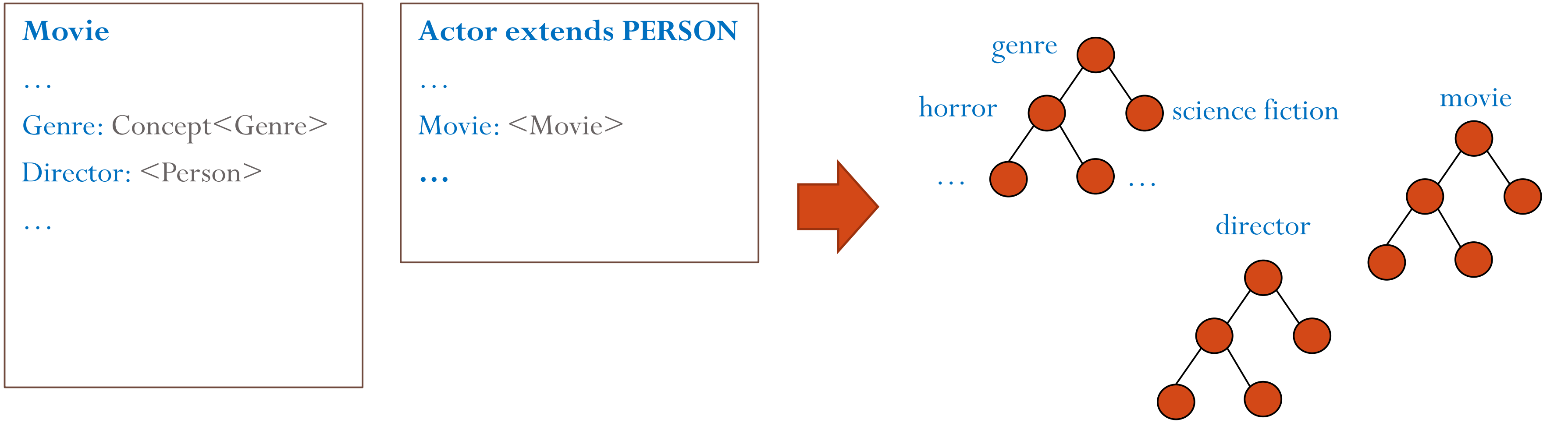
UKC: Domains

The Domain Core (DC)

- **A domain is the terminology which is needed to describe the entities that are relevant for the domain**
- **We capture domains as a set of relevant entity types selected from the etype core and corresponding terminology selected from the concept core**

