

Ontology-based User Modeling and Recommendation

Iván Cantador

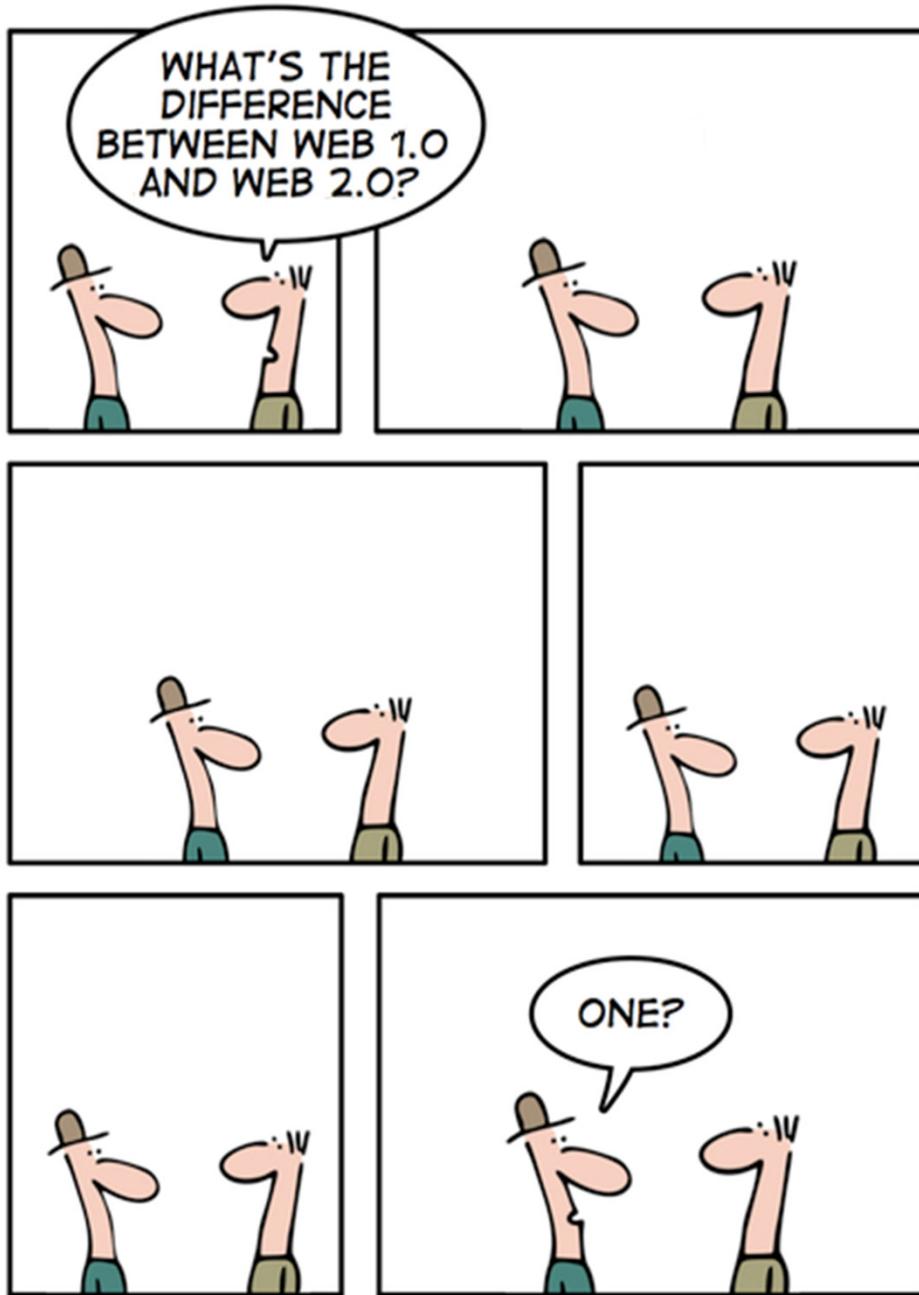
ivan.cantador@uam.es

Information Retrieval Group

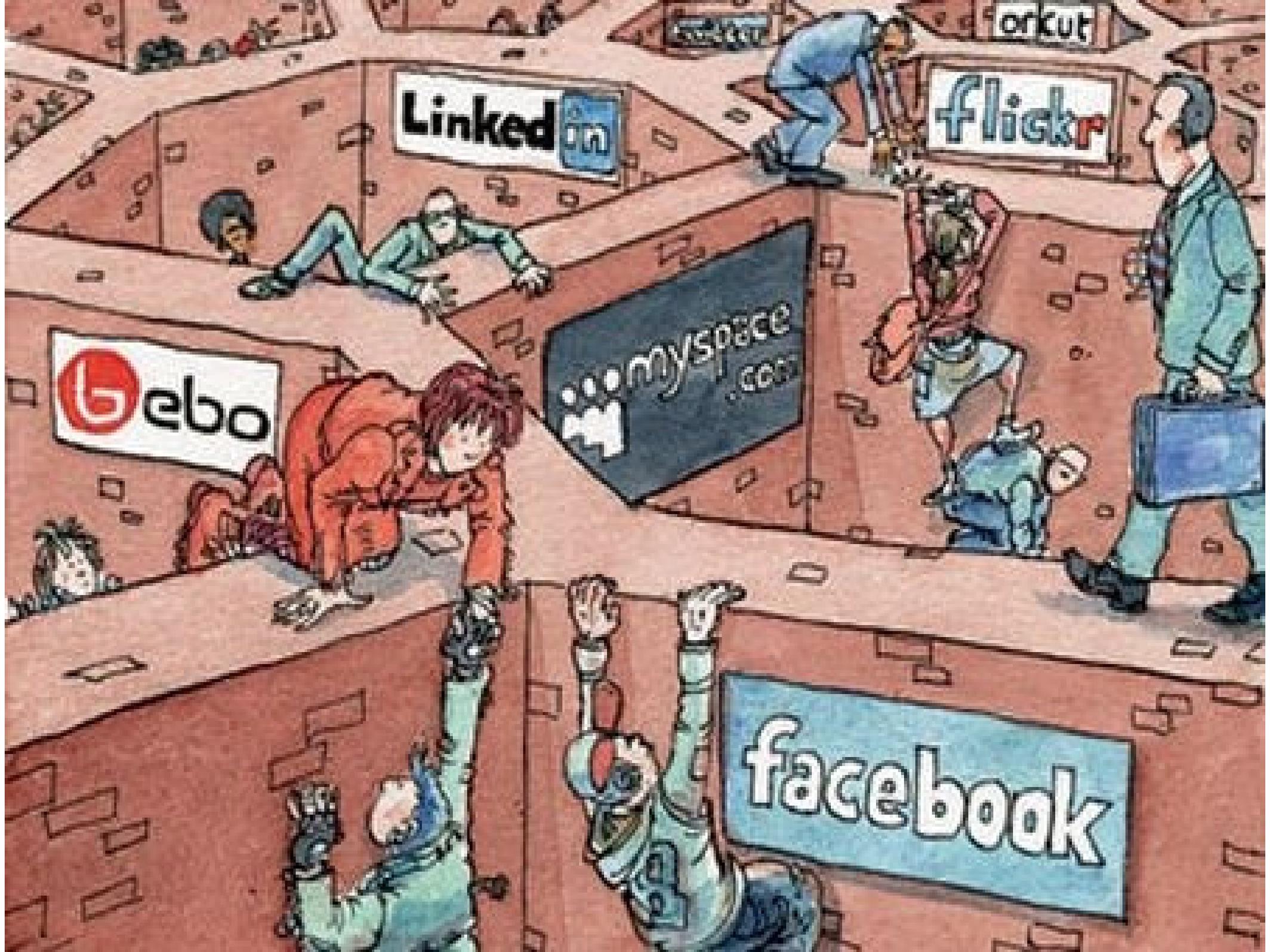
Universidad Autónoma de Madrid, Spain

- The Social Semantic Web
 - Web 2.0: The Social Web
 - Web 3.0: The Semantic Web
- Ontology-based User Modeling
 - Ontology-based User Preferences
 - Mapping Social Tags to Ontology Concepts
 - Mapping Facebook Likes to DBpedia Entities
- Ontology-based Recommendation
 - An Example of Ontology-based Recommender System
 - A Linked Data-based Cross-domain Recommendation Approach
- Conclusions

- The Social Semantic Web
 - **Web 2.0: The Social Web**
 - Web 3.0: The Semantic Web
- Ontology-based User Modeling
 - Ontology-based User Preferences
 - Mapping Social Tags to Ontology Concepts
 - Mapping Facebook Likes to DBpedia Entities
- Ontology-based Recommendation
 - An Example of Ontology-based Recommender System
 - A Linked Data-based Cross-domain Recommendation Approach
- Conclusions



Is it so easy?



Web 2.0 – Web 1.0 =

user generated contents,

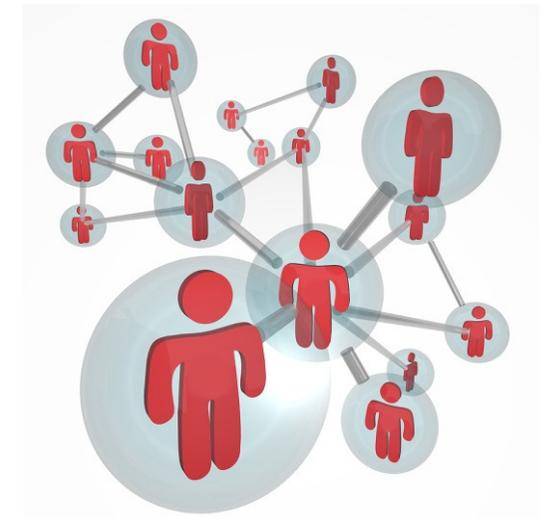
user generated contents,

...

and

user generated contents

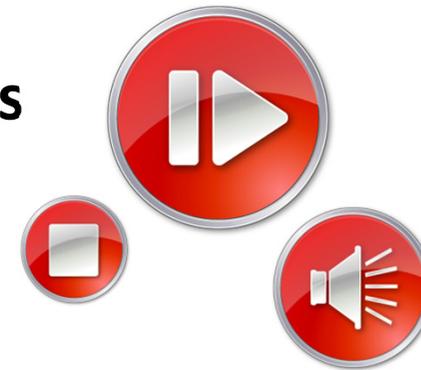
- **Web 2.0 = The Social Web**
 - Creation of diverse formats of **user generated contents**
 - People:
 - **Communicate online with contacts**
- Social networks - Facebook, Twitter, LinkedIn, ...*



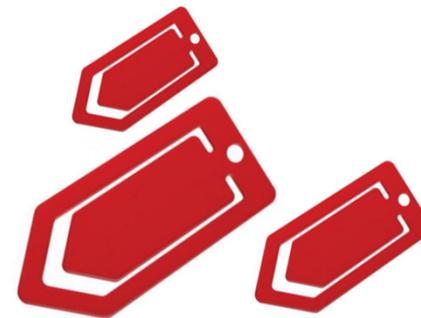
- **Web 2.0 = The Social Web**
- Creation of diverse formats of **user generated contents**
 - People:
 - Communicate online with contacts
 - **Contribute to wiki-style repositories**
Wikipedia, AboutUs, wikiHow, WikiMapia, ...



- **Web 2.0 = The Social Web**
- Creation of diverse formats of **user generated contents**
 - People:
 - Communicate online with contacts
 - Contribute to wiki-style repositories
 - **Upload or create multimedia objects**
YouTube, Flickr, Last.fm, ...



- **Web 2.0 = The Social Web**
- Creation of diverse formats of **user generated contents**
 - People:
 - Communicate online with contacts
 - Contribute to wiki-style repositories
 - Upload or create multimedia objects
 - **Maintain personal bookmarks**
Delicious, Digg, CiteULike, ...



- **Web 2.0 = The Social Web**
- Creation of diverse formats of **user generated contents**
 - People:
 - Communicate online with contacts
 - Contribute to wiki-style repositories
 - Upload or create multimedia objects
 - Maintain personal bookmarks
 - **Post comments, reviews and ratings**

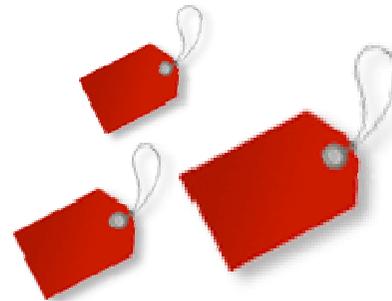
Amazon, Epinions, Blogger, Rotten Tomatoes, ...

Blogs and recommender systems



- **Web 2.0 = The Social Web**
- Creation of diverse formats of **user generated contents**
 - People:
 - Communicate online with contacts
 - Contribute to wiki-style repositories
 - Upload or create multimedia objects
 - Maintain personal bookmarks
 - Post comments, reviews and ratings
 - **Annotate resources**

Social tagging systems





Bard, M. 2010. The Big Book of Social Media: Case Studies, Stories, Perspectives.
<http://www.mirnabard.com>

A night sky filled with stars, with a rocky landscape in the foreground. The sky is dark blue and black, with many small white stars. The foreground shows large, dark rocks and some green bushes. The text is overlaid on the sky.

Ok, the Web 2.0 is out there!

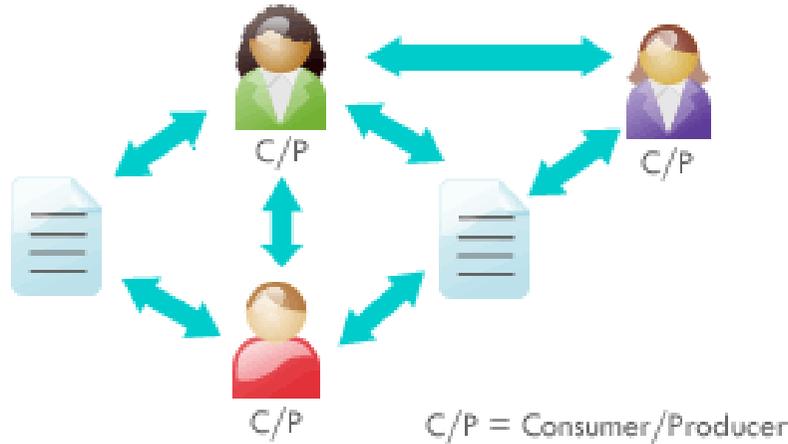
What's next!?

- The Social Semantic Web
 - Web 2.0: The Social Web
 - **Web 3.0: The Semantic Web**
- Ontology-based User Modeling
 - Ontology-based User Preferences
 - Mapping Social Tags to Ontology Concepts
 - Mapping Facebook Likes to DBpedia Entities
- Ontology-based Recommendation
 - An Example of Ontology-based Recommender System
 - A Linked Data-based Cross-domain Recommendation Approach
- Conclusions

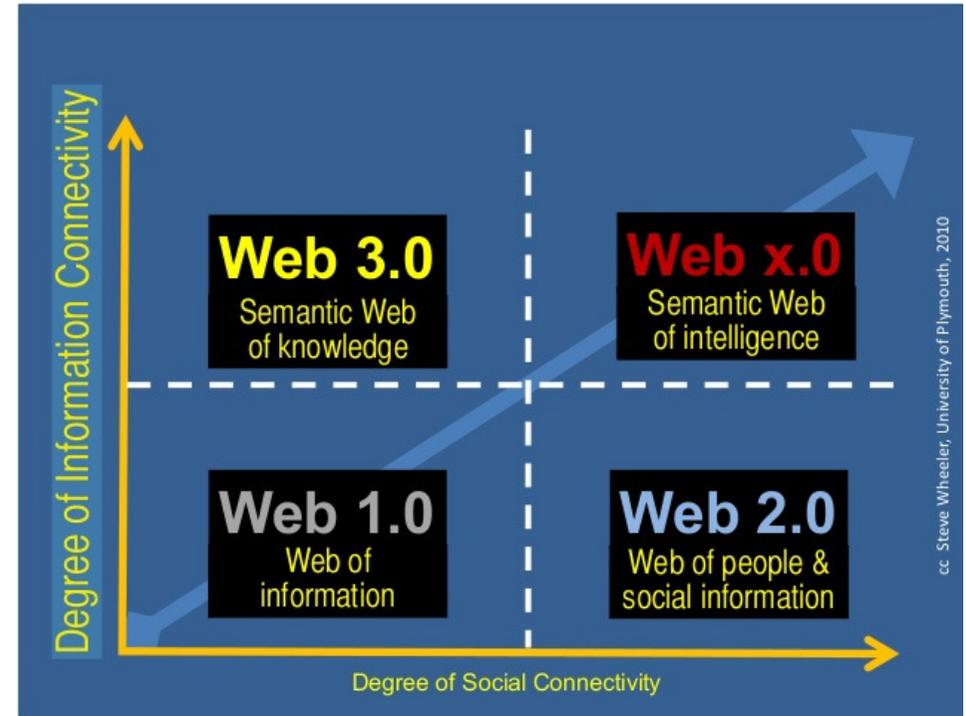
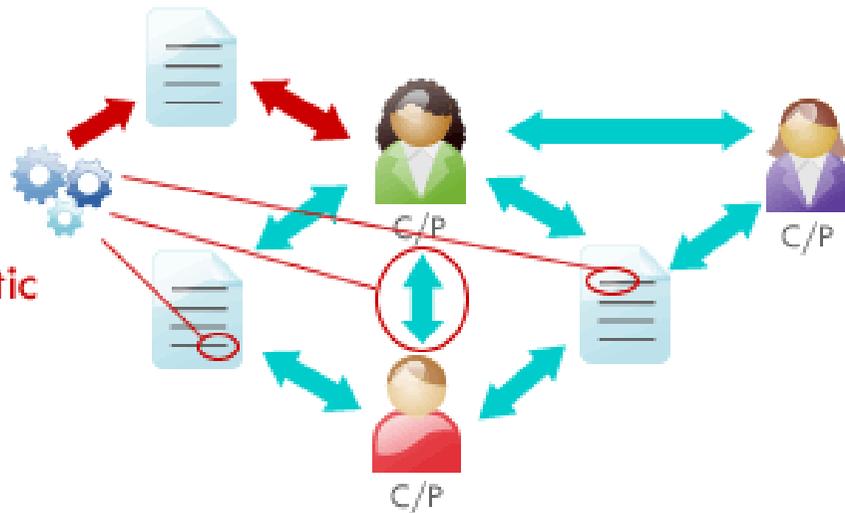
Web 1.0