The etypes in the Movie domain

The corresponding domain terminology



Movie

…

Genre: Concept<Genre>

Director: <Person>

…

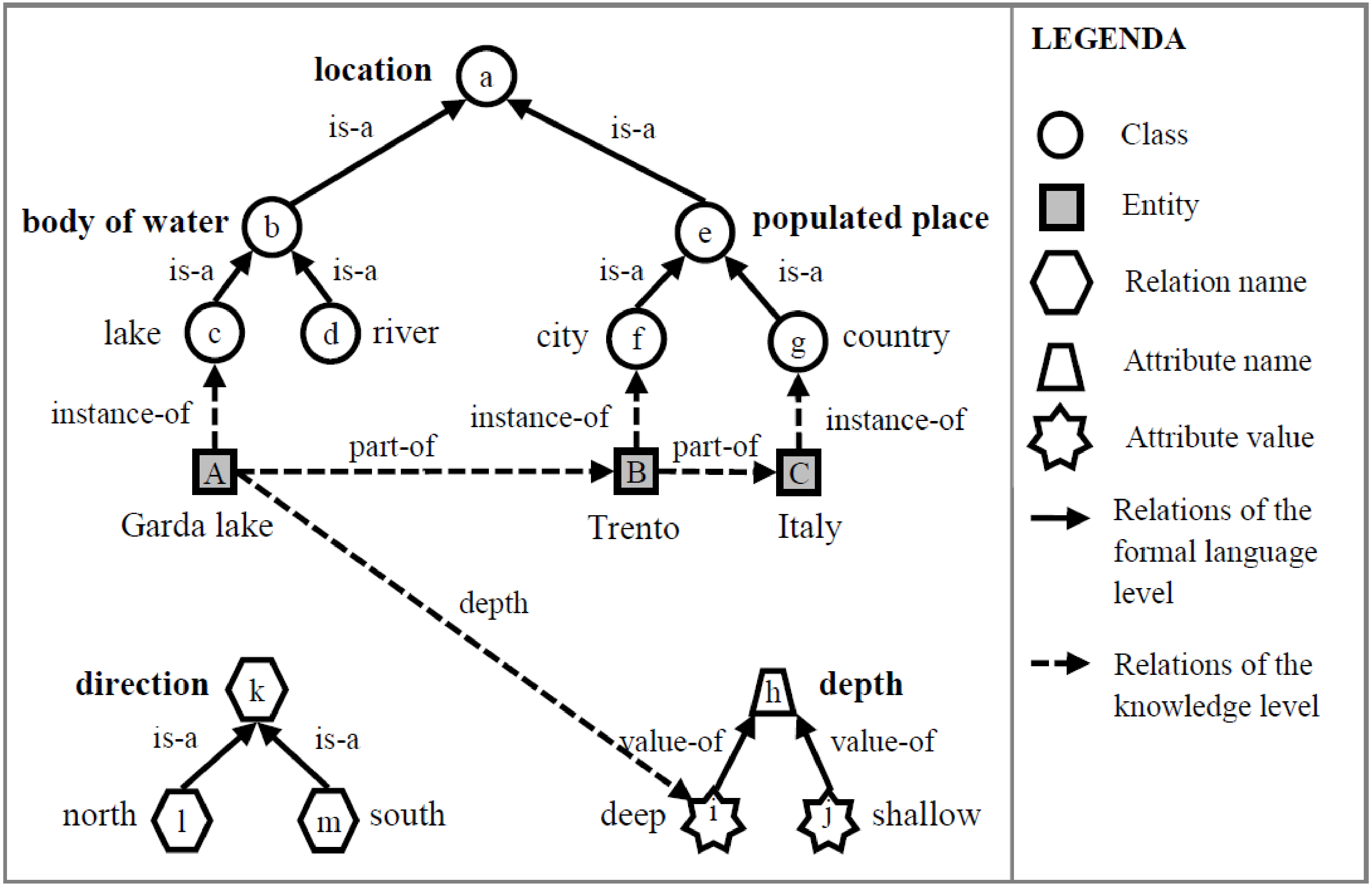
Actor extends PERSON

…

Movie: <Movie>

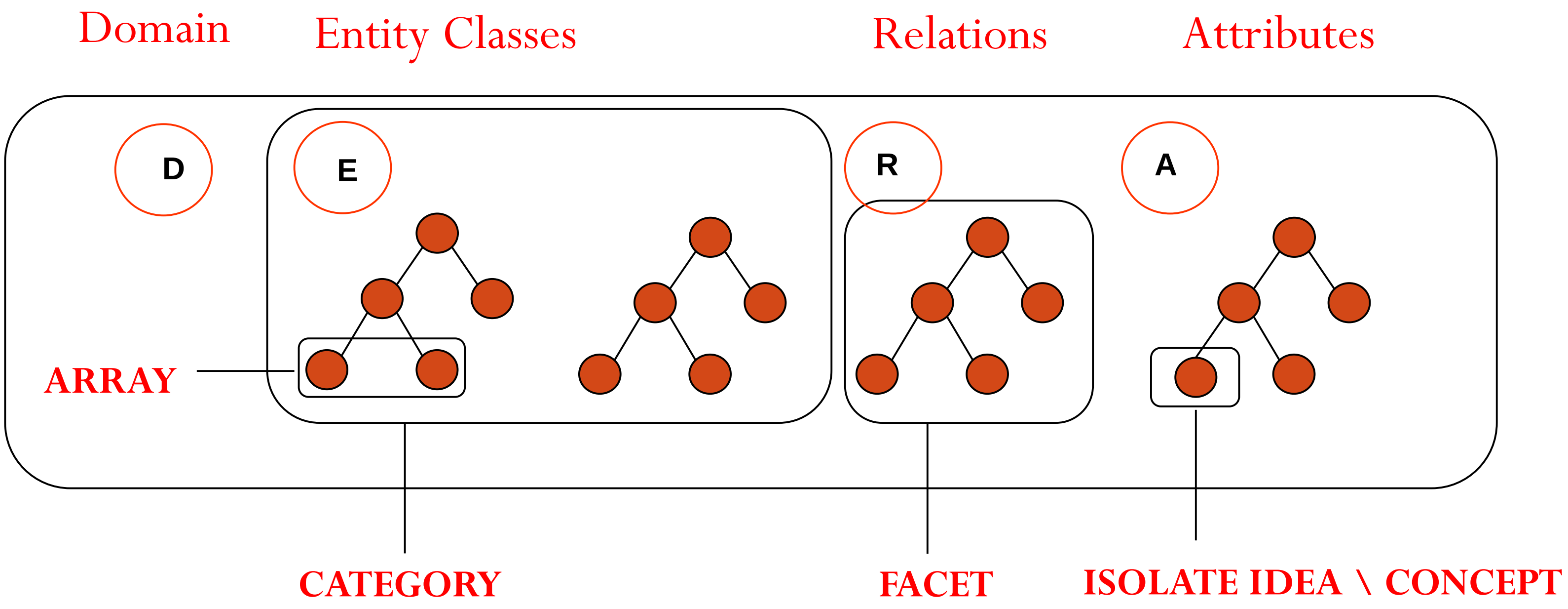
…

Descriptive ontologies [Giunchiglia et al, 2012b]



DERA [Giunchiglia et al., 2014]

- **How to build high quality and scalable descriptive ontologies?**
- **DERA is faceted as it is inspired to the principles and canons of the faceted approach by Ranganathan**
- **DERA is a KR approach as it models entities of a domain (D) by their entity classes (E), relations (R) and attributes (A)**

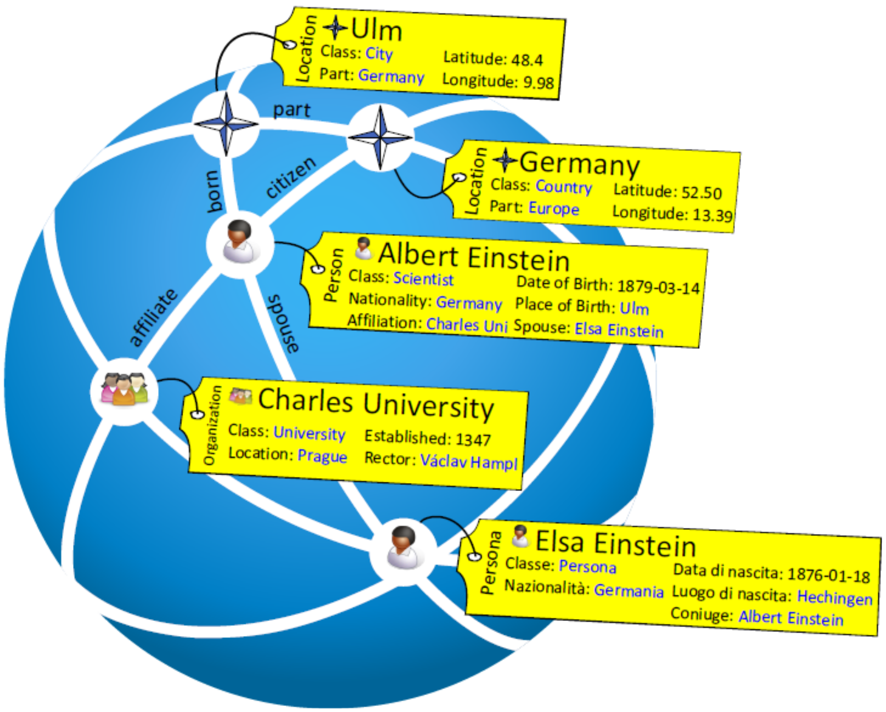


Entitypedia

Roadmap

- **Motivation and use-cases**
- **Existing approaches**
- **Diversity in Knowledge**
- **Entiy Centric representation of the world**
- **UKC**
- **Entitypedia**
- **The DERA methodology**

Entitypedia (UKC + entities)



Location
Ulm
Class: City
Part: Germany
Latitude: 48.4
Longitude: 9.98

Location
Germany
Class: Country
Part: Europe
Latitude: 52.50
Longitude: 13.39

Person
Albert Einstein
Class: Scientist
Nationality: Germany
Affiliation: Charles Uni
Date of Birth: 1879-03-14
Place of Birth: Ulm
Spouse: Elsa Einstein

Organization
Charles University
Class: University
Location: Prague
Established: 1347
Rector: Václav Hampl

Persona
Elsa Einstein
Classe: Persona
Nazionalità: Germania
Data di nascita: 1876-01-18
Luogo di nascita: Hechingen
Coniuge: Albert Einstein