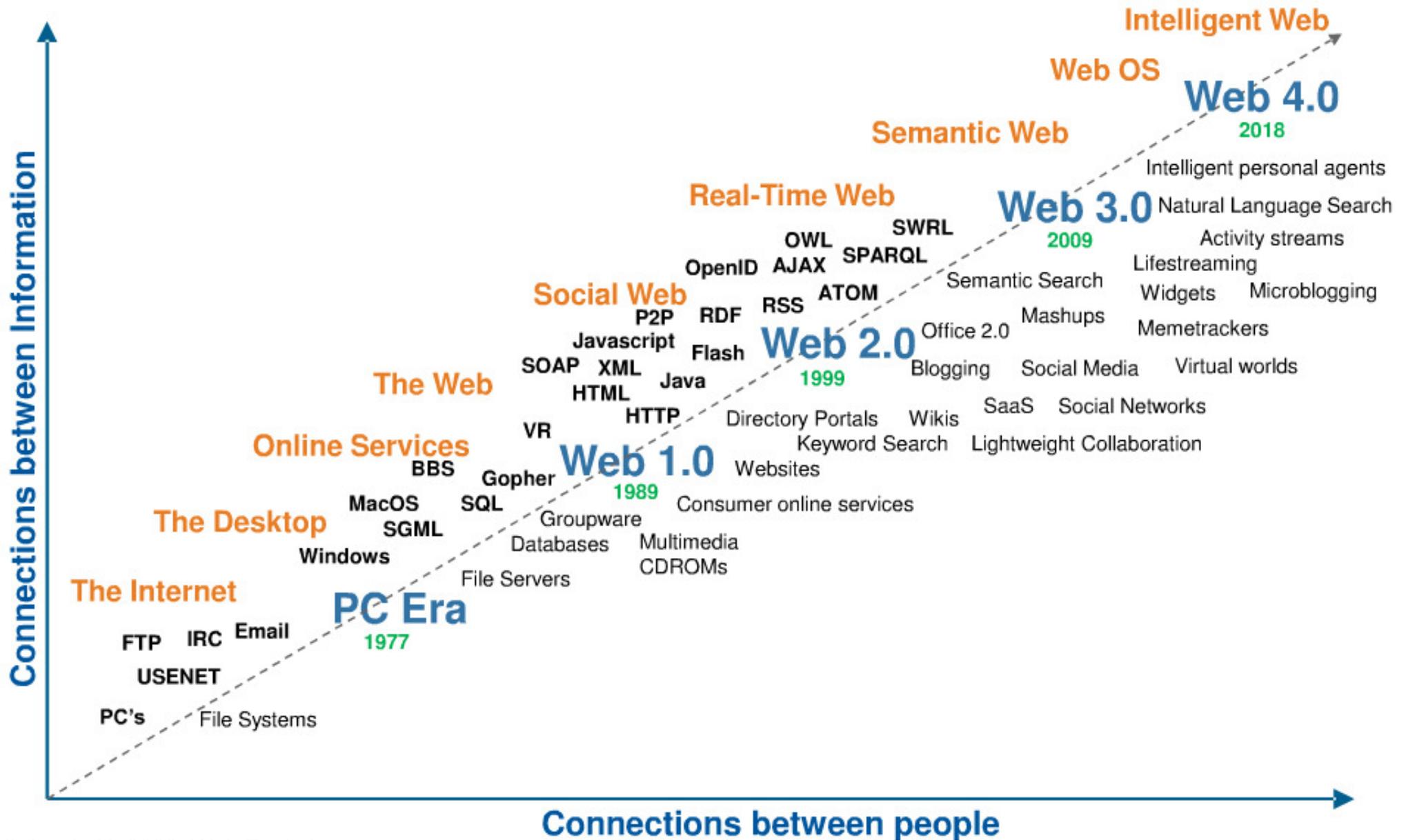


Web 2.0



The Semantic Web

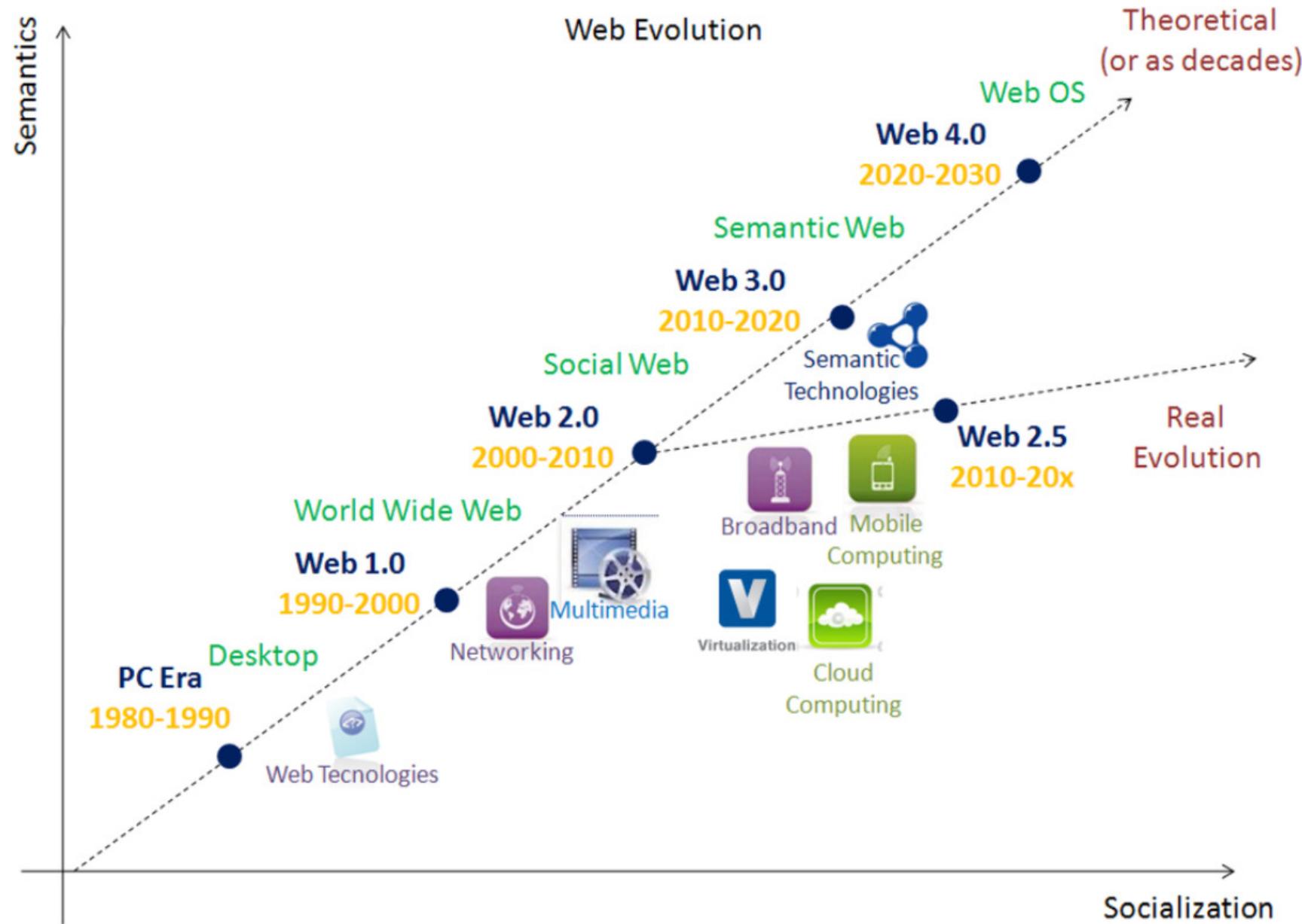




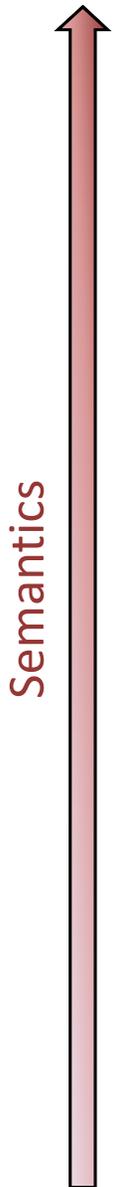
Spivack, N. 2009. Web Evolution.

<http://www.slideshare.net/novaspivack/web-evolution-nova-spivack-twine>

Web 3.0... Let's say Web 2.5 ;-)

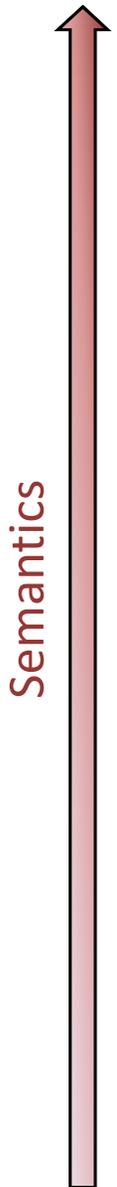


Pileggi, S.; Amor, R. 2013. Addressing Semantic Geographic Information Systems. Future Internet 5, pp. 585



- **Bags of words:** uncategorized terms
 - e.g. a folksonomy

hominidae homo human
humans man
sapiens primates



- **Taxonomies:** categories + hierarchical relations

- e.g. Linnaen taxonomy

human := [*kingdom*: Animalia, *class*: Mammalia, *order*: Primates, *family*: Hominidae, *genus*: Homo, *species*: Homo sapiens]

- **Bags of words:** uncategorized terms

- e.g. a folksonomy

hominidae homo human
humans man
sapiens primates

- **Thesauri:** categories + fixed hierarchical & associative relations

- e.g. WordNet thesaurus

human: S: (n) homo, man, human being, human (any living or extinct member of the family Hominidae characterized by superior intelligence, articulate speech, and erect carriage)

direct hypernym: S: (n) **hominid** (a primate of the family Hominidae...)

direct hyponym: S: (n) **Homo sapiens** (the only surviving hominid ...)

- **Taxonomies:** categories + hierarchical relations

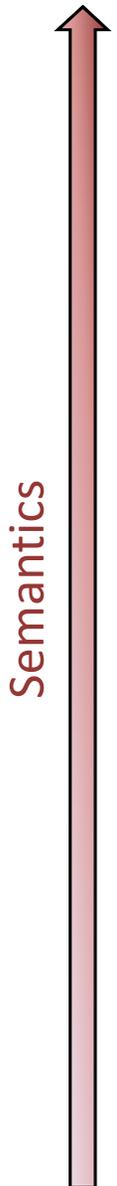
- e.g. Linnaen taxonomy

human := [*kingdom:* Animalia, *class:* Mammalia, *order:* Primates, *family:* Hominidae, *genus:* Homo, *species:* Homo sapiens]

- **Bags of words:** uncategorized terms

- e.g. a folksonomy

hominidae homo human
humans man
sapiens primates



- **Ontologies:** classes + instances + arbitrary semantic relations + rules
 - e.g. DBpedia ontology

- **Thesauri:** categories + fixed hierarchical & associative relations
 - e.g. WordNet thesaurus

human: S: (n) homo, man, human being, human (any living or extinct member of the family Hominidae characterized by superior intelligence, articulate speech, and erect carriage)

direct hypernym: S: (n) **hominid** (a primate of the family Hominidae...)

direct hyponym: S: (n) **Homo sapiens** (the only surviving hominid ...)

- **Taxonomies:** categories + hierarchical relations
 - e.g. Linnaen taxonomy

human := [*kingdom:* Animalia, *class:* Mammalia, *order:* Primates, *family:* Hominidae, *genus:* Homo, *species:* Homo sapiens]

- **Bags of words:** uncategorized terms
 - e.g. a folksonomy

hominidae homo human
humans man
sapiens primates

Semantics

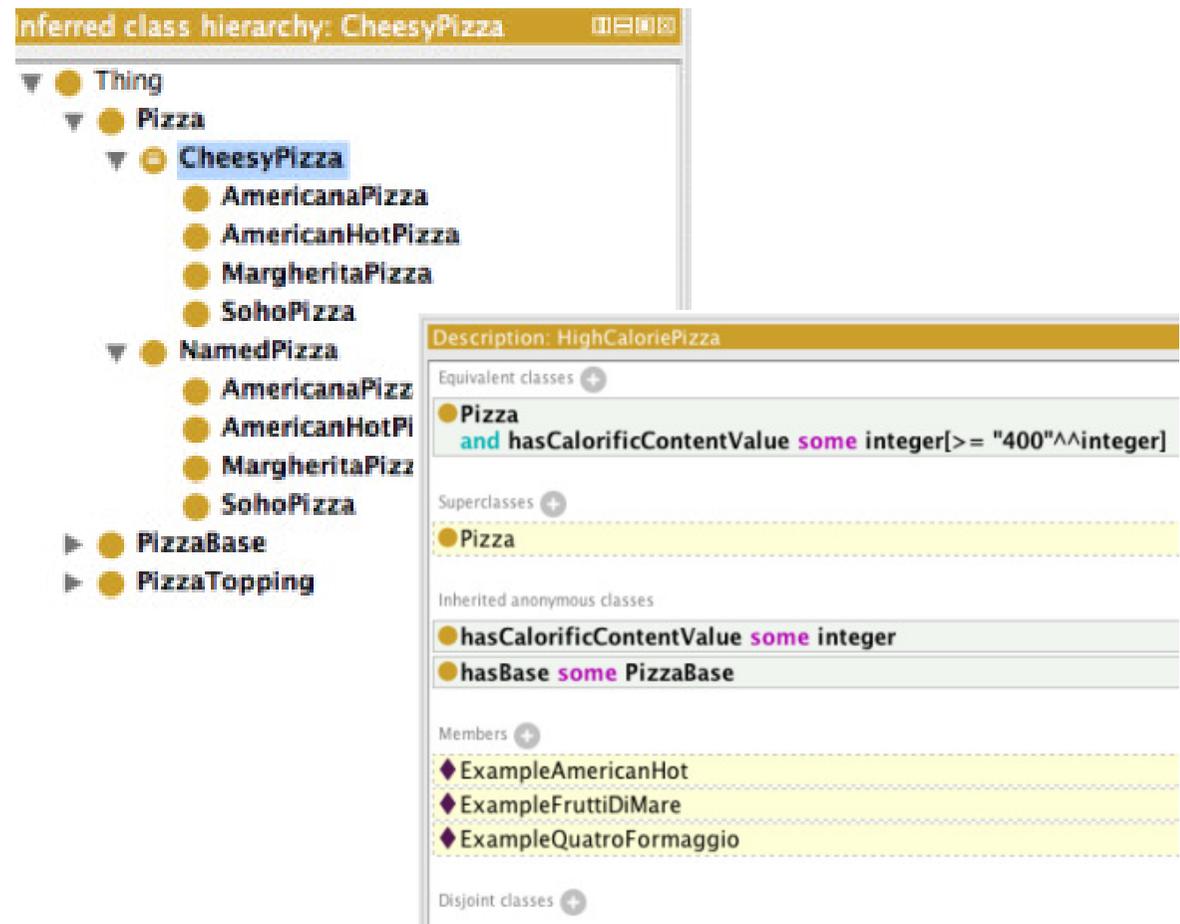
An ontology is a **formal**, **explicit** specification of a **shared** conceptualization

Formal: machine-readable

Explicit: concepts, properties, relations, functions, constraints, axioms are explicitly defined

Shared: consensual knowledge

Conceptualization: abstract model and simplified view of some phenomenon in the world that we want to represent



Gruber, T. 1995. *Toward Principles for the Design of Ontologies Used for Knowledge Sharing*. *International Journal of Human-Computer Studies* 43(5-6), pp. 907-928.

How to model the world of pizzas?

ontology

Classes

- DomainConcept
 - Country
 - Food
 - IceCream
 - Pizza
 - CheesyPizza**
 - InterestingPizza
 - MeatyPizza

Equivalent To +

- Pizza
 - and (hasTopping some CheeseTopping)

SubClass Of +

SubClass Of (Anonymous Ancestor)

- hasBase some PizzaBase

Properties

- topObjectProperty
 - hasCountryOfOrigin
- hasIngredient
 - hasBase
 - hasTopping
 - hasSpiciness
- isIngredientOf
 - isBaseOf
 - isToppingOf

Restrictions

knowledge base

Instances

Cuatro
Formaggi

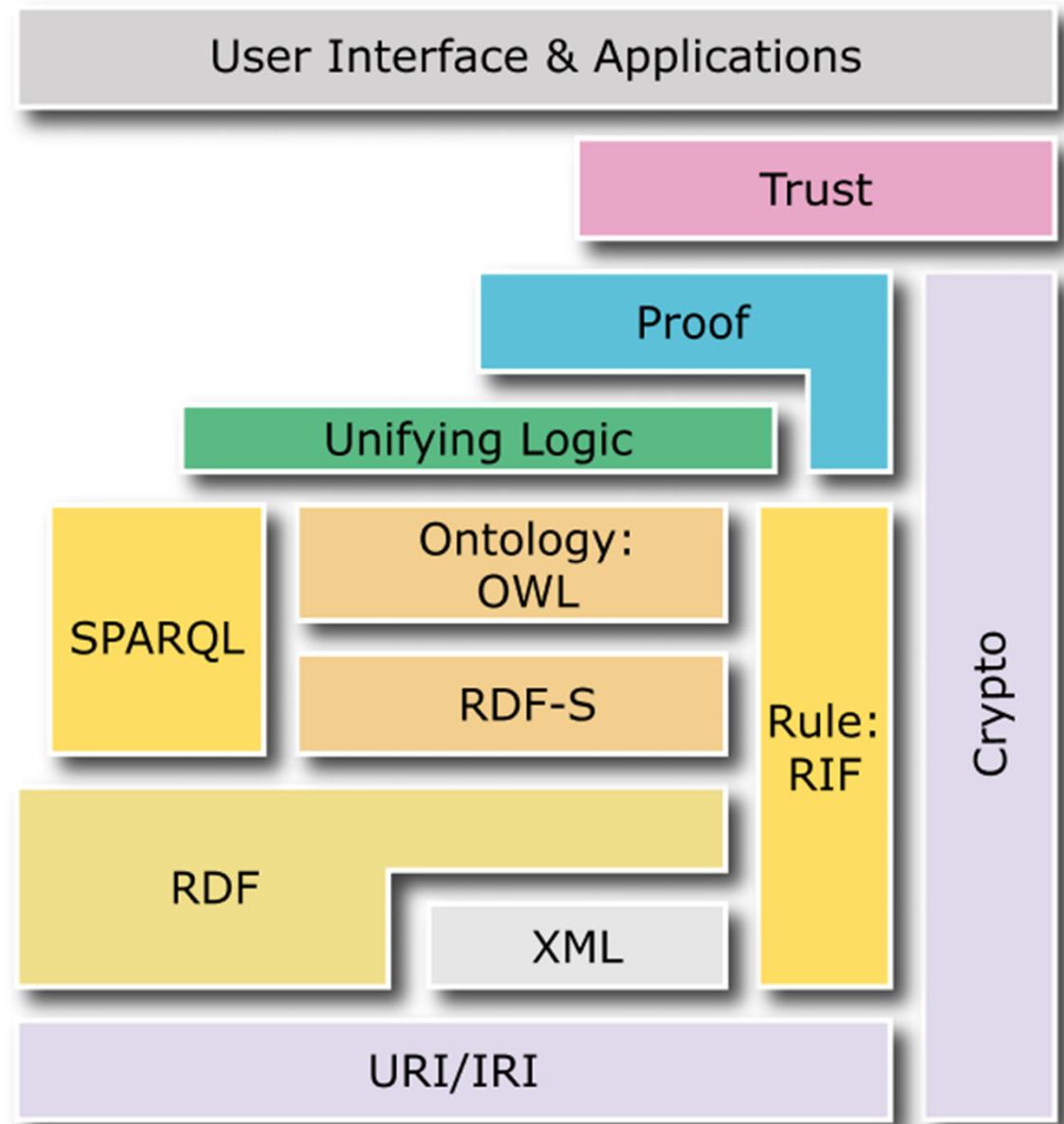


hasCountryOrigin

Italy



http://owl.cs.manchester.ac.uk/tutorials/protegeowltutorial/resources/ProtegeOWLTutorialP4_v1_3.pdf



Querying data (**SPARQL**)

Defining ontologies (**RDFS, OWL**)

Expressing and linking metadata (**RDF**)

Assigning unambiguous names (**URI**)

RDF (Resource Description Framework)

- RDF identifies things using Web identifiers (URIs), and describes **resources** with **properties** and property **values**.
 - The triple representation (**subject**, **predicate**, **object**)



```
<?xml version="1.0"?>
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:epsp="http://www.eps.uam.es/personnel#"
  <rdf:Description
    rdf:about="http://www.eps.uam.es/personnel/IvanCantador">
      <epsp:name>Ivan</epsp:name>
      <epsp:lastname>Cantador</epsp:lastname>
      <epsp:nationality>Spanish</epsp:nationality>
      <rdf:type>Associate Professor</rdf:type>
    </rdf:Description>
  </rdf:RDF>
```

- RDFS provides the framework to describe **classes** and **properties**, allowing the creation of **hierarchies**

```
<?xml version="1.0"?>
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xml:base="http://www.eps.uam.es#"
  <rdfs:Class rdf:ID="Professor" />
  <rdfs:Class rdf:ID="Associate Professor">
    <rdfs:subClassOf rdf:resource="#Professor"/>
  </rdfs:Class>
  <rdf:Property rdf:ID="teachesSubject">
    <rdfs:Domain rdf:resource="#Professor"/>
    <rdfs:Range rdf:resource="#Subject"/>
  </rdf:Property>
</rdf:RDF>
```

- OWL is a stronger language with greater **machine interpretability** than RDF/RDFS (reasoning support)
 - OWL Little / OWL DL / OWL Full

```
<?xml version="1.0"?>
```

```
<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:owl = "http://www.w3.org/2002/07/owl#"
  xml:base="http://www.eps.uam.es#
```

```
<owl:Class rdf:ID="Professor">
```

```
  <owl:restriction>
```

```
    <owl:onProperty rdf:resource="#hasAcademicTitle"/>
```

```
    <owl:hasValue>
```

```
      PhD^^http://www.w3.org/2001/XMLSchema#string
```

```
    </owl:hasValue>
```

```
  </owl:restriction>
```

```
</owl:Class>
```

```
</rdf:RDF>
```

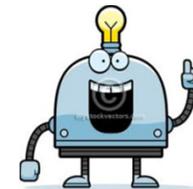
Value constraints:
owl:allValuesFrom
owl:someValuesFrom
owl:hasValue

Cardinality constraints:
owl:cardinality
owl:minCardinality
owl:maxCardinality

....