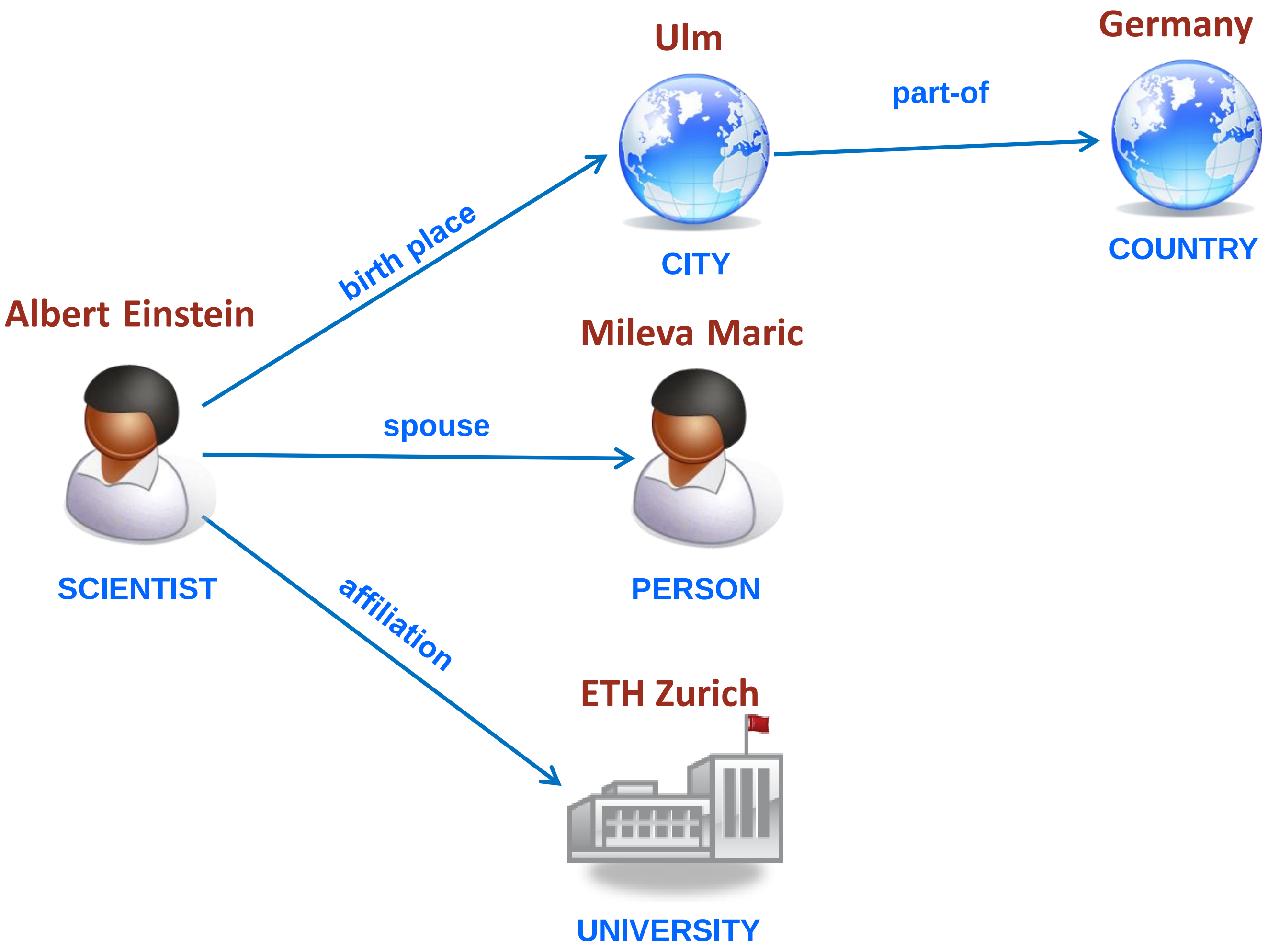
Very accurate multilingual entity base

- **Entities of different types** (e.g. location, person, organization)
- **Domains:** entities in context (e.g. Music, Sport, Politics)
- **Multi-language** (e.g. English, Italian)
- **Data quality and certification** guaranteed via a set of semi-automatic tools and expert maintenance
- **Dedicated communities**
- **Dedicated data acess APIs**

Data sources so far

- Wordnet (English) and MultiWordNet (Italian)
- GeoWordNet → 8M locations
- PAT → 20k locations
- YAGO → 700k persons, 150k organizations, 300k locations (selected and cleaned)