Iván has a PhD -> therefore
Iván can be a professor!



SPARQL (SPARQL Protocol and RDF Query Language)²⁸

- SPARQL is a query language for RDF
 - Based on the triple representation (**subject**, **predicate**, **object**)
 - SPARQL 1.1 is W3C Recommendation since 21st March 2013

```
PREFIX epsp: <http://www.eps.uam.es>
```

```
<rdf:RDF
```

```
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
```

```
  SELECT ?x ?name
```

```
  WHERE
```

```
  { ?x rdf:type epsp:Associate Professor.
```

```
    ?x epsp:name ?name.
```

```
    ?x epsp:nationality "Spanish" }
```

Can you "google" this?

SPARQL lets the construction of very powerful queries

- Filtering elements
- Querying named graphs
- Ordering by/ distinct / reduced / offset / limit
- Count / sum / avg / min / max / GroupConcat

- **Explicit semantics in the Web**
 - Microformats
 - RDFa
 - HTML5 microdata
 - Schema.org
 - Linked Open Data

Microformats

- Microformats use existing **HTML attributes** to embed structured data types in an HTML document
- Expressive power is limited as microformats are only designed to pre-defined **vocabularies**
- No interlinking between entities

```
<div class="vcard">
  <a class="fn org url" href="http://www.eps.uam.es/">EPS</a>
  <div class="adr">
    <span class="street-address">Francisco Tomas y Valiente</span >
    <span class="locality">Madrid</span>,
    <span class="postal-code">28049</span>
    <span class="country-name">Spain</span >
  </div>
</div>
```

Microformat specifications

- People and Organizations
[hCard](#), [XFN](#)
- Calendars and Events
[hCalendar](#)
- Opinions, Ratings and Reviews
[VoteLinks](#), [hReview](#)
- Licenses:
[rel-license](#)
- Tags, Keywords, Categories
[rel-tag](#)
- Lists and Outlines
[XOXO](#)
- More...
See [the list of all microformats](#)

<http://microformats.org>

- W3C standard for embedding RDF data in HTML documents
 - A set of **new HTML attributes** and specs of how to use them
- RDFa is just a syntax, the publisher has to choose the **vocabulary**
- RDFa is **domain independent**

- Syntax attributes: **@prefix, @vocab.**
- Subject attributes: **@about.**
- Predicate attributes: **@property, @rel, @rev.**
- Resource attributes: **@resource, @href, @src.**
- Literal attributes: **@datatype, @content, @xml:lang** Or **@lang.**
- Macro attributes: **@typeof, @inlist.**

```
<html>
<body vocab="http://xmlns.com/foaf/0.1/">
...
<h2 property="name">Ivan Cantador</h2>
<p><span property="gender">undefined</span></p>
...
</body>
</html>
```

Web standard since June 2012
<http://www.w3.org/TR/rdfa-core/>

- Used by Facebook! ... among others
 - **RDF vocabulary** used in conjunction with RDFa

```
<html xmlns:og="http://opengraphprotocol.org/schema/">
```

```
<head>
```

```
<title>The Rock (1996)</title>
```

```
<meta property="og:title" content="The Rock" />
```

```
<meta property="og:type" content="movie" />
```

```
<meta property="og:url" content="http://www.imdb.com/title/tt0117500/" />
```

```
<meta property="og:image" content="http://ia.media-imdb.com/images/rock.jpg" />
```

```
...
```

```
</head> ...
```

```
</html>
```



- HTML5 introduces **inline elements** (text-level semantics) to describe specific type of information

```
<div itemscope itemtype="http://data-vocabulary.org/Person">  
  My name is <span itemprop="name">Ivan</span>  
  Here is my home page:  
  <a href="http://arantxa.ii.uam.es/~cantador/" itemprop="url">  
http://arantxa.ii.uam.es/~cantador/  
  </a>  
  I live in Madrid, and work as an  
  <span itemprop="title">Associate Professor</span> at  
  <span itemprop="affiliation">UAM</span>.  
</div>
```

- Itemid
- Itemprop
- Itemref
- Itemscope
- itemtype

<http://dev.w3.org/html5/md-LC/#encoding-microdata>

- Schema.org provides a collection of **vocabularies** and a micro-data format, which are recognized by the major search engine providers (de facto standard)

```
<div itemscope itemtype=http://schema.org/Movie>
```

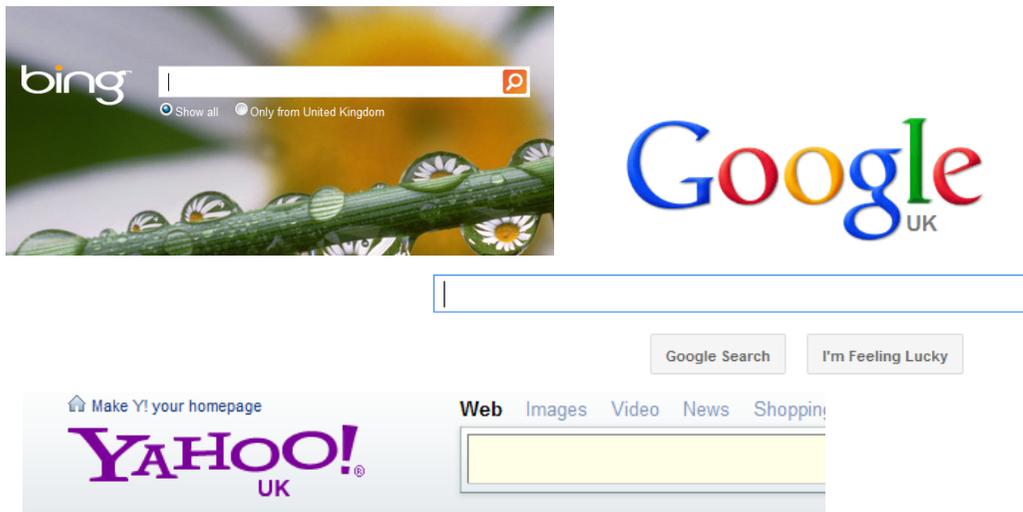
```
<h1 itemprop="name">Pirates of the Carribean: On Stranger Tides (2011)</h1>
```

Director:

```
<div itemprop="director" itemscope itemtype="http://schema.org/Person">
```

```
<span itemprop="name">Rob Marshall</span>
```

```
</div>
```



Organization of Schemas

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

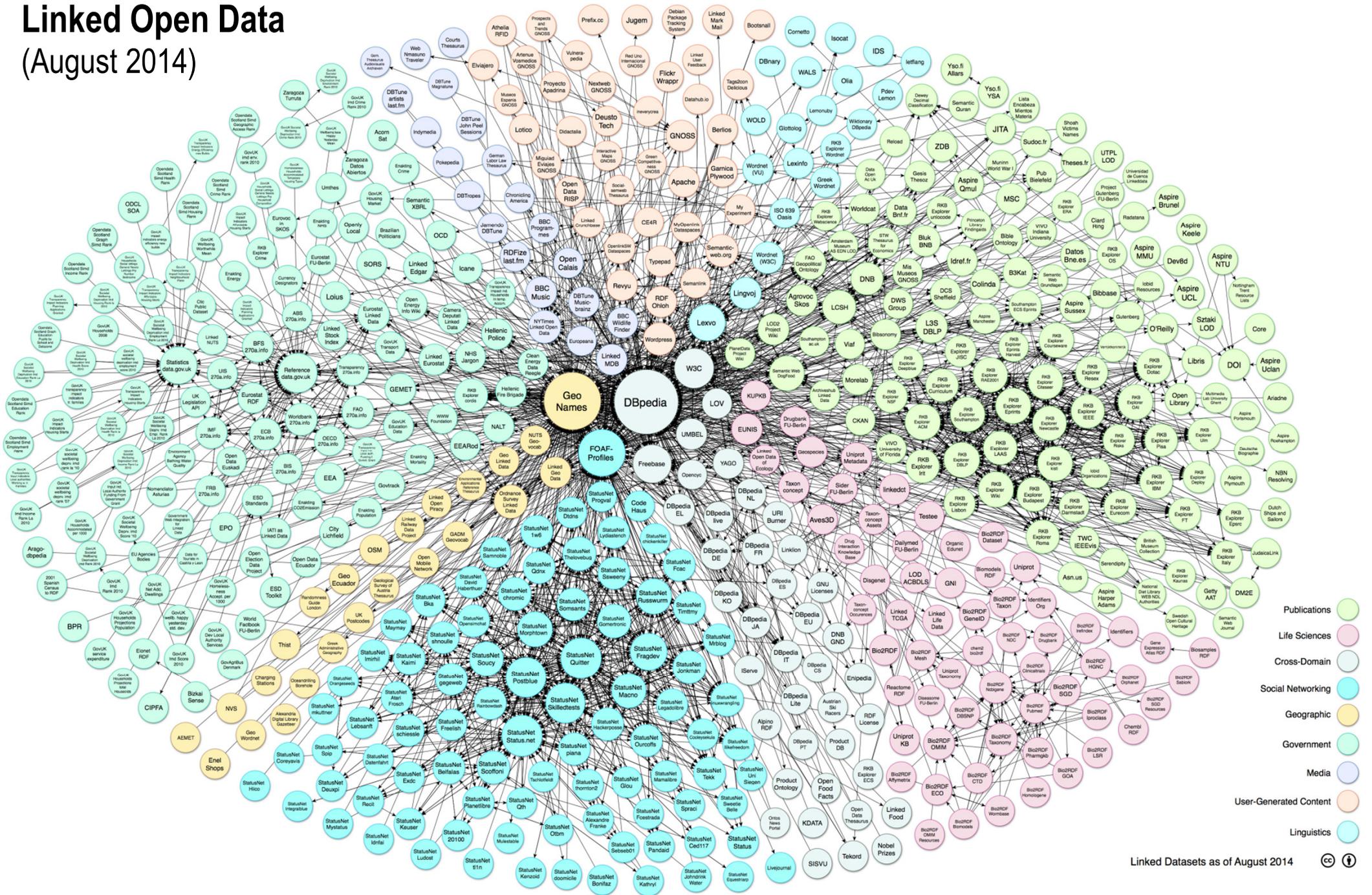
Browse the full hierarchy:

- [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](#)

Linked Open Data (August 2014)



DBpedia: The Wikipedia ontology and knowledge base

DBpedia, <http://dbpedia.org>

Wikipedia

The screenshot shows the Wikipedia article for 'Universidad Autónoma de Madrid'. The page includes a search bar, navigation tabs (Artículo, Discusión), and a sidebar with various tools and language options. The main content area contains a detailed description of the university, its history, and its academic structure. A right-hand sidebar provides a structured overview of the university's key information, including its acronym (UAM), location, founding year (1968), and website.

Universidad Autónoma de Madrid

La **Universidad Autónoma de Madrid (UAM)** es una **universidad pública** que se fundó en 1968, momento en que sus facultades estaban dispersas por diversos edificios de la capital española. No obstante, la localización actual de esta universidad es el campus de Cantoblanco, al norte de la ciudad de **Madrid**, junto a **Alcobendas** y **San Sebastián de los Reyes**. Dicho campus, con 2.252.000 m² de superficie total, se inauguró el 25 de octubre de 1971.

Es una de las seis universidades públicas de la Comunidad de Madrid junto a la Universidad Complutense de Madrid, la Universidad Carlos III de Madrid, la Universidad Politécnica de Madrid, la Universidad de Alcalá y la Universidad Rey Juan Carlos.

Cuenta con siete facultades: Ciencias, Derecho, Filosofía y Letras, Psicología, Medicina (situada fuera del Campus de Cantoblanco), Ciencias Económicas y Empresariales, Formación de Profesorado y Educación y la Escuela Politécnica Superior, además de siete Escuelas Universitarias adscritas (cinco de Enfermería, una de Fisioterapia y una de Magisterio); todo ello estructurado en 70 Departamentos.² También cuenta con numerosos Institutos de investigación propios y centros del Consejo Superior de Investigaciones Científicas (CSIC) asociados. Su estudios de Medicina, gracias a su Hospital Universitario La Paz y a las excelentes calificaciones que sus alumnos obtienen en el MIR son especialmente prestigiosos, así como en Biología, por su estrecha relación con diversos centros del CSIC, Derecho e Historia. Su Departamento de Matemáticas figura entre los cincuenta mejores del mundo en los estudios más reconocidos. Es la Universidad de España con más investigadores de impacto, según los rankings internacionales.

Historia

Al igual que la Universidad Autónoma de Barcelona y la Universidad de Bilbao (posteriormente llamada Universidad País Vasco), la Universidad Autónoma de Madrid nació gracias al Decreto-Ley 5/1968 aprobado por el Consejo d

The screenshot shows the DBpedia page for 'Universidad Autónoma de Madrid'. The page displays the entity's type (University) and its location within the DBpedia data space. A table lists various properties and their corresponding values, such as the acronym (UAM), affiliation (UNICA, EUA, CRUE, CRUMA), city (Madrid), and founding year (1968). The page also includes a search bar and navigation icons.

About: La Autónoma

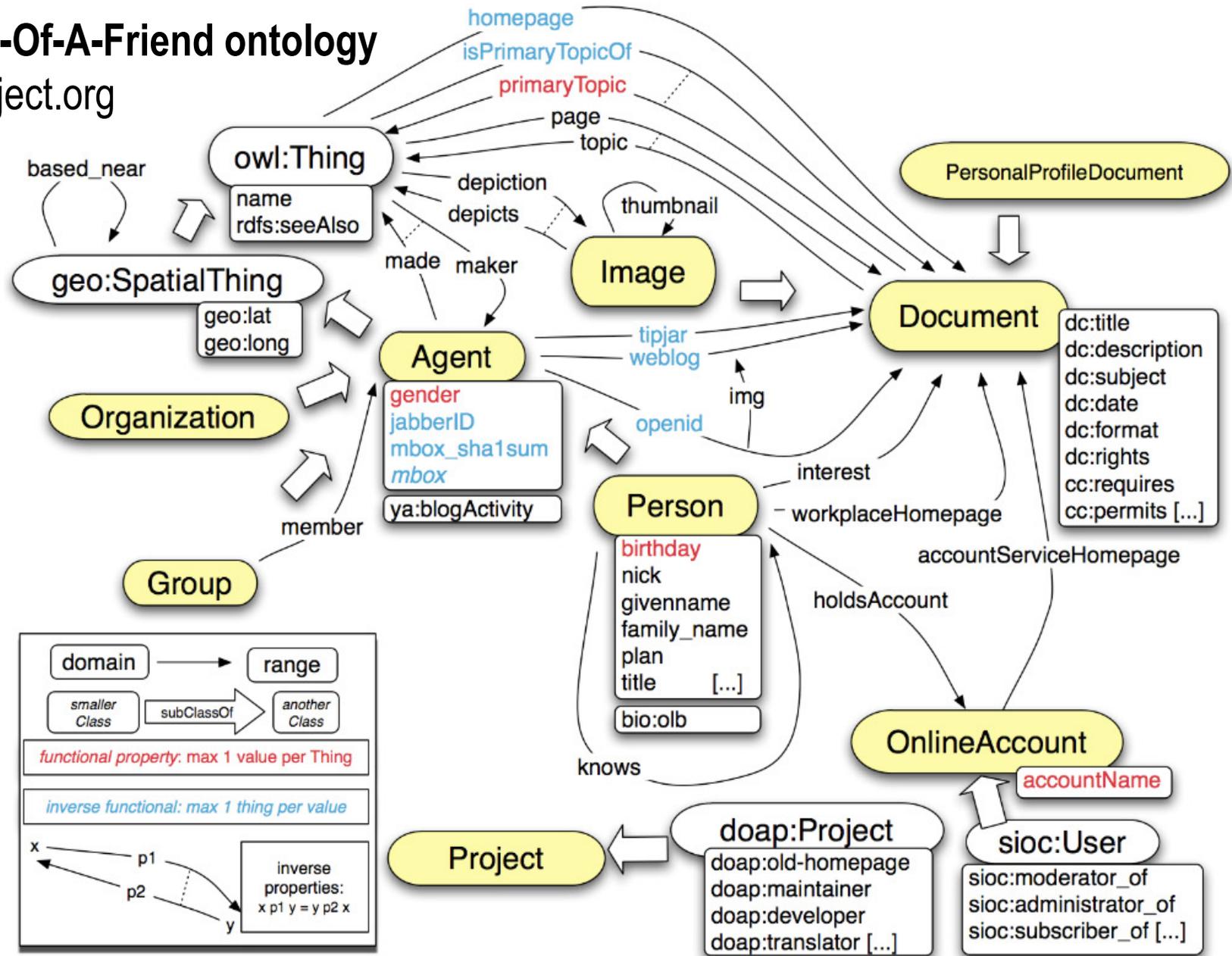
An Entity of Type : [University](#), from Named Graph : <http://es.dbpedia.org>, within Data Space : <es.dbpedia.org>