Example of entities

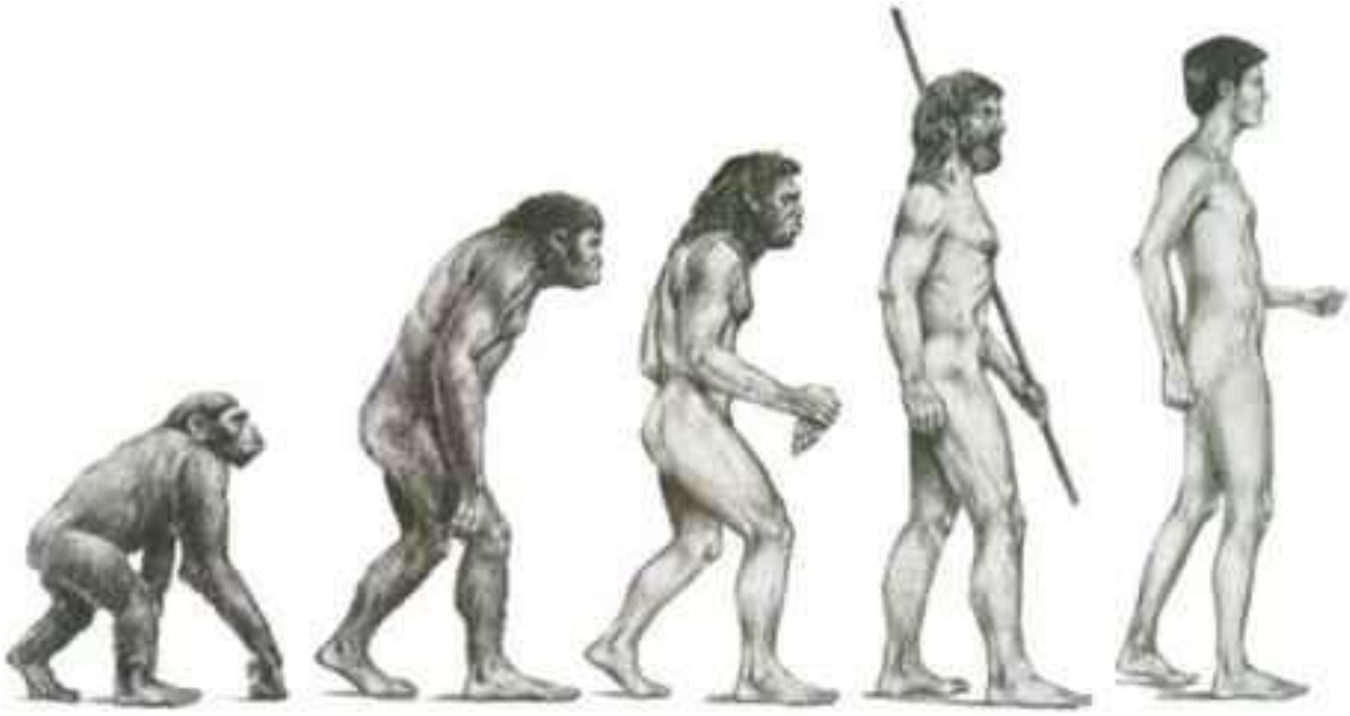


TBox

ABox

The DERA methodology

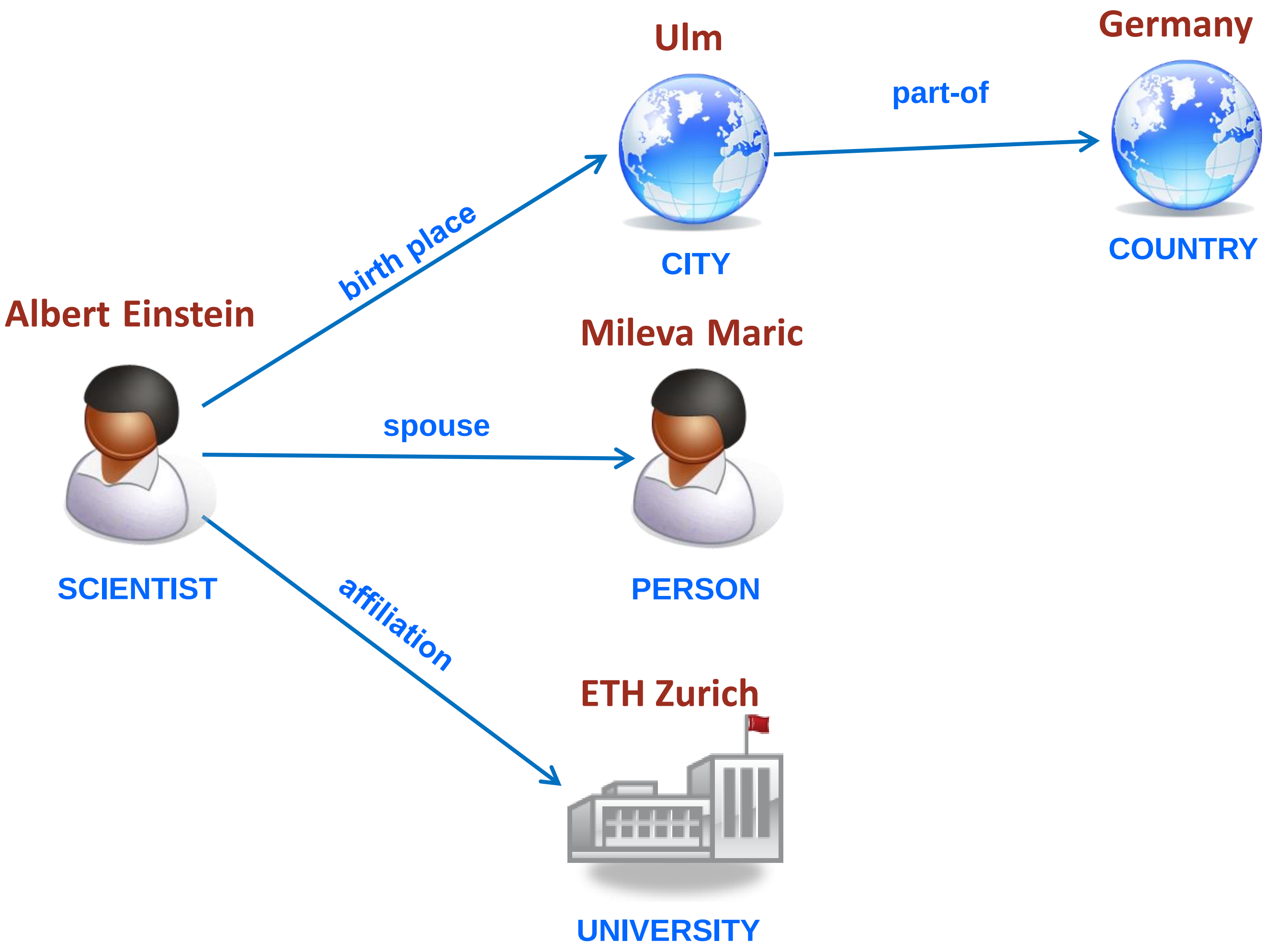
WHY DO WE NEED A METHODOLOGY?
BECAUSE SMALL DIFFERENCES MATTER…



Humans and chimps share a surprising 98.8 percent of their DNA.

How to build ontologies which are of the highest quality possible?

Example of entities



TBox

ABox

Back to entities

Entity Class	Class:	River
Attributes	Name:	Thames
	Latitude:	51.50
	Longitude:	0.61
	Length:	346 km (long)
Relations	Part-of:	UK

Thames river

Each of the terms above comes from a DERA ontology in KB

Data fragmentation

Courses

ID	Professor	Course	Year
05	Fausto Giunchiglia	Logic	2010

Research papers

ID	Title	Author	Subject
09	Theory of Contexts	F. Giunchiglia	AI

Projects

ID	Project	Coordinator
35	Smart Society	Fausto Giunchiglia

Exams

ID	Student	Course	Mark
09	Mary Chen	Logics	28

- **Data come from different sources**
- **Each data source contains a subset of the information about a certain entity (a course, a person, a project, a paper …)**

Data heterogeneity

ID	Type	Title	Author	Subject	Year
09	Scholarly article	Theory of Contexts	F. Giunchiglia	AI	2003

ID	Kind	Title	Author	Topic
43	Book	Intelligent robots	A. Smith	Artificial intelligence
44	Paper	Theory of Contexts	Giunchiglia Fausto	Automated reasoning

- **Each data source describes data in different ways and with different terminology**