Property	Value
http://es.dbpedia.org/property/acrónimo	■ UAM
http://es.dbpedia.org/property/afiliación	■ UNICA, EUA, CRUE, CRUMA
http://es.dbpedia.org/property/ciudad	■ dbpedia:Madrid
http://es.dbpedia.org/property/dirección	■ 28049 (xsd:integer) ■ Ctra. de Colmenar km. 15 ■ Ciudad Universitaria de Cantoblanco
http://es.dbpedia.org/property/estudiantes	■ 32446 (xsd:integer)
http://es.dbpedia.org/property/fundación	■ 1968 (xsd:integer)
http://es.dbpedia.org/property/lema	■ Quid Ultra Faciam?
http://es.dbpedia.org/property/lemaesp	■ ¿Qué más debemos hacer?
http://es.dbpedia.org/property/nombre	■ Universidad Autónoma de Madrid
http://es.dbpedia.org/property/pais	■ dbpedia:España
http://es.dbpedia.org/property/rector	■ dbpedia:José_María_Sanz_Martinez
http://es.dbpedia.org/property/sitioWeb	■ http://www.uam.es
http://es.dbpedia.org/property/sobrenombre	■ La Autónoma
http://es.dbpedia.org/property/teléfono	■ +34 91 497 50 00
http://es.dbpedia.org/property/tipo	■ Pública
http://es.dbpedia.org/property/wikiPageUsesTemplate	■ dbpedia:Plantilla:Ficha_de_universidad
dcterms:subject	■ dbpedia:Categoría:Universidad Autónoma de Madrid ■ dbpedia:Categoría:Campus de Excelencia Internacional
georss:point	■ 40.5453 -3.69611
rdf:type	■ owl:Thing ■ http://www.opengis.net/gml/_Feature ■ dbpedia-owl:Agent ■ dbpedia-owl:Organisation ■ http://schema.org/Organization ■ dbpedia-owl:University ■ dbpedia-owl:EducationalInstitution ■ http://schema.org/EducationalOrganization ■ http://schema.org/CollegeOrUniversity
geo:lat	■ 40.545300 (xsd:float)
geo:long	■ -3.696110 (xsd:float)
foaf:homepage	■ http://www.uam.es

Linked Open Data

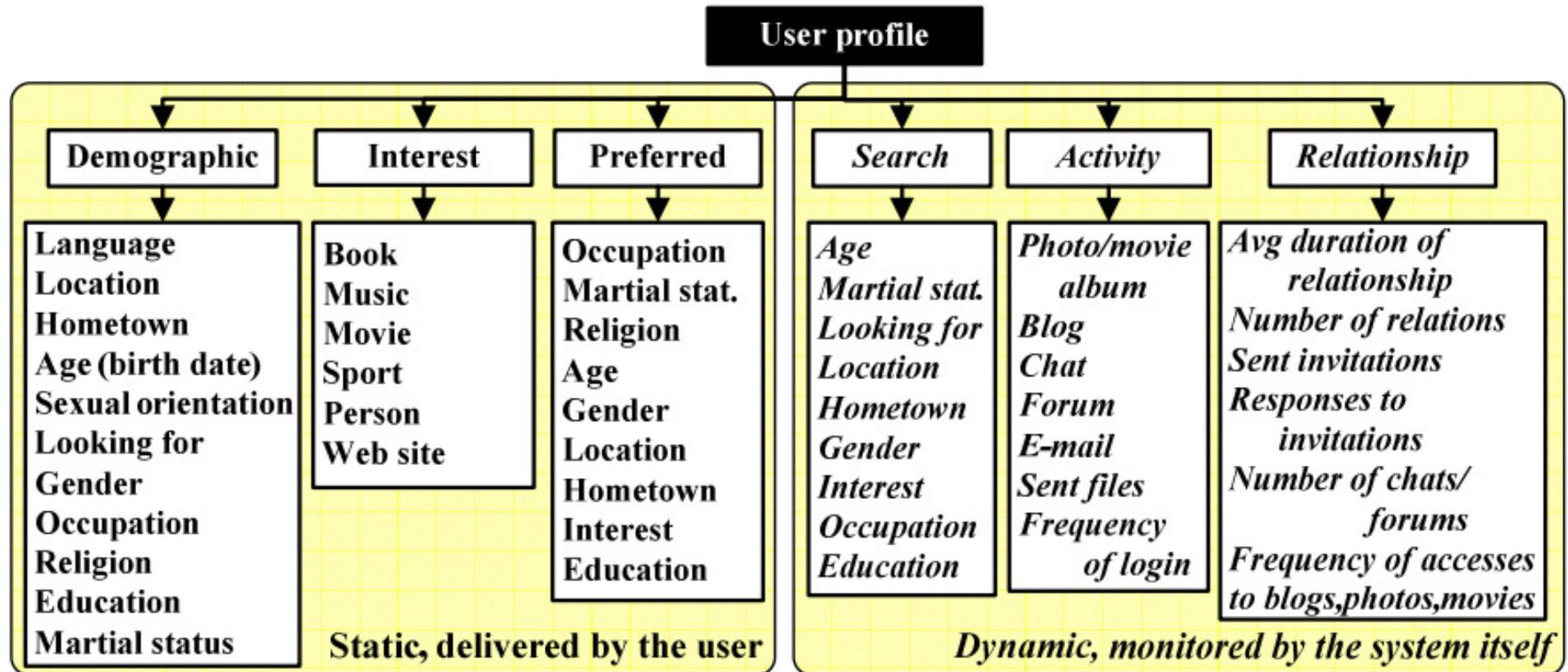
FOAF: The Friend-Of-A-Friend ontology

<http://www.foaf-project.org>

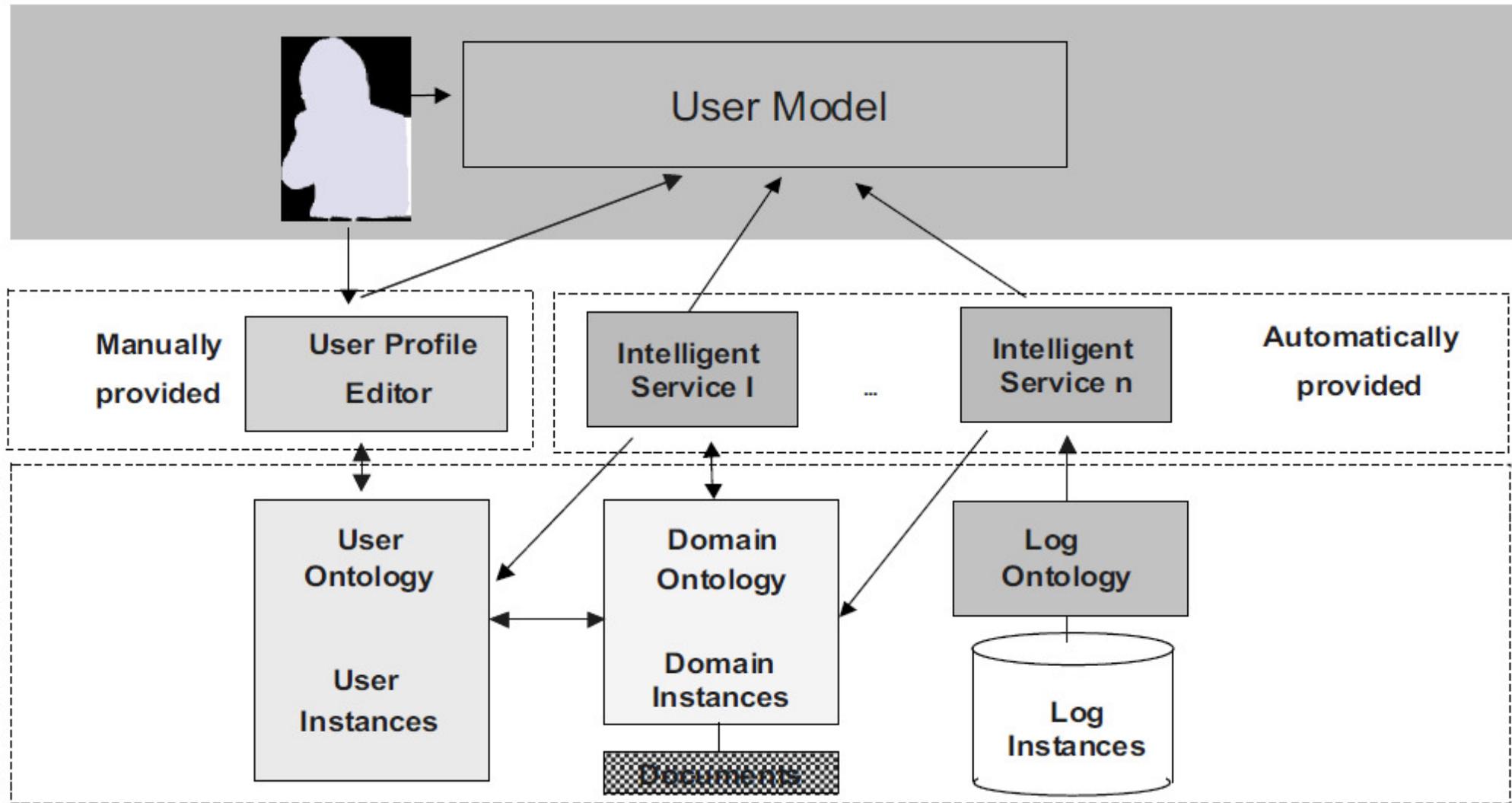


- The Social Semantic Web
 - Web 2.0: The Social Web
 - Web 3.0: The Semantic Web
- **Ontology-based User Modeling**
 - **Ontology-based User Preferences**
 - Mapping Social Tags to Ontology Concepts
 - Mapping Facebook Likes to DBpedia Entities
- Ontology-based Recommendation
 - An Example of Ontology-based Recommender System
 - A Linked Data-based Cross-domain Recommendation Approach
- Conclusions

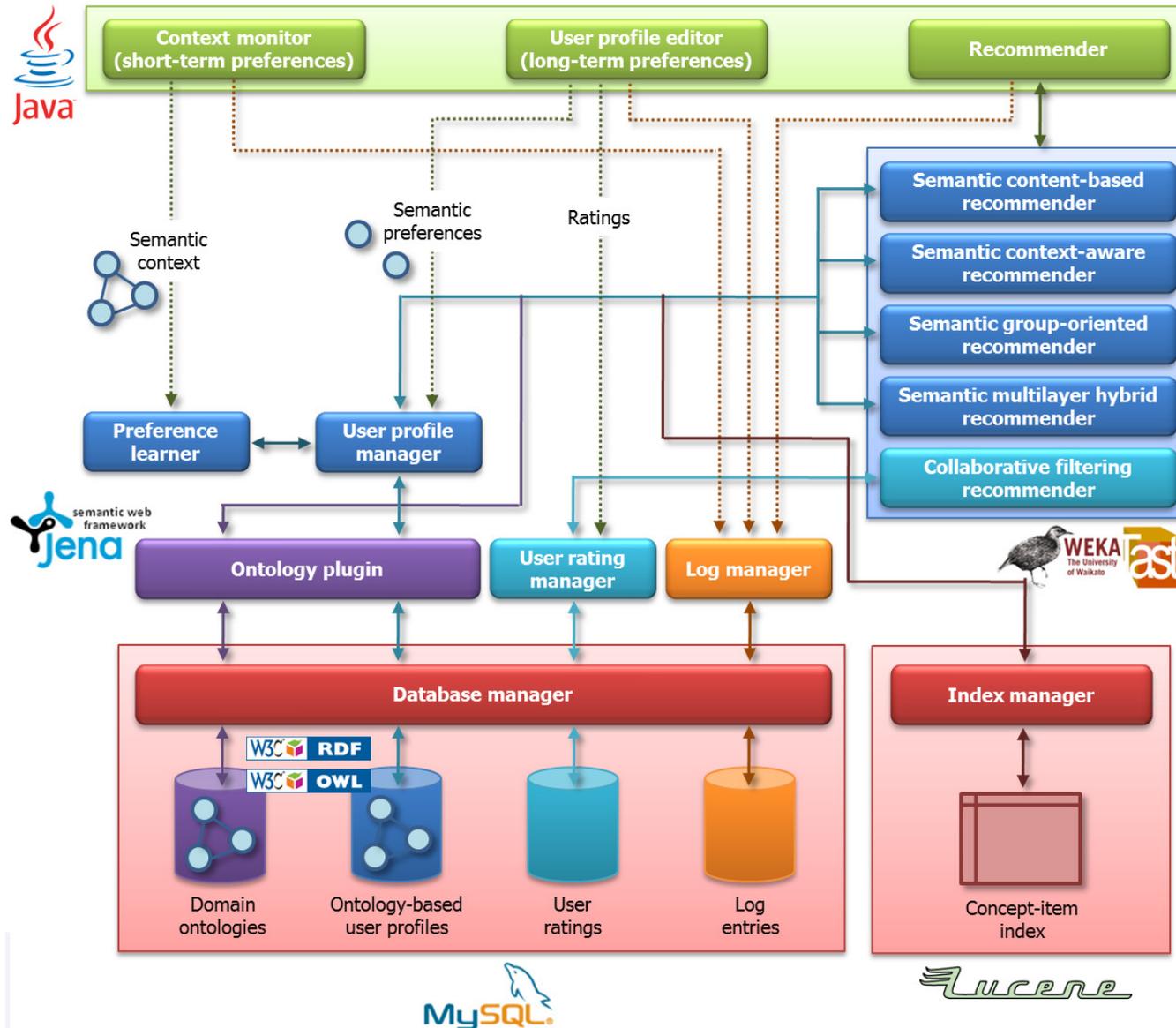
- What data from a user could be modeled, gathered, learned, and exploited?
 - Among other things...



Musial, K. 2009. *Recommender System for Online Social Network*. Lambert Academic Publishing.



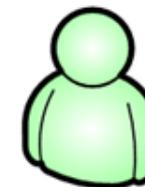
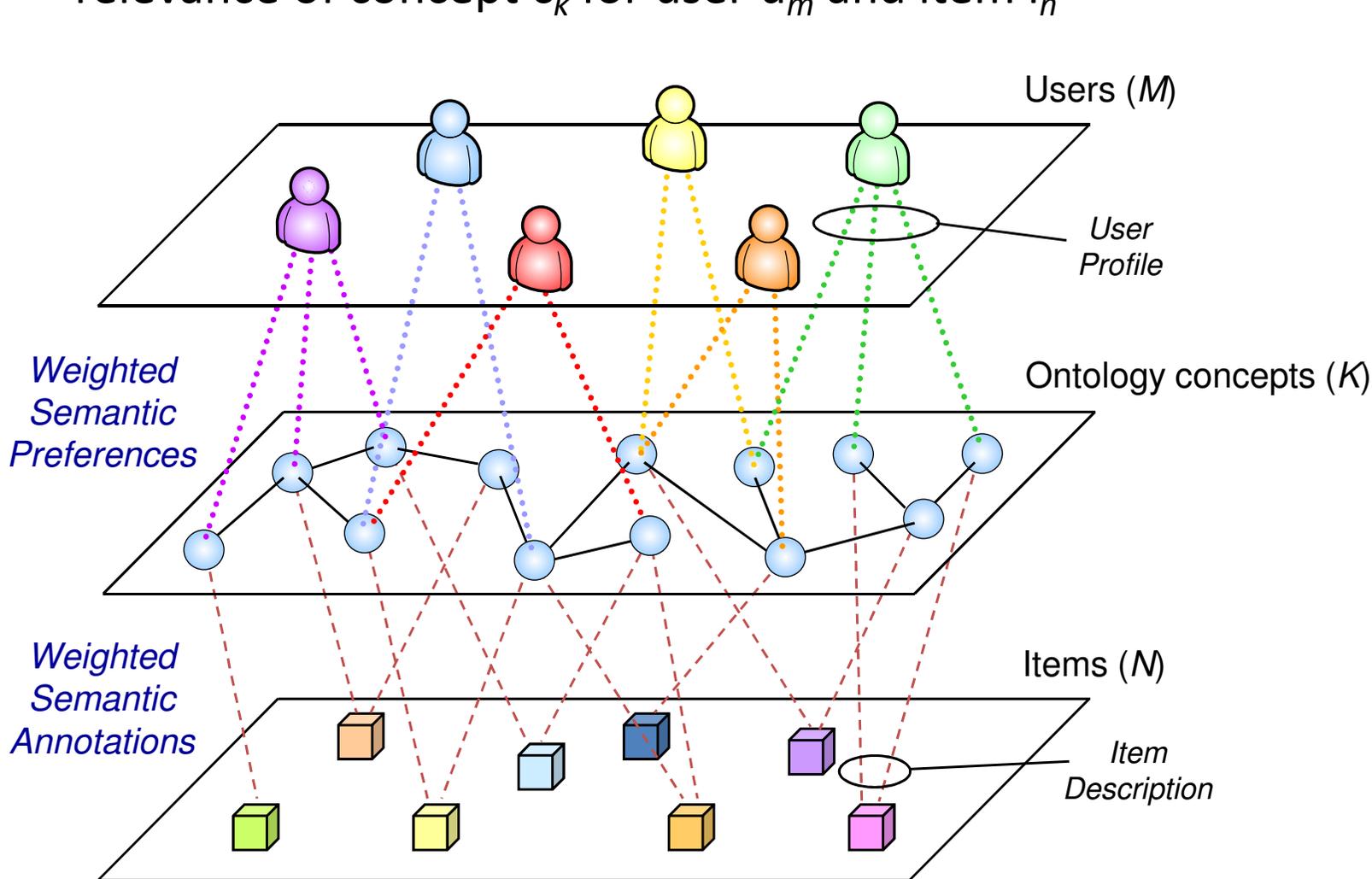
Razmerita, L., Angehrn, A., Maedche, A. 2003. *Ontology-Based User Modeling for Knowledge Management Systems*. In *Proceedings of the 9th International Conference on User Modeling (UM 2003)*, pp. 213-217.



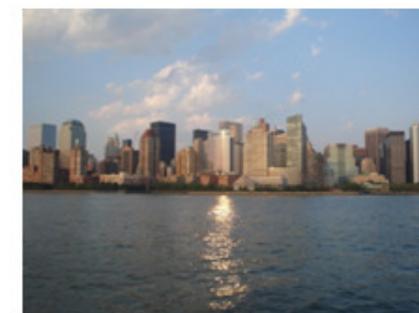
Cantador, I. 2008. *Exploiting the Conceptual Space in Hybrid Recommender Systems: A Semantic-based Approach*. PhD thesis, Universidad Atónoma de Madrid.

Ontology-based User Preferences

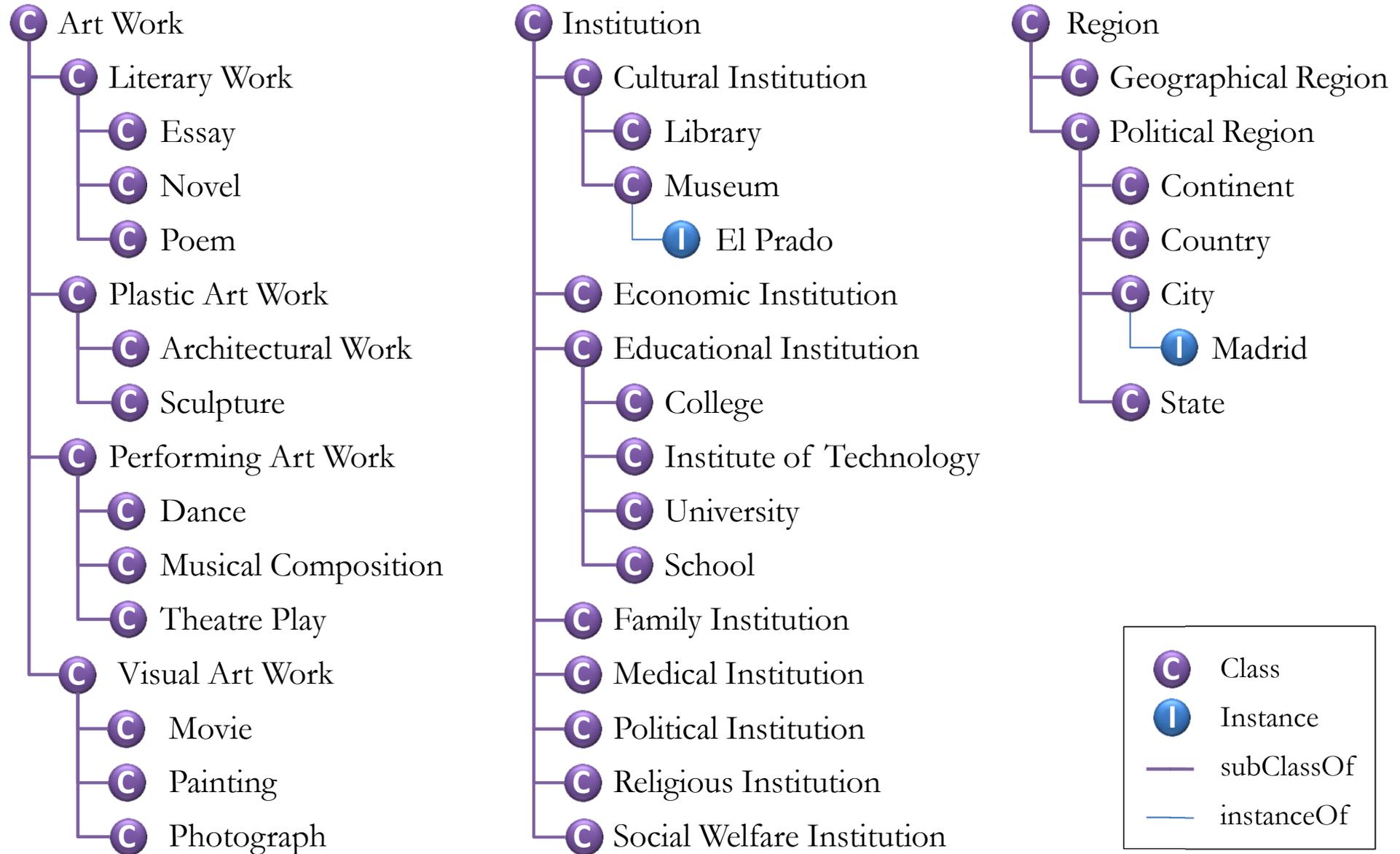
- User and item profiles are represented as vectors $\mathbf{u}_m = (u_{m,1}, \dots, u_{m,K}) \in [-1, +1]^K$ and $\mathbf{i}_n = (i_{n,1}, \dots, i_{n,K}) \in [0, +1]^K$, where $u_{m,k}, i_{n,k}$ are the weights that measure the relevance of concept c_k for user u_m and item i_n

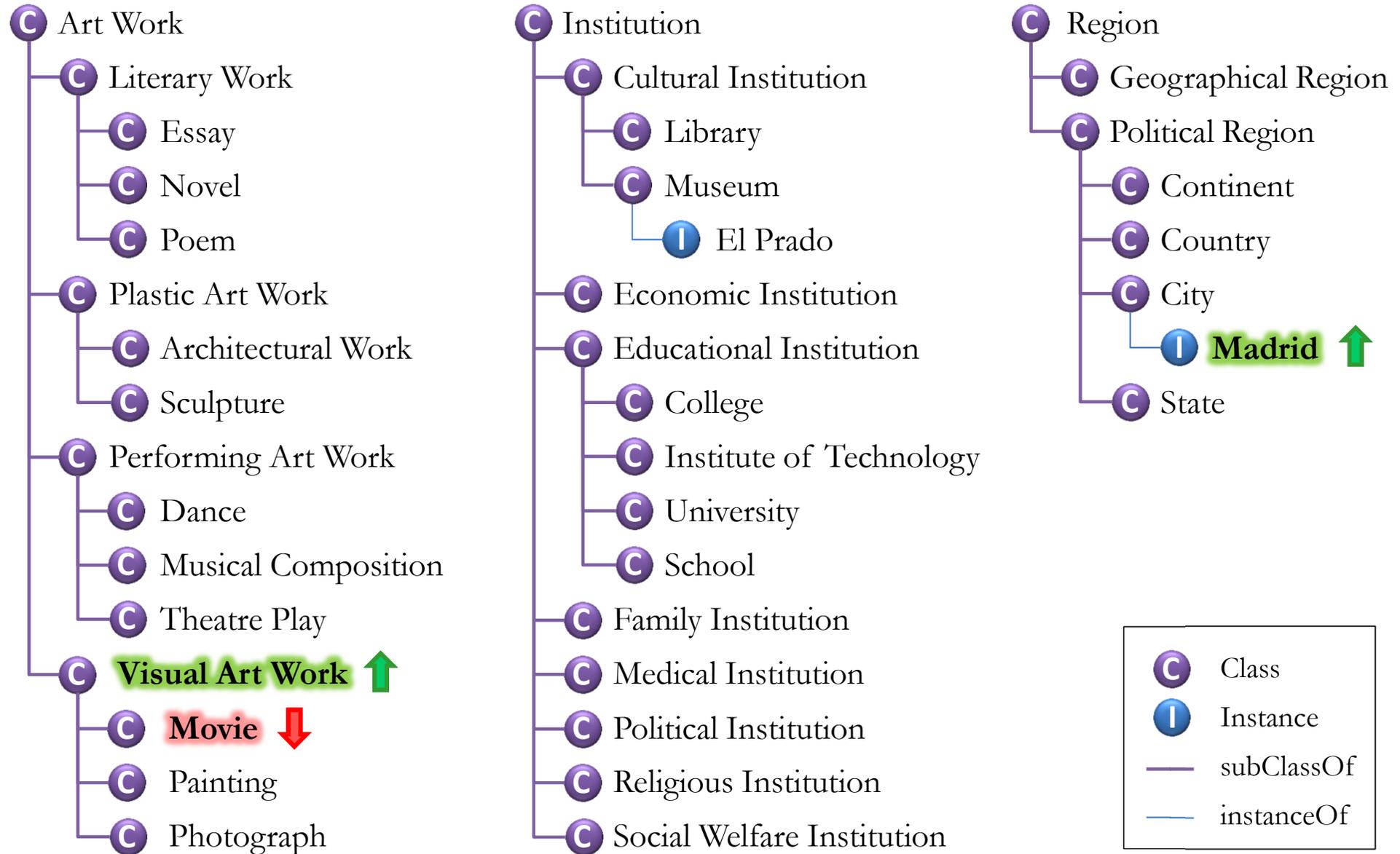


User preferences	
Class	Weight
Flower	-0.3
Mountain	0.9
Sea	0.7
Sky	0.5
Tree	-0.1

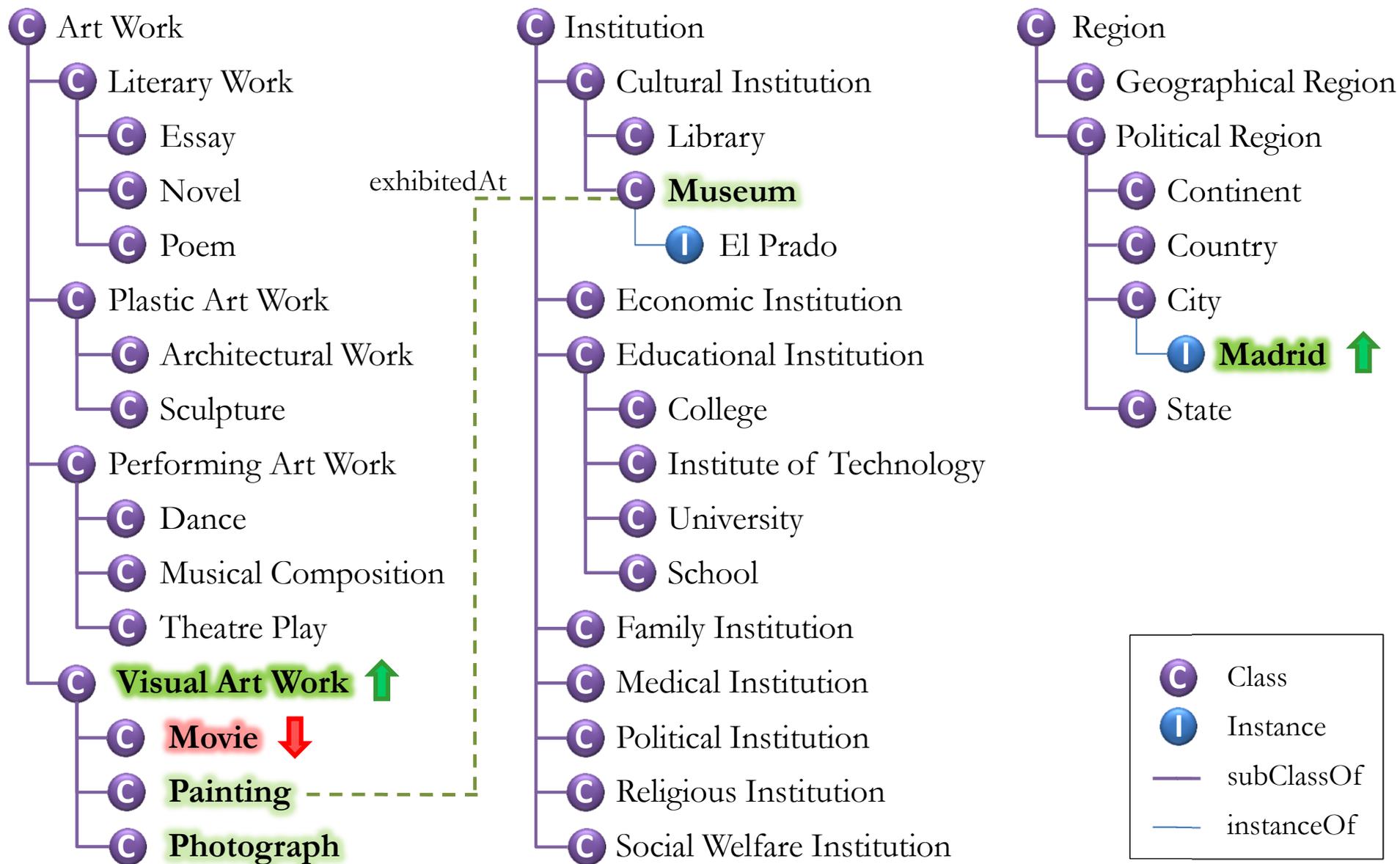


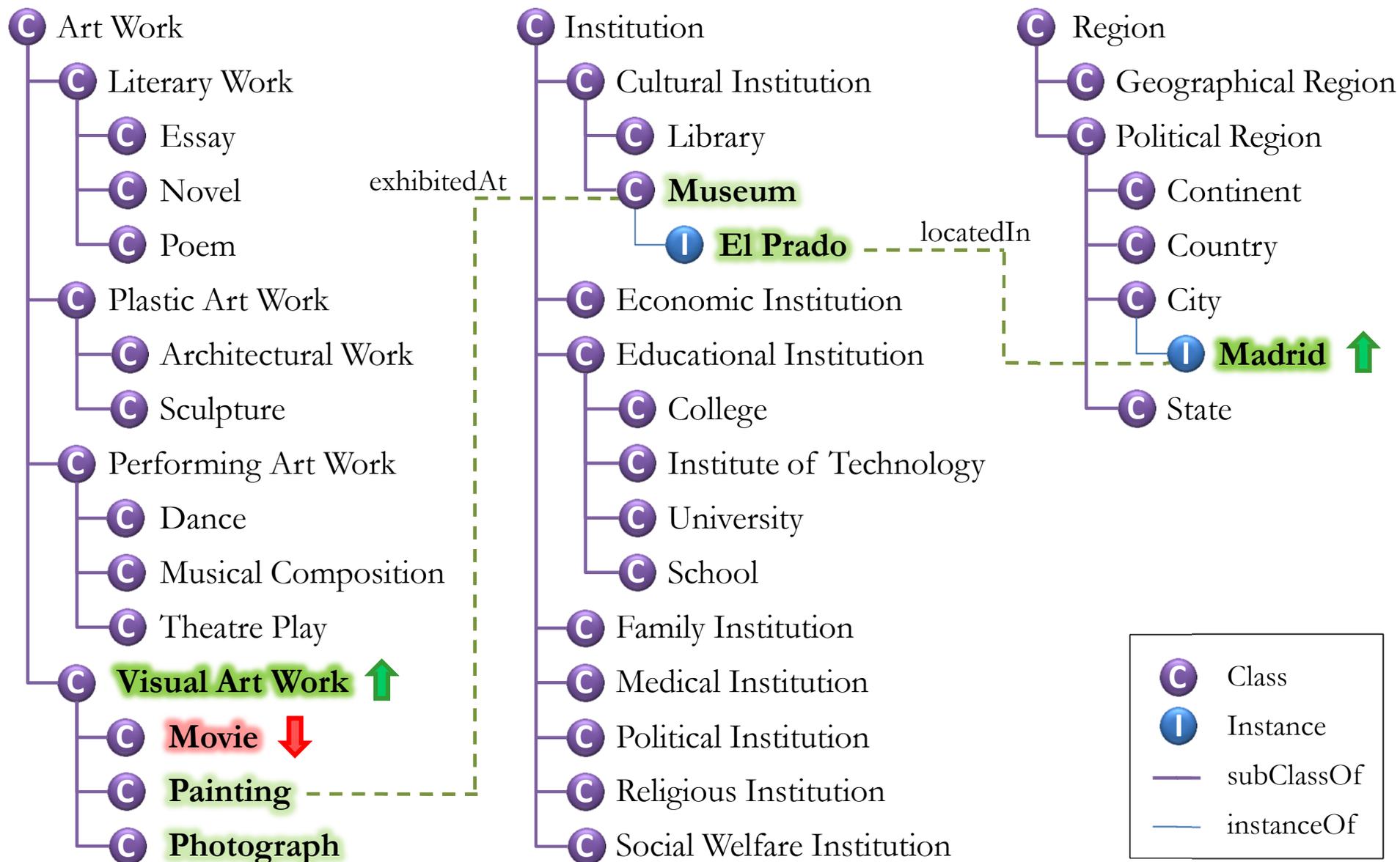
Content metadata	
Class	Weight
Building	0.8
Sea	0.6
Sky	0.4





Ontology-based User Preferences





- The Social Semantic Web
 - Web 2.0: The Social Web
 - Web 3.0: The Semantic Web
- Ontology-based User Modeling
 - Ontology-based User Preferences
 - **Mapping Social Tags to Ontology Concepts**
 - Mapping Facebook Likes to DBpedia Entities
- Ontology-based Recommendation
 - An Example of Ontology-based Recommender System
 - A Linked Data-based Cross-domain Recommendation Approach
- Conclusions

- In social tagging systems, **users** upload or create content (**items**), annotate it with freely chosen words (**tags**), and share these annotations with others
- The nature of tagged items is manifold:
 - *photos* (Flickr)
 - *music tracks* (Last.fm)
 - *video clips* (YouTube)
 - *movies* (MovieLens)
 - *web pages* (Delicious)
 - *scientific articles* (CiteULike, BibSonomy)
 - ...



- **Social tagging** is a practice of **manually** creating and managing tags to **annotate** and **categorize content**
 - This practice is also known as *social classification*, *social indexing*, and ~~*collaborative tagging*~~
- **Folksonomy**, a term coined by Thomas Vander Wal from a mailing list conversation with Gene Smith, is a portmanteau of **folk** and **taxonomy**

Smith, G. 2004. Folksonomy: Social Classification

http://atomiq.org/archives/2004/08/folksonomy_social_classification.html

Vander Wal, T. 2007. Folksonomy Coinage and Definition

<http://vanderwal.net/folksonomy.html>

- Delicious, <http://www.delicious.com>

The screenshot shows the Delicious website interface. At the top, there are navigation links: "Join Now!", "What's New?", "Learn more", "Help", and "Sign In". Below these are menu items: "delicious", "Home", "Bookmarks", "People", and "Tags". A search bar contains the text "bolzano university" and a "Search" button. Below the search bar, it says "Show: Everybody's bookmarks (60)".

The search results are displayed in a list. Each result includes a title, a URL, and a set of tags. The number of users who tagged the item is shown in a blue box on the right of each result.

Item	URL	Tags	Users
Free University of Bozen-Bolzano - Home Unibz	www.unibz.it/	italy, uni, www.provincia.bz.it, design, bolzano	19
Free University of Bozen · Bolzano - Main page	www.unibz.it/en/public/university/default.html	education, italy, lampo, bolzano, university	8
Free University of Bozen · Bolzano - Welcome	www.unibz.it/en/students/welcome/default.html	bolzano, university	
European Masters Program in Computational Logic	www.computational-logic.eu/home.php	logic, master, learning, ar, course	20
Free University of Bozen · Bolzano - News Overview	www.unibz.it/en/search/newsoverview.html?NewsID=11650&language=EN	italy, chatbot, english, library2.0, university	
Zibaldino	web.mac.com/edamiani/Site/Blog/Blog.html	professor, damiani, fub, blog, university	

tagged item (web page)

users who tagged the item

tags assigned to the item

- Delicious, <http://www.delicious.com>

tag-based search

The screenshot shows the Delicious website interface. At the top, there are navigation links: "Join Now!", "What's New?", "Learn more", "Help", and "Sign In". Below these are menu items: "delicious", "Home", "Bookmarks", "People", and "Tags". A search bar contains the text "bolzano university" and a "Search" button. Below the search bar, there is a "Show:" section with a checked box for "Everybody's bookmarks (60)". The search results are displayed in a list format. The first result is "Free University of Bozen-Bolzano - Home Unibz" with 19 bookmarks and tags: "italy", "uni", "www.provincia.bz.it", "design", "bolzano". The second result is "Free University of Bozen - Bolzano - Main page" with 8 bookmarks and tags: "education", "italy", "lampo", "bolzano", "university". The third result is "Free University of Bozen - Bolzano - Welcome" with tags: "bolzano", "university". The fourth result is "European Masters Program in Computational Logic" with 20 bookmarks and tags: "logic", "master", "learning", "ar", "course". The fifth result is "Free University of Bozen - Bolzano - News Overview" with tags: "italy", "chatbot", "english", "library2.0", "university". The sixth result is "Zibaldino" with tags: "professor", "damiani", "fub", "blog", "university".

Room for improvement!

non-relevant results

- Tags are free text, and thus suffer from various vocabulary problems
 - **Polysemy:** bridge, jaguar, java
 - **Synonym:** biscuit, cookie
 - **Morphological deviations:** blog, blogs, blogging
 - **Multilinguality:** spain, españa, spagna, espagne
 - **Contemporaneous terminology:** podcast, ipod, diy
 - **Misspellings:** barclona (*instead of barcelona*)
 - **Compound nouns:** new_york, thelordoftheringsmovie
 - **Subjective/personal concepts:** funny, my wife, to read

Room for improvement!

Mathes, A. 2004. Folksonomies - Cooperative Classification and Communication through Shared Metadata. Computer Mediated Communication - LIS590CMC, Graduate School of Library and Information Science, University of Illinois Urbana-Champaign, IL, USA.



To address the above problems,
folskonomy-based information retrieval and filtering systems
have to identify and exploit the underlying **semantics** of the tags

- Let's tag this Flickr photo!