Data Quality

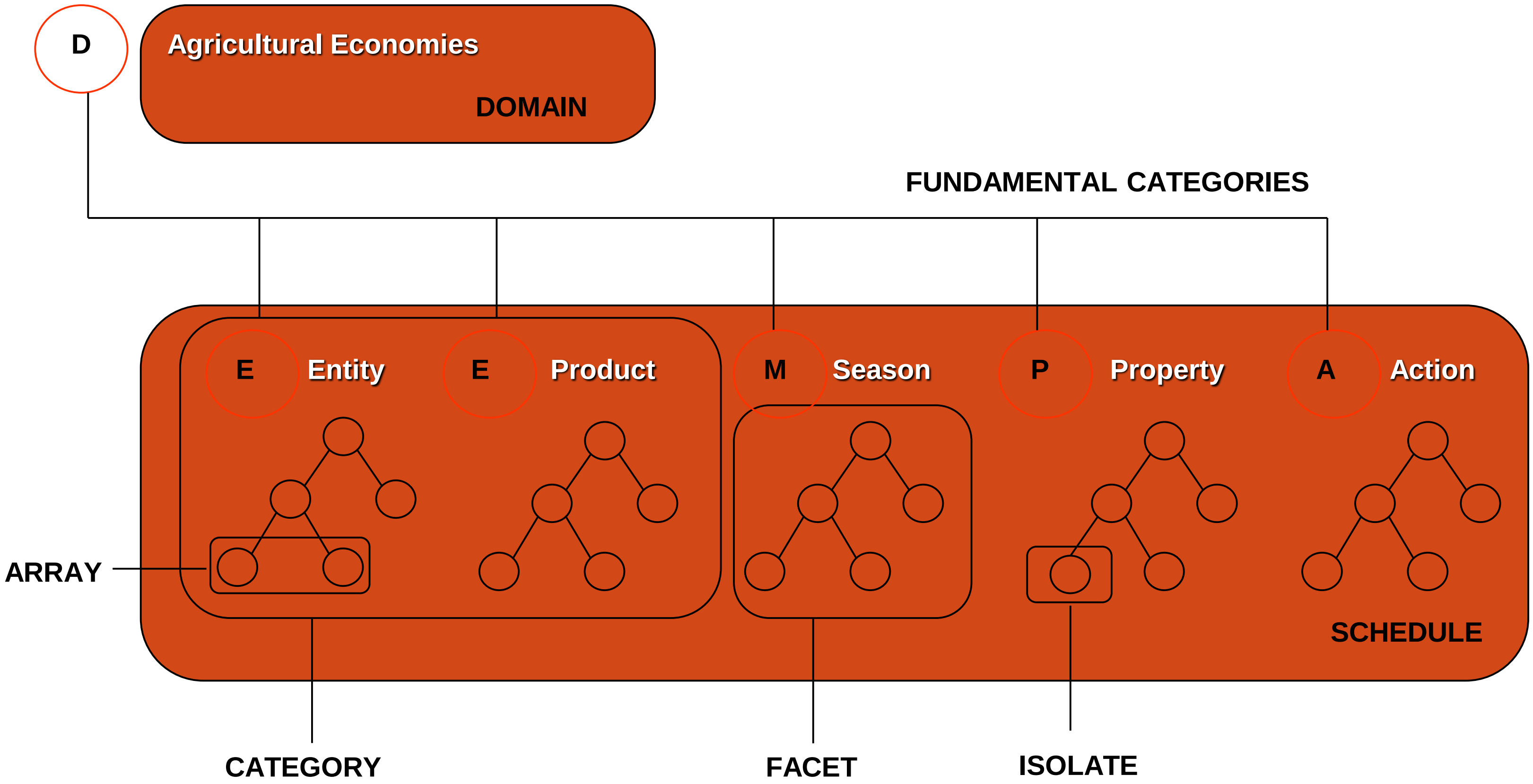
stringa_autori	autori_interni	keyword
F. Giunchiglia, J. Doe; P. Lewis	Fausto Giunchiglia (ID = 123)	Computer science, Ontology matching; semantics - tools

citta_editore	luogo_convegno	titolo	titolo_libro
	Crete (GR)	Semantic Matching	

- **Data sources are not normalized (several entities in one record)**
- **There is partial identity control (e.g. what is Crete? What is GR?)**
- **Data is poorly formatted and ambiguous (rules are not followed)**
- **Data is partial (missing values)**

The Faceted approach [Ranganathan, 1967]

- **Analysis**: relevant terms of the domain are identified and disambiguated
- **Synthesis**: within *fundamental categories* identified terms are categorized into *facets* according to their distinguishing *characteristics*

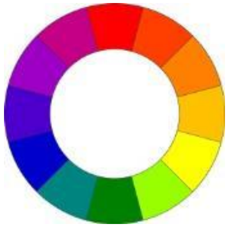


Not all ontologies are the same [Giunchiglia et al., 2014]

- **WHAT:** DERA aims to develop high quality descriptive ontologies
- **HOW:** We employ a faceted methodology with precise principles and guidelines
- **WHY:** DERA facets are integral part of the content of the Universal Knowledge Core (UKC) that in our view is a fundamental tool to enable semantic interoperability across cultures world-wide.
- **IN WHICH WAY:**
  - In order to scale and capture the diversity of the world, in the UKC we partition terminology into domains.
  - Each domain is defined in the UKC as a set of relevant entity types (etypes) each of them providing constraints, in form of templates, on the attributes and relations that entities of specific kinds (e.g. locations, organizations, persons, events) can instantiate and the language that can be used to express them.
  - Entity types impose a certain level of standardization, still giving the users the flexibility to define their own metadata and use different natural language terms (natural language level) to denote the same concept (formal language level).

Primitive notions

- **Concept:** a formal notion denoting an element of a DERA domain. Concepts constitute the formal language that can be used to describe entities.
- **Entity**: a (digital) description of any real world physical or abstract object so important to be denoted with a proper name. A single person, a place or an organization are all examples of entities.
- **Relation**: any object property used to connect two entities. Typical examples of relations include part-of, friend-of and affiliated-to.
- **Attribute**: any data property of an entity. Each attribute has a name and one or more values taken from a range of possible values.



DERA facets

- The language required to describe entities of a certain entity type in a given domain (D) correspond to entity classes (E), relations (R) and attribute (A) names as well as corresponding values.