- Let's tag this Flickr photo!



bolzano · trentino · italy ·
alps · landscape · mountains ·
cathedral

- Let's tag this Flickr photo!



bolzano · trentino · italy ·
alps · landscape · mountains ·
cathedral ·
italia · alpi · montagna · duomo

- Let's tag this Flickr photo!



bolzano · trentino · italy ·
alps · landscape · mountains ·
cathedral ·
italia · alpi · montagna · duomo ·
beautiful · unforgattablepictures ·
holidays · summer · 2010 ·
nikon · 50mm



Users have different intentions when tagging, and not all tags available in a folksonomy are related to the content of the annotated items

- Tagging purposes/intentions

Cantador et al.	Xu et al.	Sen et al.	Golder et al.	Bischoff et al.
Content-based	Content-based	Factual	What or who is about	Topic
	Attribute		What it is	Type
			Who owns it	Author/owner
Context-based	Context-based		Refining other categories	Time
		Location		
Subjective	Subjective	Subjective	Qualities/characteristics	Opinions/qualities
Organisational	Organisational	Personal	Task organisation	Usage context
			Self reference	Self reference

Cantador, I., Konstas, I., Jose, J. M. 2010. *Categorising Social Tags to Improve Folksonomy-based Recommendations*. *Journal of Web Semantics*. Elsevier. In press.

Xu, Z., Fu, Y., Mao, J., Su, D. 2006. *Towards the Semantic Web: Collaborative Tag Suggestions*. In *Proceedings of the WWW'06 Collaborative Web Tagging Workshop*.

Sen, S., Lam, S. K., Rashid, A. M., Cosley, D., Frankowski, D., Osterhouse, J., Harper, M. F., Riedl, J. 2006. *Tagging, Communities, Vocabulary, Evolution*. In *Proceedings of the 20th ACM Conference on Computer Supported Cooperative Work (CSCW'06)*, 181-190.

Golder, S. A., Huberman, B. A. 2006. *Usage Patterns of Collaborative Tagging Systems*. *Journal of Information Science* 32(2), 198-208.

Bischoff, K., Firan, C. S., Nejd, W., Paiu, R. 2008. *Can All Tags Be Used for Search?* In *Proceeding of the 17th ACM Conference on Information and Knowledge Management (CIKM'08)*, 203-212.

- 4 categories
 - **Content-based tags**
 - Describe the content of the items; e.g. `vehicle, dog, tree`
 - **Context-based tags**
 - Provide contextual information; e.g. `bolzano, mountain, summer, holidays`
 - **Subjective tags**
 - Express opinions and qualities; e.g. `happy, sunny, contemporary art`
 - **Organizational tags**
 - Define personal usages and tasks; e.g. `to look at, scan for print, myself, our best friends`

- A couple of examples

glasgow → *context*



kilt → *content*



- How do we infer the above categorization?

- A couple of examples

glasgow → *context*



kilt → *content*



- **How do we infer the above categorization?**

glasgow → *city in Scotland* → *location* → *context*

kilt → *Scottish piece of cloth* → *physical entity* → *content*



- **Research questions**

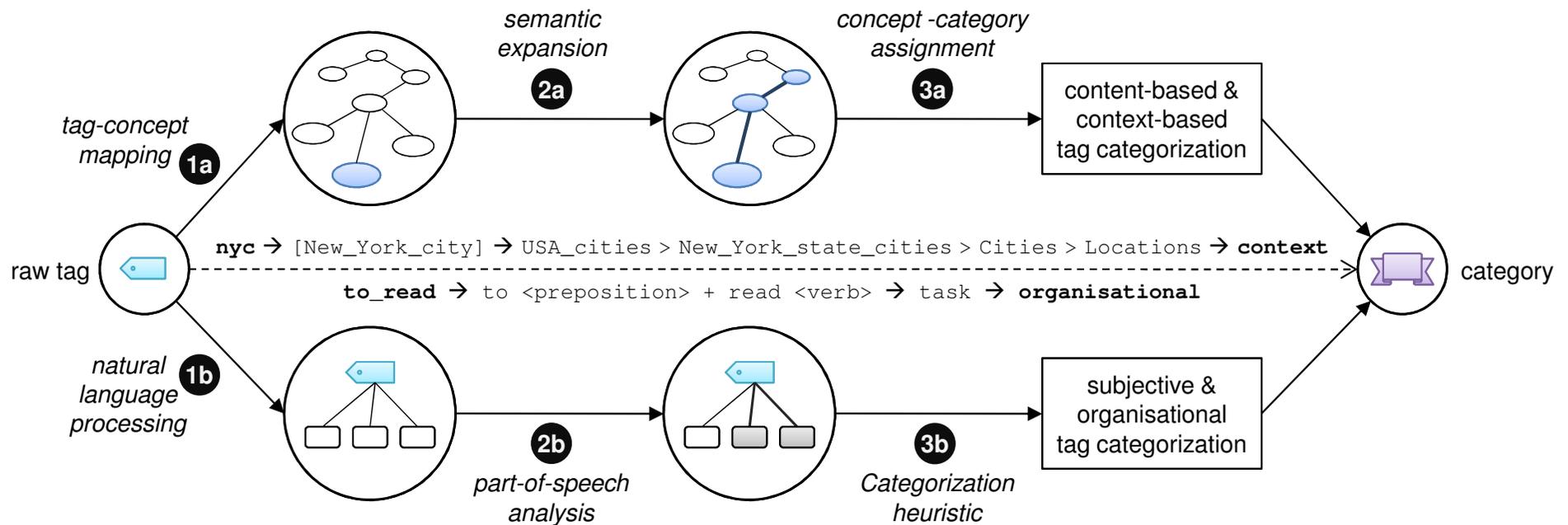
- **RQ1:** How to **automatically categorize tags** based on their purpose/intention?
- **RQ2:** Is a purpose-oriented categorization of tags useful for folksonomy-based **recommendation** strategies?



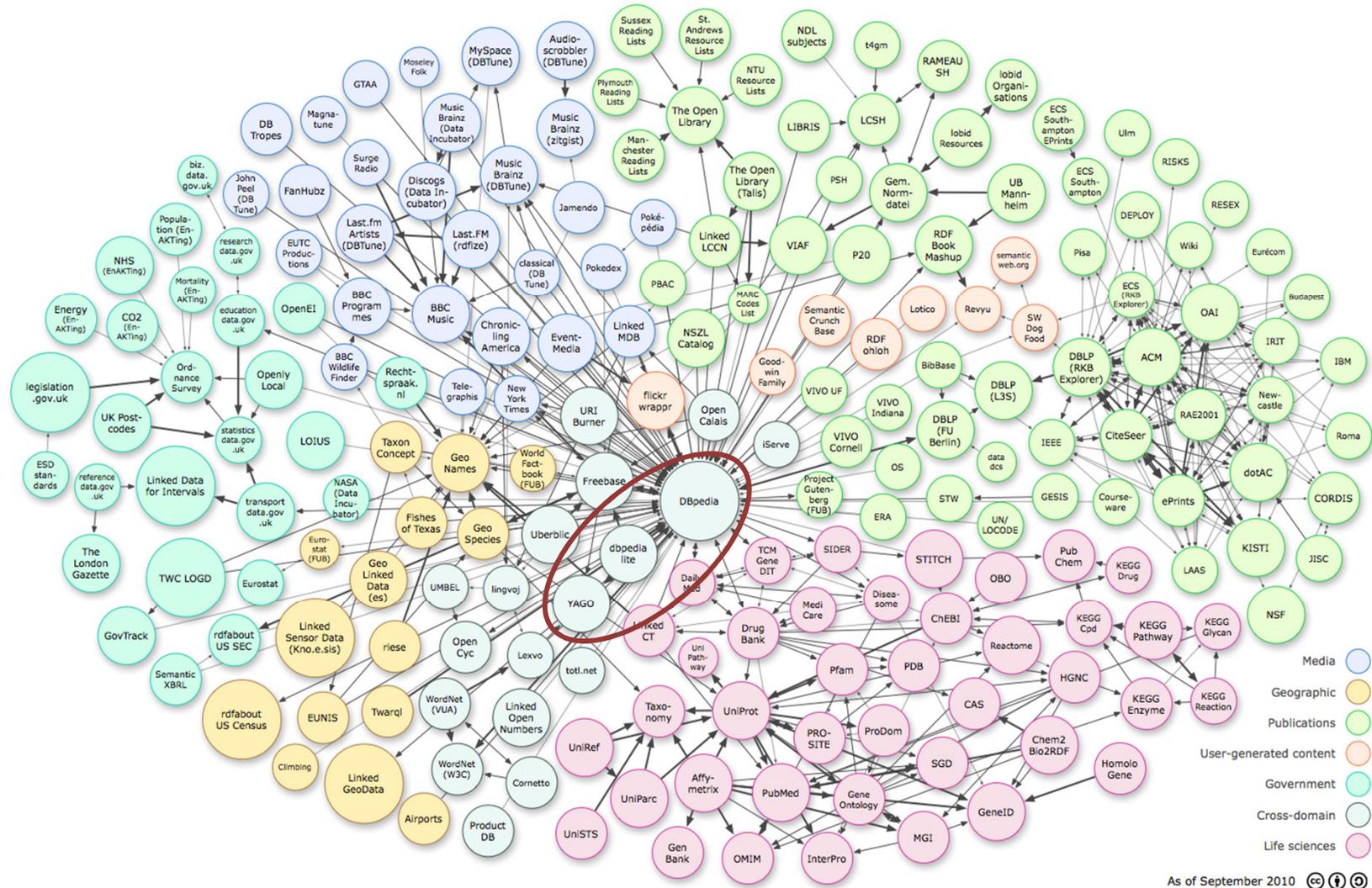
- **Research questions**

- **RQ1:** How to **automatically categorize tags** based on their purpose/intention?
- **RQ2:** Is a purpose-oriented categorization of tags useful for folksonomy-based **recommendation** strategies? **Yes** 😊

- Proposed approach



- Path A: tag-concept mapping
 - **Linked Data** <http://linkeddata.org>



- **Path A:** tag-concept mapping

- **DBPedia** <http://dbpedia.org> – Giving structure to **Wikipedia**

- 3.4M entities: *300K people, 400K places, 146K species, 140K organizations, 94K music albums, 49K films, 15K video games, ...*
- 565K Wikipedia categories
- 75K YAGO categories
- 1.5M links to images
- 5.5M links to web pages
- 4.9M links to RDF repositories

```
{{Infobox Town AT |
name = Innsbruck |
image_coa = InnsbruckWappen.png |
image_map = Karte-tirol-I.png |
state = [[Tyrol]] |
regbzk = [[Statutory city]] |
population = 117,342 |
population_as_of = 2006 |
pop_dens = 1,119 |
area = 104.91 |
elevation = 574 |
lat_deg = 47 |
lat_min = 16 |
lat_hem = N |
lon_deg = 11 |
lon_min = 23 |
lon_hem = E |
postal_code = 6010-6080 |
area_code = 0512 |
licence = I |
mayor = Hilde Zach |
website = [http://innsbruck.at] |
}}
```

Innsbruck	
	
Country	Austria
State	Tyrol
Administrative region	Statutory city
Population	117,342 (2006)
Area	104.91 km ²
Population density	1,119 /km ²
Elevation	574 m
Coordinates	47°16′ N 11°23′ E ⓘ
Postal code	6010-6080
Area code	0512
Licence plate code	I
Mayor	Hilde Zach
Website	www.innsbruck.at ⓘ

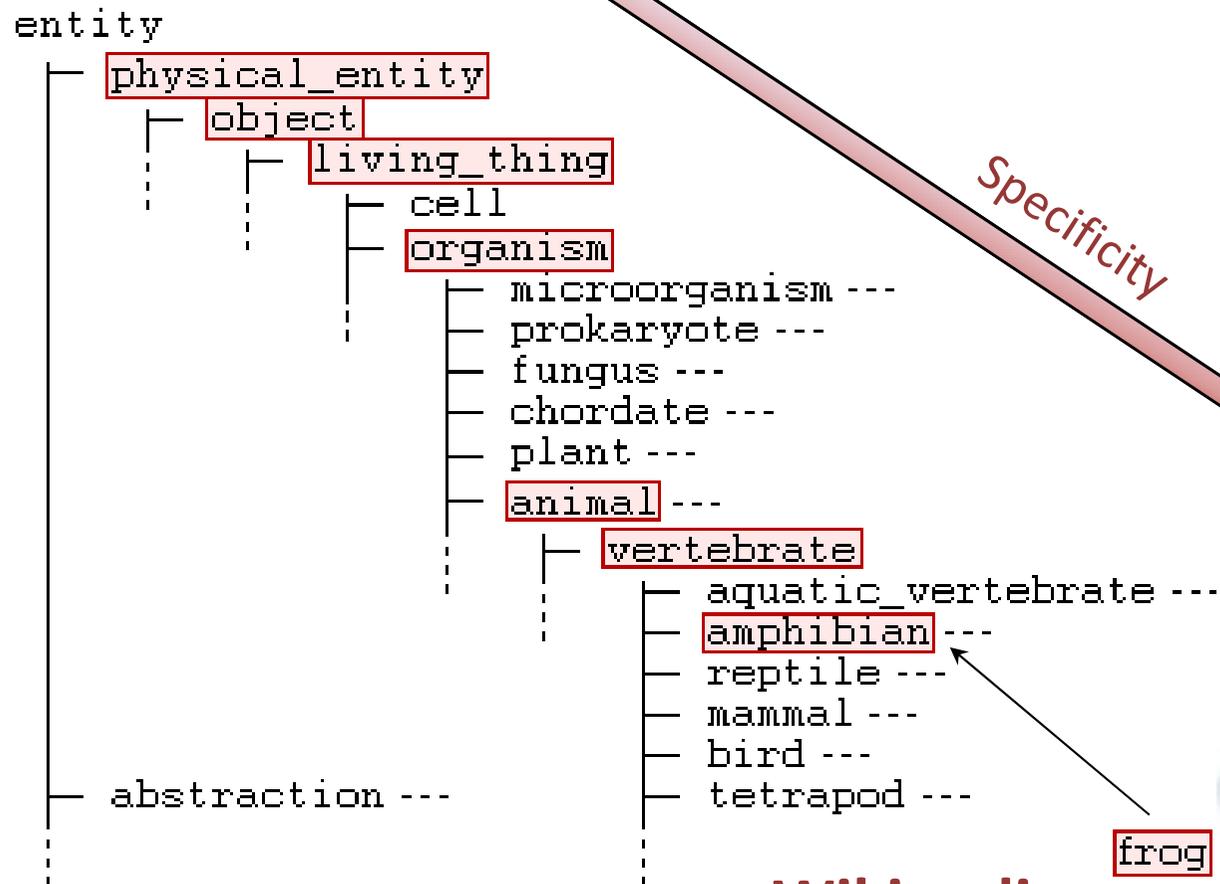
- **Path A**: tag-concept mapping
 - **YAGO (Yet Another Great Ontology)** <http://www.mpi-inf.mpg.de/yago>
 - **DBPedia** concepts + **WordNet** structure
 - >2M entities
 - Manually confirmed accuracy of 95%

Suchanek, F. M. , Kasneci, G., Weikum, G. 2007. YAGO: A Core of Semantic Knowledge. In Proceedings of the 16th International Conference on World Wide Web (WWW'07), 697-706.

Suchanek, F. M. , Kasneci, G., Weikum, G. 2008. YAGO: A Large Ontology from Wikipedia and WordNet. Journal of Web Semantics 6(3), 203-217.

- **Path A:** tag-concept mapping
 - Exploiting YAGO structure: “reference classes”

WordNet



Wikipedia



- **Path A:** tag-concept mapping
 - Exploiting YAGO structure: “reference classes”

Category	Subcategory	YAGO reference classes
<i>Content-based</i>	<i>Physical entity</i>	physical_entity
	└ <i>Artefact</i>	artifact
	└ <i>Living entity</i>	living_thing, life_form, live_body
	└ <i>Animal</i>	animal
	└ <i>Person</i>	person, human_body, kin
	└ <i>Plant</i>	plant, plant_part
	<i>Non-physical entity</i>	abstraction
	└ <i>Organisation</i>	organization
<i>Context-based</i>	<i>Location</i>	location, land, geological_formation, social_group
	<i>Time</i>	time, time_interval, time_period, time_unit

YAGO reference classes associated to the considered content- and context-based subcategories

- **Path A:** tag-concept mapping
 - Tag-concept mapping examples (from a Flickr dataset)

Category	Subcategory	Flickr tag examples
<i>Content-based</i>	<i>Physical entity</i>	food, glue, heart, ice
	└ <i>Artefact</i>	comb, finger, helicopter, table
	└ <i>Living entity</i>	cell, clone, life, mushroom
	└ <i>Animal</i>	caterpillar, frog, pigeon, pet
	└ <i>Person</i>	boy, daniel, friend, sister
	└ <i>Plant</i>	cactus, cereal flower, tree
	<i>Non-physical entity</i>	cloud, feminism, noise, tennis
	└ <i>Organisation</i>	bmw, ibm, religion, rolling stones
<i>Context-based</i>	<i>Location</i>	california, rome, spain, wedding
	<i>Time</i>	halloween, march, sixties, winter

- **Path A:** tag-concept mapping
 - **Ambiguity**

java

**Subcategory: animal**

Java_(chicken)

wikicategory_Chicken_breeds

Subcategory: location

Java,_Georgia

wikicategory_Cities,_towns_and_villages_in_Georgia

Java,_South_Dakota

wikicategory_Towns_in_South_Dakota

Java,_New_York

wikicategory_Towns_in_New_York

Java,(island)

wikicategory_Islands_of_Indonesia

Subcategory: non-physical

Java_(band)

wikicategory_French_hip_hop_groups

Java_(board_game)

wikicategory_Economic_simulation_board_games

Java_(programming_language)

wikicategory_Java_specification_requests

Subcategory: person

Java_(actor)

wikicategory_Film_actors

- Categorization results
 - accuracy
 - 30 subjects, 4K tag assignments

	Without translation	With translation
<i>Content-based</i>	27.8	32.9
<i>Context-based</i>	11.3	13.0
<i>Subjective</i>	13.4	14.6
<i>Organisational</i>	6.7	7.2
<i>Unknown</i>	40.8	32.4

Main category	Subcategory	Category	#evaluated assignments	Accuracy (subcategory)	Accuracy (category)	Accuracy (main category)
Content-based	<i>Physical entity</i>	-	296	55.0%	-	77.5%
	<i>Artefact</i>	<i>Physical entity</i>	537	70.0%	72.5%	72.5%
	<i>Living entity</i>	<i>Physical entity</i>	124	67.8%	75.0%	82.1%
	<i>Animal</i>	<i>Living entity</i>	49	55.0%	55.0%	65.0%
	<i>Person</i>	<i>Living entity</i>	73	75.0%	75.0%	83.3%
	<i>Plant</i>	<i>Living entity</i>	202	86.7%	88.4%	96.0%
	<i>Non-physical entity</i>	-	643	63.1%	-	84.4%
	<i>Organisation</i>	<i>Non-physical entity</i>	8	85.0%	95.0%	96.5%
	<i>Entity</i>	-	108	57.5%	-	57.5%
Context-based	<i>Location</i>	-	723	75.6%	-	75.6%
	<i>Time</i>	-	57	90.0%	-	90.0%
Subjective	<i>Opinion</i>	-	461	70.0%	-	97.5%
	<i>Quality</i>	-	409	82.5%	-	82.5%
Organisational	<i>Self-reference</i>	-	180	85.7%	-	91.4%
	<i>Task</i>	-	9	60.0%	-	66.7%
	<i>Action</i>	-	36	75.0%	-	75.0%
				72.1%	76.8%	80.8%

- **Open research lines**

- need for improving **disambiguation** process

- definition of more POS-patterns

- revision of some POS-patterns:

- e.g. subjective → content-based tags

- [*<adjective><noun>*]: *big house* → *house*

What is the meaning of **sf**?

Contextualization and disambiguation of tags ⁷⁶

- What is the meaning of **sf**?



sf · sanfrancisco · california · usa · city · bridge · goldengate · sea



sf · sciencefiction · fiction · starwars · movie · tatooine · androids · c3po · r2d2

Different meanings!

Contextualization and disambiguation of tags ⁷⁷

- What is the meaning of **web**?



web · webbrowser · browser · browsing · opera · explorer · firefox · chrome · safari



web · socialweb · web20 · socialmedia · facebook · twitter · flickr · lastfm

Different semantic contexts!



- Research questions
 - **RQ1:** How to accurately and efficiently identify the **actual meaning or semantic context** of a tag?
 - **RQ2:** Is a semantic contextualization of tags useful for folksonomy-based user modelling and **recommendation** strategies?

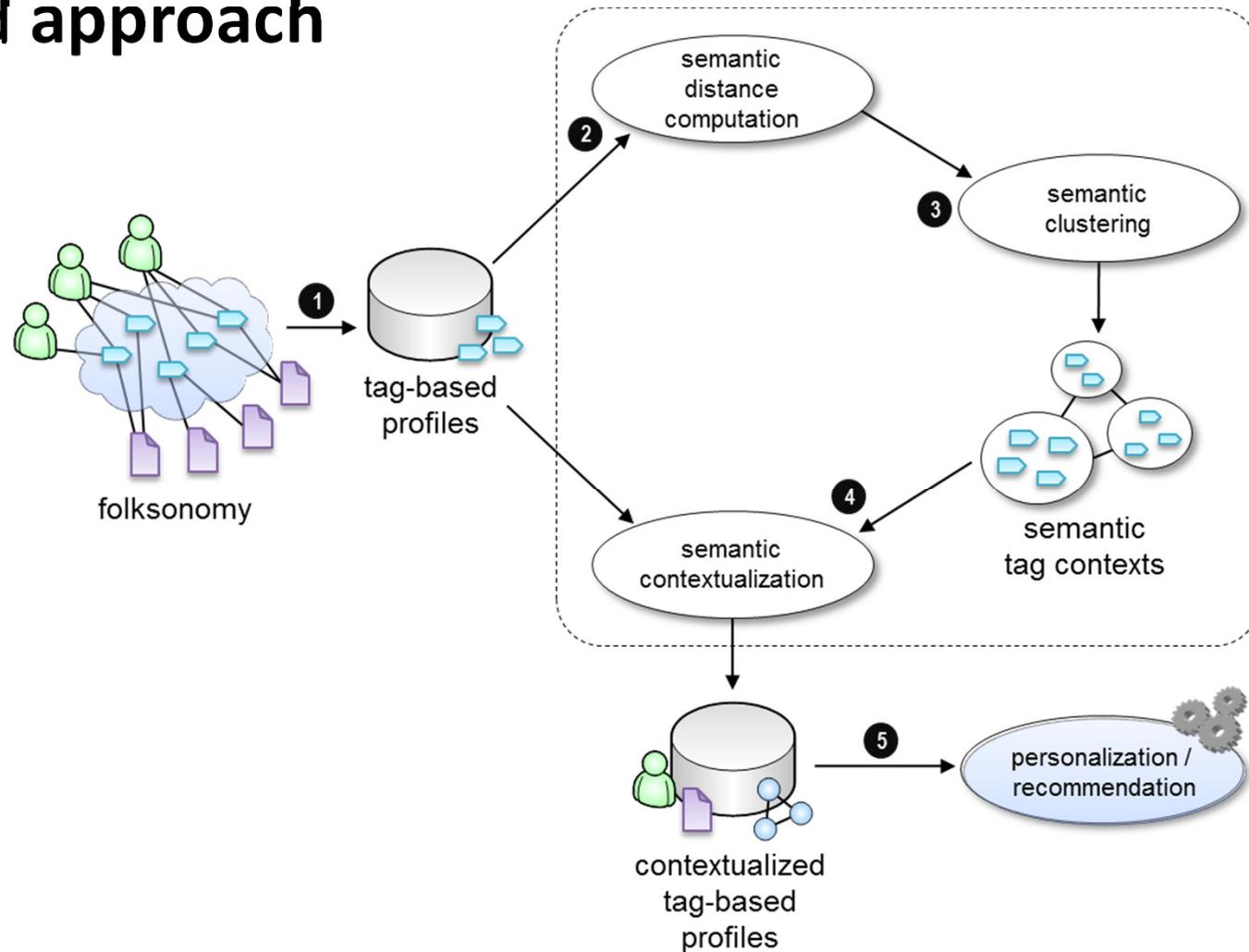


- **Research questions**

- **RQ1:** How to accurately and efficiently identify the **actual meaning or semantic context** of a tag?
- **RQ2:** Is a semantic contextualization of tags useful for folksonomy-based user modelling and **recommendation** strategies? Yes, again! 😊

Contextualization and disambiguation of tags 80

- Proposed approach



Cantador, I., Bellogín, A., Fernández-Tobías, I., López-Hernández, S. 2011. *Semantic Contextualization of Social Tag-based Item Recommendations*. In *Proceedings of the 12th International Conference on E-Commerce and Web Technologies (EC-Web 2011)*, pp. 101-113.

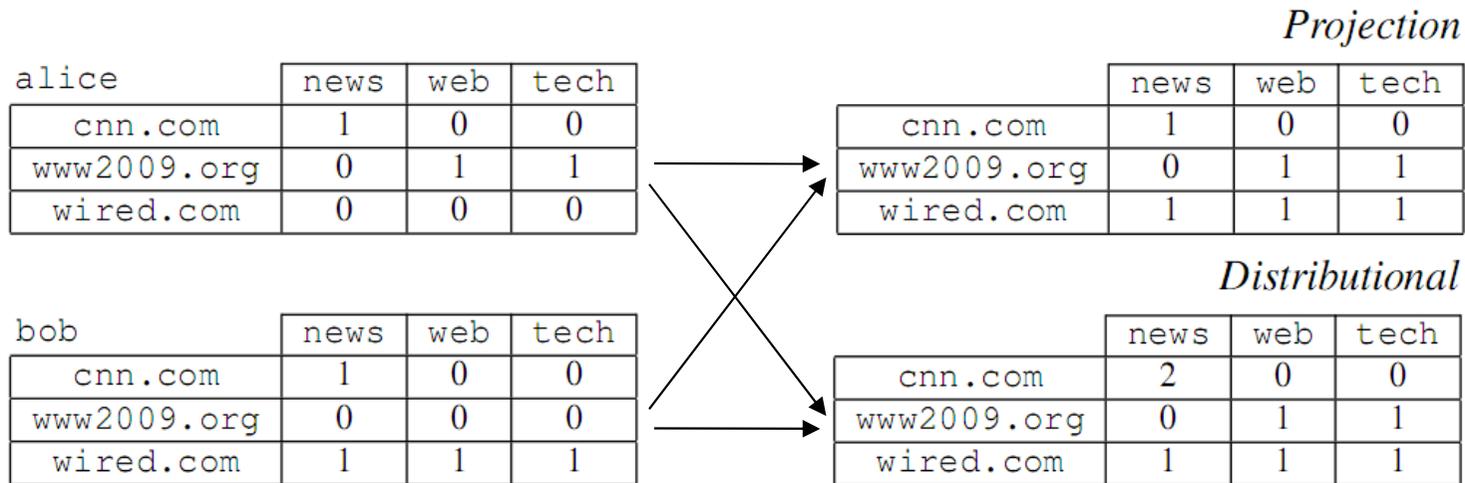
Contextualization and disambiguation of tags ⁸¹

- **Aggregation methods** ([user, item, tag] → [item, tag])
 - Projection
 - Distributional
 - Macro-aggregation
 - Collaborative
- **Similarity measures** ([item, tag] → [tag, tag])
 - Matching
 - Overlapping
 - Jaccard
 - Dice
 - Cosine
 - Mutual information

Markines, B., Cattuto, C., Menczer, F., Benz, D., Hotho, A., Stumme, G. 2009. Evaluating Similarity Measures for Emergent Semantics of Social tagging. In Proceedings of the 18th International Conference on World Wide Web (WWW'09), 641-650.

Contextualization and disambiguation of tags 82

- **Aggregation methods** ([user, item, tag] → [item, tag])



- **Similarity measures** ([item, tag] → [tag, tag])

Overlap similarity

$$\text{sim}(t_1, t_2) = \sum_{r \in \mathcal{R}} w_{t_1} \cdot w_{t_2} = |\mathcal{R}_{t_1} \cap \mathcal{R}_{t_2}|$$

Jaccard similarity

$$\text{sim}(t_1, t_2) = \frac{|\mathcal{R}_{t_1} \cap \mathcal{R}_{t_2}|}{\min(|\mathcal{R}_{t_1}|, |\mathcal{R}_{t_2}|)}$$

Dice similarity

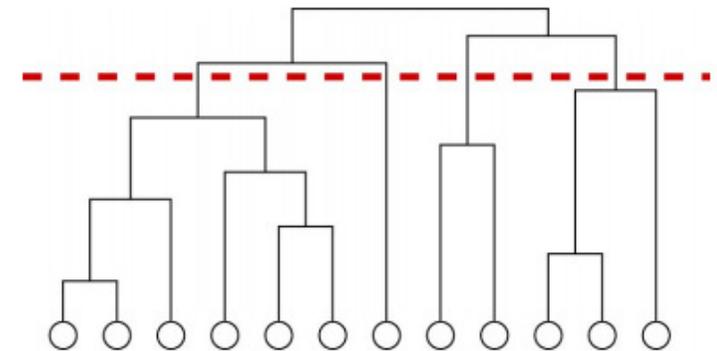
$$\text{sim}(t_1, t_2) = \frac{|\mathcal{R}_{t_1} \cap \mathcal{R}_{t_2}|}{|\mathcal{R}_{t_1} \cup \mathcal{R}_{t_2}|}$$

Cosine similarity

$$\text{sim}(t_1, t_2) = \frac{\mathcal{R}_{t_1}}{\sqrt{|\mathcal{R}_{t_1}|}} \cdot \frac{\mathcal{R}_{t_2}}{\sqrt{|\mathcal{R}_{t_2}|}} = \frac{|\mathcal{R}_{t_1} \cap \mathcal{R}_{t_2}|}{\sqrt{|\mathcal{R}_{t_1}| \cdot |\mathcal{R}_{t_2}|}}$$

Contextualization and disambiguation of tags ⁸³

- Compute the semantic distance between each pair of tags
- Create the whole tag **graph** G based on the above distances
- For each tag t , extract a **subgraph** G_t from G with the tags “most” similar to t
- **Clustering** G_t by using the algorithm proposed by Newman and Girvan, 2004
 - Fast computation
 - Automatic clustering stop criterion

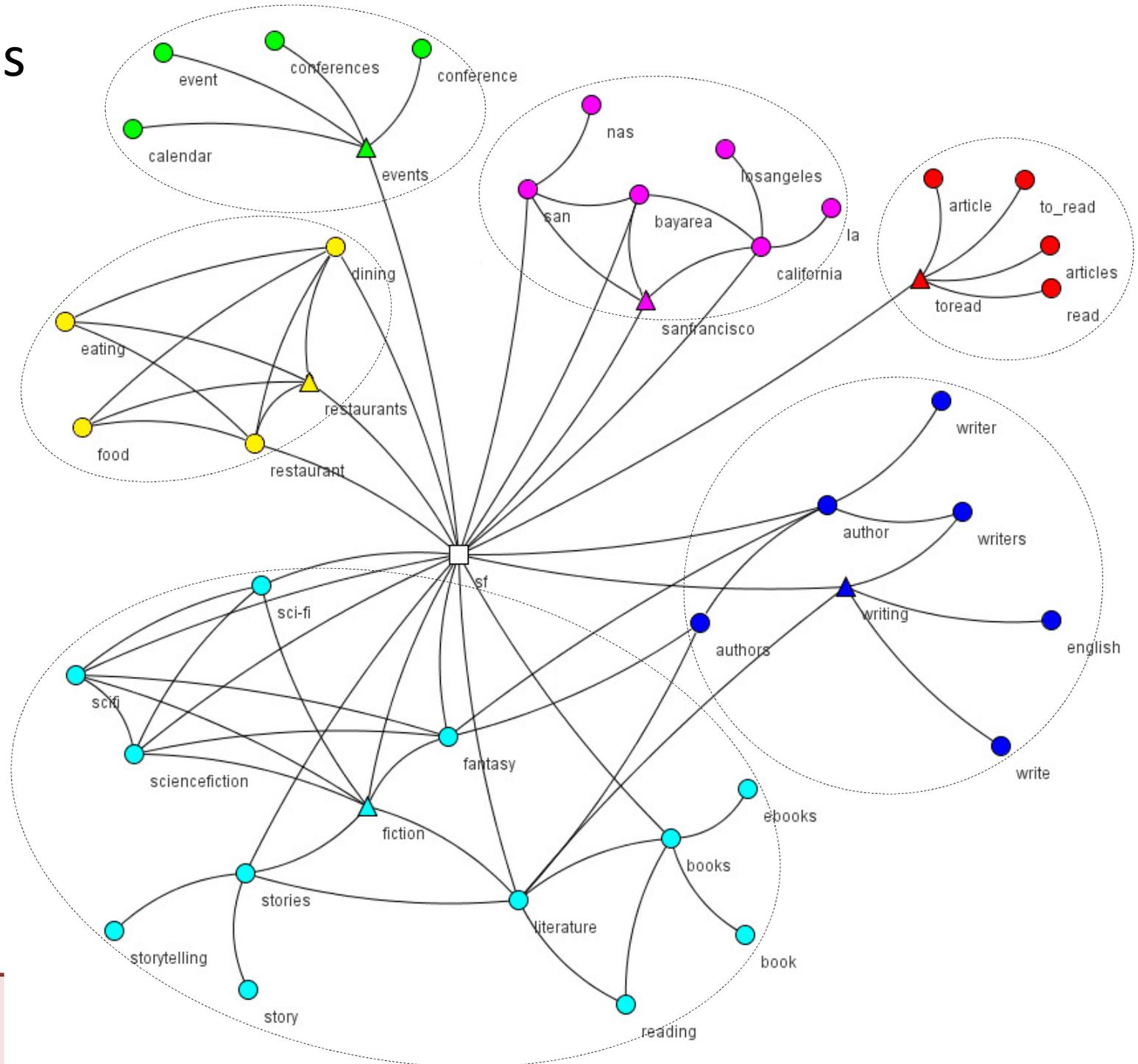


Newman, M. E. J., and Girvan, M. 2004. *Finding and Evaluating Community Structure in Networks*. *Physical Review*, E 69, 026113

Newman, M. E. J. 2006. *Finding Community Structure in Networks Using the Eigenvectors of Matrices*. *Physical Review*, E 74, 036104.

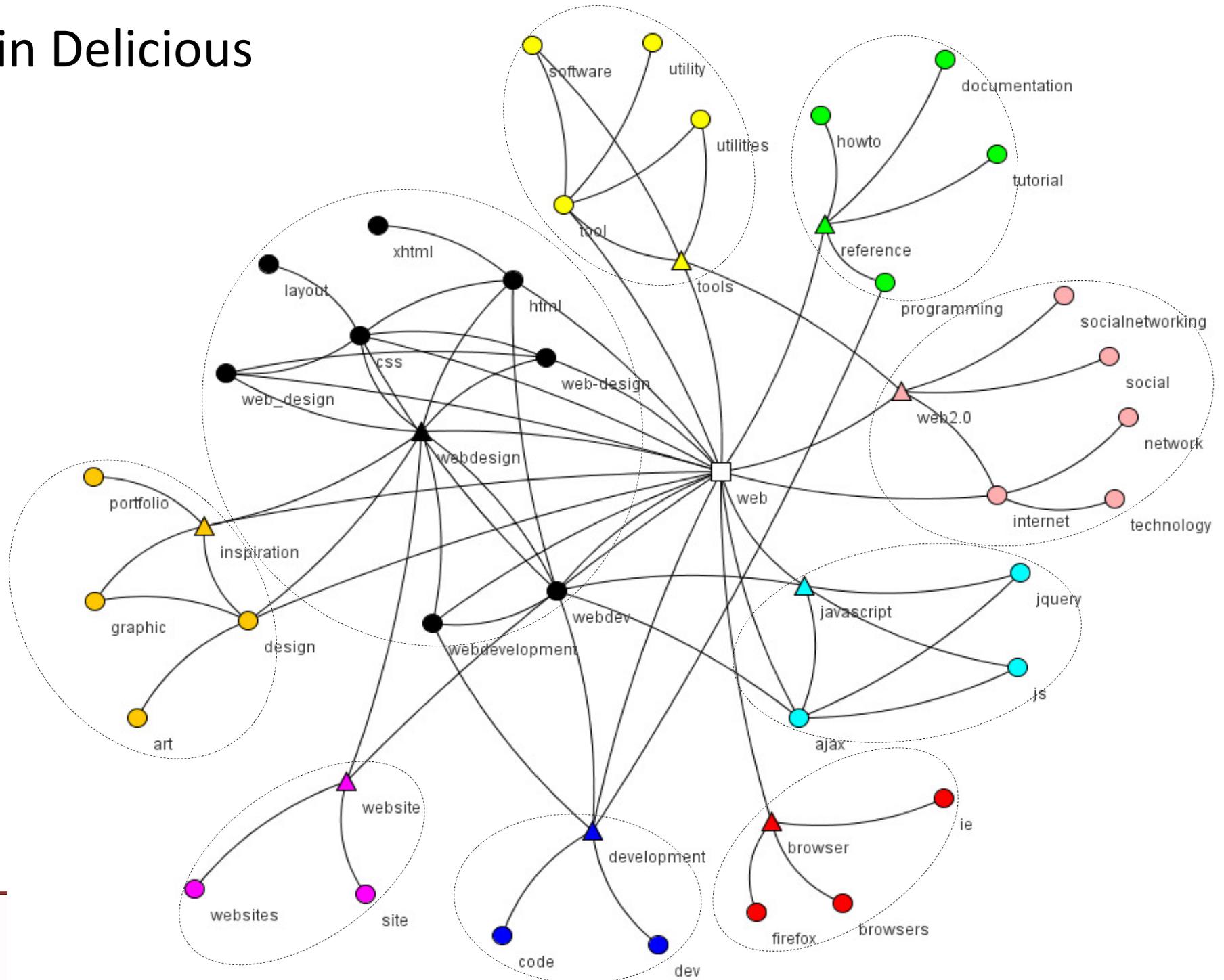
Contextualization and disambiguation of tags 84

- **sf** in Delicious

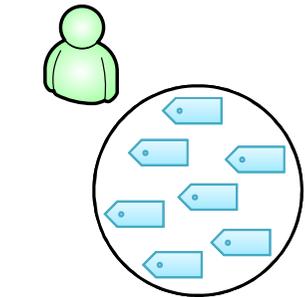


Contextualization and disambiguation of tags 85

- **web** in Delicious



- Profile models



Basic

tag	weight
java	0.29
j2ee	0.17
web	0.42
tag	0.12

Single-context

tag	tag context	weight
java	programming	0.29
j2ee	java	0.17
web	socialweb	0.42
tag	socialweb	0.12

Multiple-context

tag	tag context	weight
java	programming	0.20
java	tools	0.14
j2ee	java	0.13
web	socialweb	0.23
web	webdev	0.21
tag	socialweb	0.08

Extended single-context

tag	weight
java	0.27
programming	0.05
j2ee	0.15
web	0.34
socialweb	0.09
tag	0.10

Extended multiple-context

tag	weight
java	0.30
programming	0.03
tools	0.02
j2ee	0.11
web	0.37
socialweb	0.05
webdev	0.04
tag	0.07

Tag: java

Contextualized tag: java | programming

Cantador, I., Bellogín, B., Vallet, D. 2010. Content-based Recommendation in Social Tagging Systems. In Proceedings of the 4th ACM Conference on Recommender Systems (RecSys'10), 237-240.