- According to the DERA methodology, concepts and semantic relations between them form hierarchies of homogeneous nature at formal language level called facets, each of them codifying a different aspect of the domain.

ENTITY

- Location
  - Landform
  - (is-a) Natural elevation
    - (is-a) Continental elevation
      - (is-a) Mountain
      - (is-a) Hill
    - (is-a) Oceanic elevation
      - (is-a) Seamount
      - (is-a) Submarine hill
  - (is-a) Natural depression
    - (is-a)Continental depression
      - (is-a) Valley
      - (is-a) Trough
    - (is-a) Oceanic depression
      - (is-a) Oceanic valley
      - (is-a) Oceanic trough
  - Body of water
  - (is-a) Flowing body of water
    - (is-a) Stream, Watercourse
      - (is-a) River
      - (is-a) Brook
  - (is-a) Still body of water
    - (is-a) Lake
    - (is-a) Pond

RELATION

- Direction
  - (is-a) East
  - (is-a) North
  - (is-a) South
  - (is-a) West

- Relative level
  - (is-a) Above
  - (is-a) Below

- Containment
  - (is-a) part-of

ATTRIBUTE

- Name
- Latitude
- Longitude
- Altitude
- Area
- Population

- Depth
  - (value-of) deep
  - (value-of) shallow

- Length
  - (value-of) long
  - (value-of) short

Advantages of DERA

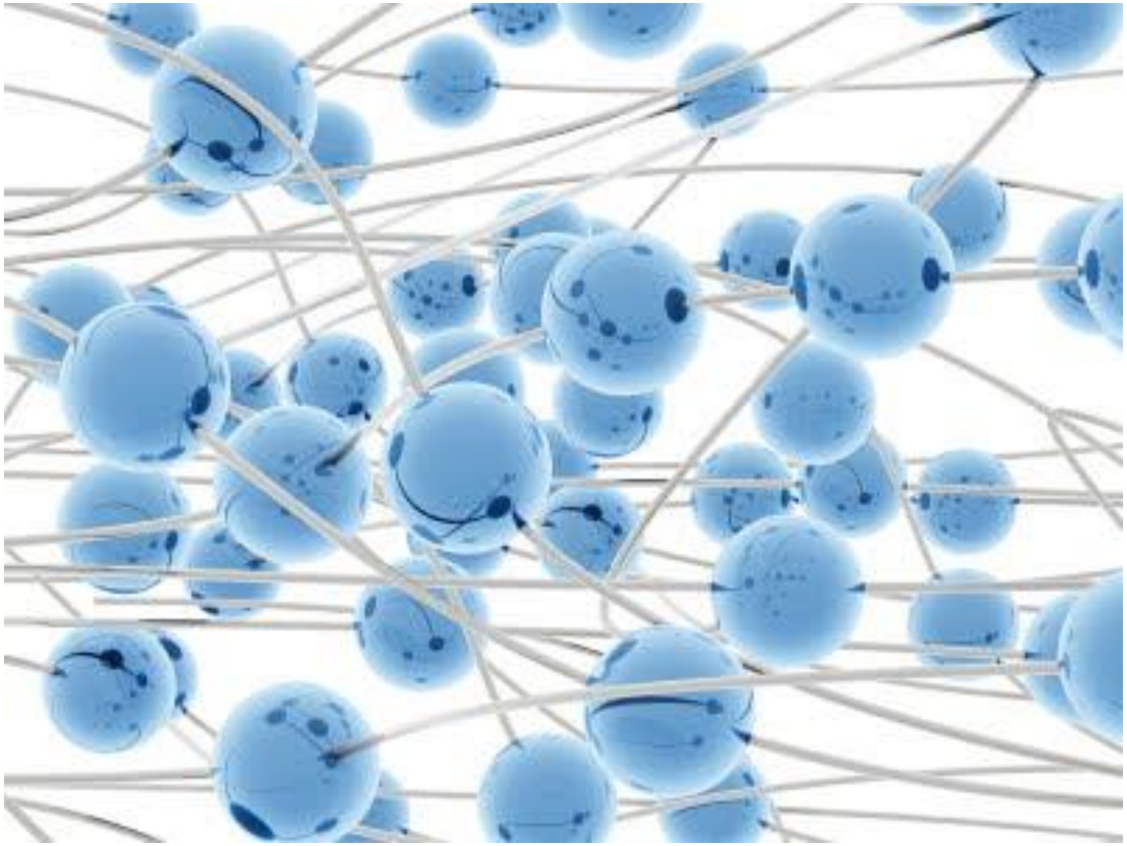
- DERA facets have **explicit semantics** and are modeled as descriptive ontologies

- DERA facets inherits all the important properties of the faceted approach, such as robustness and scalability

- DERA allows for **automated reasoning** via the formalization into Description Logics ontologies. In particular, DERA allows for a very expressive search by any entity property

DERA allows to improve onWordnet difficulties

DERA STEPS



Making sense of terms and concepts…

Step 2:analysis (I)

- ***Identification of the relevant authoritative resources:*** by initially inspecting the terms identified during the previous step and by consulting dictionaries, available standards, and sources on line (e.g. Wikipedia), the purpose is to identify the key resources necessary to deeply understand the identified terms.

- ***Study of the domain:*** to effectively start the analysis, it is fundamental to study the domain under examination. This allows the identification of the core terms, i.e. the terms which play a central role in the domain.

- ***Deep analysis of each term:*** each term must be analyzed separately such that we can clearly understand the similarities and differences with respect to each other. Collect existing definitions for the terms at the purpose of identifying their essential (i.e. true in all contexts) *genus* and *differentia.*

Step 2:analysis (II)

- ***Rationalize terms and concepts:*** the main result of the deep analysis should be the identification of:
  - terms with same meaning, i.e. synonyms, that should be grouped together as lexicalizations of the same atomic concept
  - redundant concepts and individuals that should be eliminated (e.g. WordNet contains “Winchester College”)

- ***Categorize concepts:*** identified atomic concepts are distinguished into:
  - entity classes,
  - relations,
  - attribute names and attribute values

For instance, while analyzing educational institutions we identified the attribute:
educational mode: a way or manner in which lessons are given

> (value-of) **regular**: of or relating to the mode of teaching based on fixed schedule
>
> (value-of) **corresponding**: of or relating to the mode of teaching through broadcasting or remotely

Step 3: synthesis

- With this step, we give shape to each facet by grouping similar concepts together.
- In practice, this may require subsequent iterations of analysis and synthesis to progressively refine the facets and to ensure that the various principles are met.
- ***Identification of the main characteristics of division:*** in arranging identified concepts, the high-level characteristics of divisions which are peculiar of the domain under examination need to be identified. They should be reflected in the differentia of the various analyzed terms.

Educational institutions can be distinguished at first level *by level of complexity* from preschool to university; at second level we distinguish secondary schools *by programme orientation*.

DERA PRINCIPLES

We need principles to guide the development…

Selection of the characteristics

- ***Principle of ascertainability:*** every single design choice must be verifiable by consulting authoritative sources such as dictionaries and any other source that is relevant for the domain under examination. All the relevant material used to take a design choice must be reported in the final documentation.
- ***Principle of permanence:*** each characteristic should reflect permanent properties of entities. The selected characteristics of division should correspond to essential properties (as opposed to accidental), i.e. properties which are inherent in the nature of the entities and do not vary in time.
- ***Principle of relevance:*** the selection of the characteristics that are used to form the facets should reflect the purpose, scope and subject of the facet.

Formation of arrays

- ***Principle of exhaustiveness:*** concepts in the same array should be totally exhaustive w.r.t. their respective common parent concept in the facet hierarchy. For example, to classify concepts immediately under the attribute *release frequency* (of publications) we should provide all the following values: *daily, weekly, monthly, quarterly, seasonal, annually, and bi-annually*.
- ***Principle of exclusiveness:*** the characteristics should be selected in such a way that sibling concepts in each array must be mutually exclusive. Moreover, all the characteristics used to classify a concept must be mutually exclusive, i.e. no two facets can overlap in content.
- ***Principle of helpful sequence:*** the order of the concepts within each array should reflect the purpose, scope and subject of the facet. It should be applied consistently and should not be changed. For example, the facet *populated place* may include *hamlet, village, town* and *city* in ascending order according to population.

Selection of the terminology

- ***Principle of currency:*** the terms chosen to denote concepts in a certain language should be those of current usage in the subject field. This is particularly important for the preferred term.
- ***Principle of reticence:*** the terms chosen to denote concepts and their glosses should not reflect any bias or prejudice (e.g. of gender, cultural, religious) or express any personal opinion of the person who develops the facet. By doing so, the aim is to minimize the cultural gap and decrease the probability of disagreement between different users.
- ***Principle of context:*** the position of a concept in the facet is a function of its meaning. Therefore the term used to denote a certain concept should be selected by taking into account the position of the concept itself in the facet (i.e. the path from the root). Moreover, the gloss used to describe the concept in a certain language should reflect the position of the concept w.r.t. its immediate parent (the genus) and its siblings.
- ***Principle of consistency:*** the terms should be used consistently (i.e. with same meaning) if they appear in the same context (e.g. in a certain sub-tree or related to a certain kind of entity).

Some reference material

[Ranganathan, 1967] Prolegomena to library classification. Asia Publishing House.

[Gruber, 1993] A translation approach to portable ontology specifications. Knowledge Aquisition, 5 (2), 199–220.

[Pollock, 2002] Integration's Dirty Little Secret: It's a Matter of Semantics. Whitepaper, The Interoperability Company.

[Guarino and Welty, 2002] Guarino, N., Welty, C. (2002). Evaluating ontological decisions with OntoClean. Communications of the ACM, 45(2), 61-65.

[Uschold and Gruninger, 2004] Ontologies and semantics for seamless connectivity. SIGMOD Rec., 33(4), 58–64.

[Varzi, 2006] A note on the transitivity of parthood. Applied Ontology, 1 (2), 141-146.

[Giunchiglia et al., 2009] Faceted Lightweight Ontologies. In: Conceptual Modeling: Foundations and Applications, LNCS Springer.

[Giunchiglia et al., 2012a] A facet-based methodology for the construction of a large-scale geospatial ontology. Journal on Data Semantics, 1 (1), pp. 57-73.

[Giunchiglia et al., 2012b] Domains and context: first steps towards managing diversity in knowledge. Journal of Web Semantics, special issue on Reasoning with Context in the Semantic Web.

[Maltese, 2012] Dealing with semantic heterogeneity in classifications. PhD thesis: http://eprints-phd.biblio.unitn.it/700/ (see chapters 1.1. and 2.1)

[Giunchiglia et al., 2014] From Knowledge Organization to Knowledge Representation. Knowledge Organization. 41(1), 44-56.

Thank you!
Questions?