Contextualization and disambiguation of tags ⁸⁷

- Still evaluating the recommendation models by using semantically contextualized tag-based profiles...
- Some (preliminary) results with a Delicious dataset

	P@5	P@10	P@20	MAP	R@5	R@10	R@20	MRR	NDCG
<i>tf</i>	0.073	0.056	0.041	0.023	0.024	0.036	0.054	0.182	0.061
<i>tfidf</i>	0.135	0.103	0.074	0.044	0.044	0.067	0.096	0.330	0.113
<i>bm25</i>	0.149	0.109	0.077	0.048	0.048	0.071	0.100	0.357	0.121
<i>tf-single</i>	0.093	0.069	0.049	0.029	0.030	0.045	0.064	0.239	0.077
<i>tfidf-single</i>	0.162	0.117	0.083	0.052	0.053	0.076	0.107	0.388	0.131
<i>bm25-single</i>	0.171	0.123	0.085	0.069	0.055	0.080	0.109	0.403	0.136

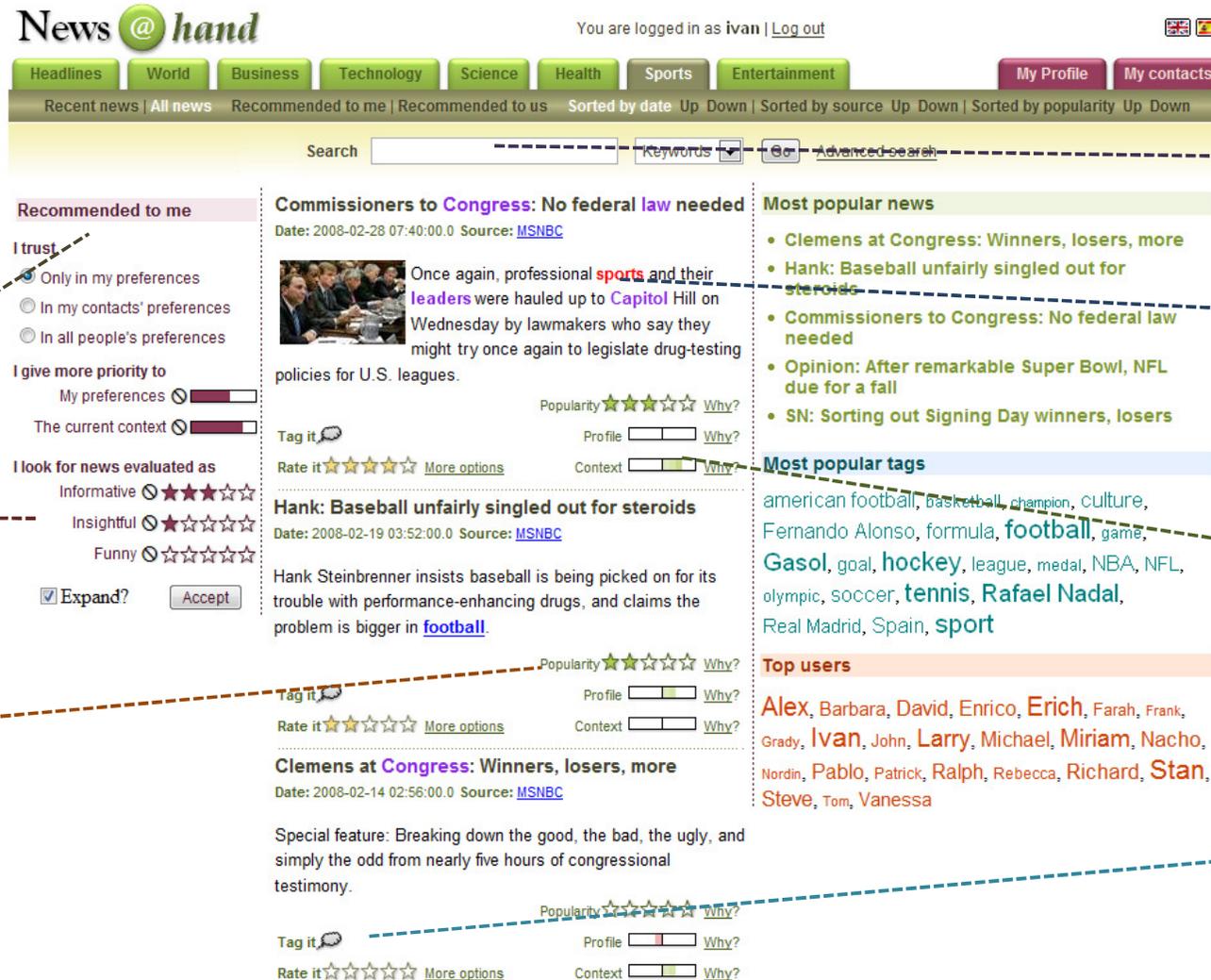
<i>tf-single</i>	27.20%	23.18%	18.54%	23.77%	28.40%	23.98%	19.25%	31.24%	24.81%
<i>tfidf-single</i>	19.68%	14.49%	12.15%	18.07%	19.37%	14.18%	11.62%	17.31%	18.07%
<i>bm25-single</i>	15.25%	13.09%	9.85%	16.97%	15.09%	12.57%	9.13%	12.96%	12.64%

Improvements!

- The Social Semantic Web
 - Web 2.0: The Social Web
 - Web 3.0: The Semantic Web
- **Ontology-based User Modeling**
 - Ontology-based User Preferences
 - Mapping Social Tags to Ontology Concepts
 - **Mapping Facebook Likes to DBpedia Entities**
- Ontology-based Recommendation
 - An Example of Ontology-based Recommender System
 - A Linked Data-based Cross-domain Recommendation Approach
- Conclusions

- The Social Semantic Web
 - Web 2.0: The Social Web
 - Web 3.0: The Semantic Web
- Ontology-based User Modeling
 - Ontology-based User Preferences
 - Mapping Social Tags to Ontology Concepts
 - Mapping Facebook Likes to DBpedia Entities
- Ontology-based Recommendation
 - **An Example of Ontology-based Recommender System**
 - A Linked Data-based Cross-domain Recommendation Approach
- Conclusions

- Recommendation of news items



Personalised and group-oriented content retrieval

Multi-criteria evaluations

Collaborative content-based recommendations

Semantic search

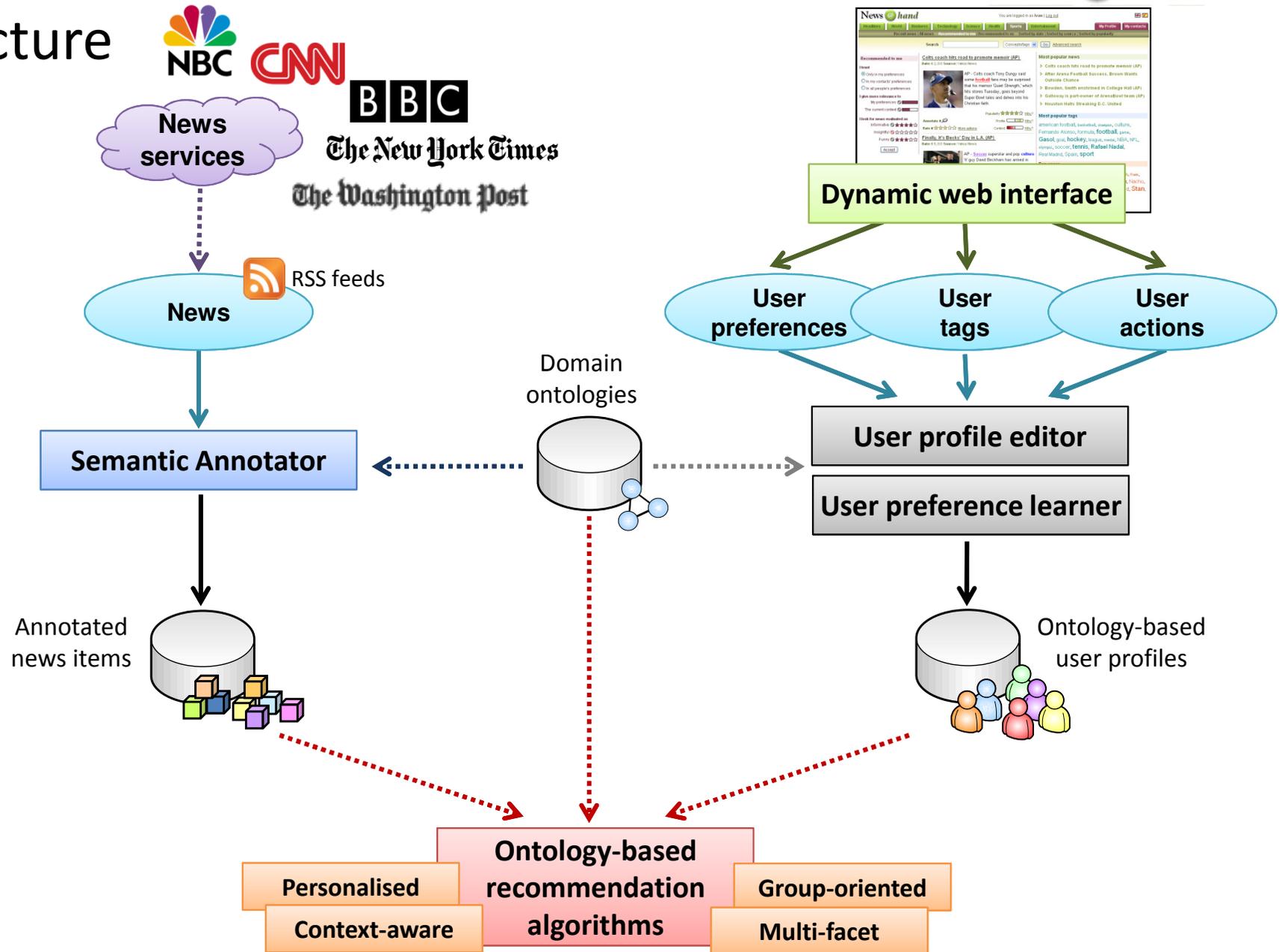
Automatic semantic annotation

Long-term (profile) and short-term (context) user preferences

Tags, comments and ratings

Cantador, I., Bellogín, A., Castells, P. 2008. *News@hand: A Semantic Web Approach to Recommending News*. In Proc. of the 5th Intl. Conference on Adaptive Hypermedia and Adaptive Web-Based Systems (AH 2008), 279-283.

- Architecture



- User profile editor

News@hand

You are logged in as ivan | [Log out](#)

Headlines World Business Technology Science Health Sports Entertainment

My Profile My contacts

Personal data My preferences My ratings My tags

Choose an interest situation At home

My preferences

Concept	Weight	Public for		
		All people	My contacts	Only me
soccer	<input type="range"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
tennis	<input type="range"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
basketball	<input type="range"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
baseball	<input type="range"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
earthquake	<input type="range"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>

hurricane

Insert a defined concept or select it from the categories

Vocabulary about disaster and accident

- disaster and accident
 - incidents
 - accidents
 - disaster
 - natural disasters
 - flood
 - earthquake
 - drought
 - cyclone
 - hurricane

h

- hurricane dennis
- hurricane ivan
- hurricane katrina
- hurricane rita
- hurricane stan
- hurricane wilma
- hurricane katrina
- hurricane rita
- hurricane stan
- hurricane wilma

This system has been developed by Iván Cantador and Alejandro Bellogín (Networked Semantics Team, <http://nets.ii.uam.es>)

Semantic user preferences

Ontology class browser

Ontology instance viewer

- **Problem 1:** the ontologies did not have **instances**

- named entities: people, places, organizations, etc.

- e.g. `teide`

- Teide should be an instance of the class **Volcano** (subclass of **Mountain**)*

- **Problem 2:** the news items did not have semantic **annotations**

- **Problem 3:** the users had to manually search for and select ontology concepts for building their **profiles**



- **Research questions**

- **RQ1:** How to automatically **populate ontologies** and **analyze news items**?
- **RQ2:** How to facilitate users to build their **ontology-based profiles** without asking them to search for concepts of interests?

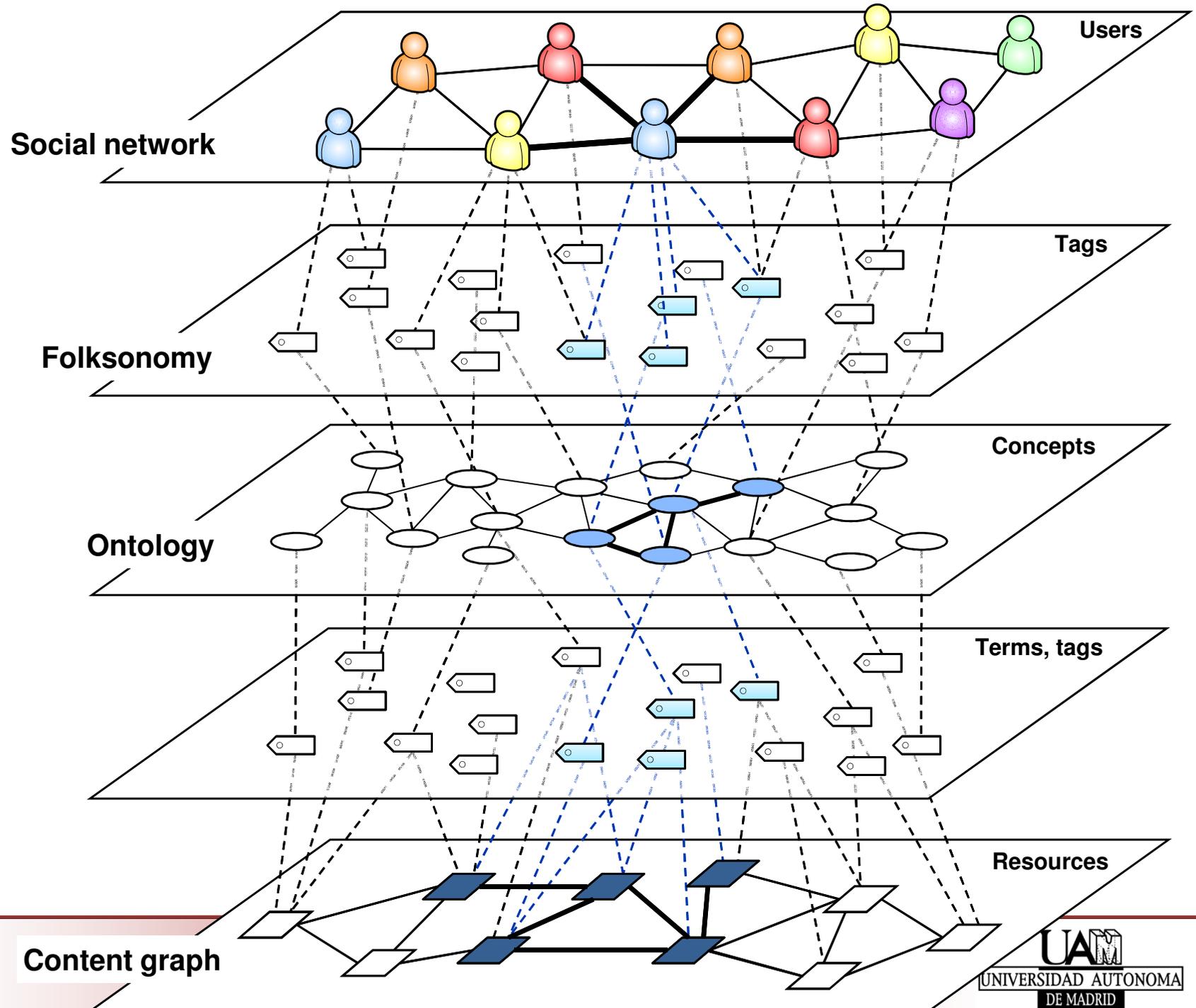
- **Approach**

- Mapping terms and social tags to ontology concepts

*Cantador, I., Castells, P., Bellogín, A. 2011. **An Enhanced Semantic Layer for Hybrid Recommender Systems: Application to News Recommendation**. International Journal on Semantic Web and Information Systems 7(1), pp. 44-77.*

*Fernández, M., Cantador, I., López, V., Vallet, D., Castells, P., Motta, E. 2011. **Semantically Enhanced Information Retrieval: An Ontology-based Approach**. Journal of Web Semantics 9(4), pp. 434-452.*

- Proposal



- Mapping a term or tag to a **Wikipedia entry** (+/- dealing with ambiguity)
 - `teide` → <http://en.wikipedia.org/wiki/Teide>
- Information extraction and processing of **Wikipedia categories**
 - `teide` → volcanoes, mountains, spain, ...
- Mapping of the term categories to **ontology classes**
 - `teide` → instance of “volcano” class

Teide

From Wikipedia, the free encyclopedia

Mount Teide or, in Spanish, **El Teide**, is an active though dormant **volcano** which last erupted in 1909 from the El Chinyero vent on the Santiago (northwestern) rift and is located on **Tenerife, Canary Islands**. The volcano and its surrounds comprise the **Teide National Park (Parque Nacional del Teide** in Spanish). The park has an area of 18900 ha and was named a **World Heritage Site** by **UNESCO**^[1] on **June 29, 2007**.

At 3718 m above sea level, and approximately 7500 m above the floor of the Atlantic Ocean, Teide is the highest mountain in **Spain** and the highest point in the **Atlantic Ocean**.^[2] (Note: The actual summit stands 3 metres (10 ft) higher than the triangulation station, and associated bench mark, which has an altitude of 3,715 m (12,188 ft)). The island of Tenerife itself is the third largest volcanic ocean island on Earth by volume. Teide is also the third highest volcano on a volcanic ocean island.^[3] It is also unstable and possibly in a more advanced stage of deformation and failure than the much publicised **Cumbre Vieja**.^[4] The United Nations Committee for Disaster Mitigation have designated Teide as a **Decade Volcano**. It is considered to be the 13th most dangerous volcano in the world due to its proximity to several major towns and the nearby city of Puerto de la Cruz.

Teide together with its neighbour **Pico Viejo** and **Montaña Blanca** forms the Central Volcanic Complex.



Categories: [World Heritage Sites in Spain](#) | [Tenerife](#) | [Mountains of Spain](#) | [Volcanoes of Spain](#) | [National parks of Spain](#) | [Stratovolcanoes](#) | [Hotspot volcanoes](#) | [Decade Volcanoes](#)

• Ontology population

- 17 ontologies (built from IPTC taxonomy)
- 744 classes; 121,135 instances
- 69.9% / 84.4% of accuracy (20 people evaluating 10,200 instances)

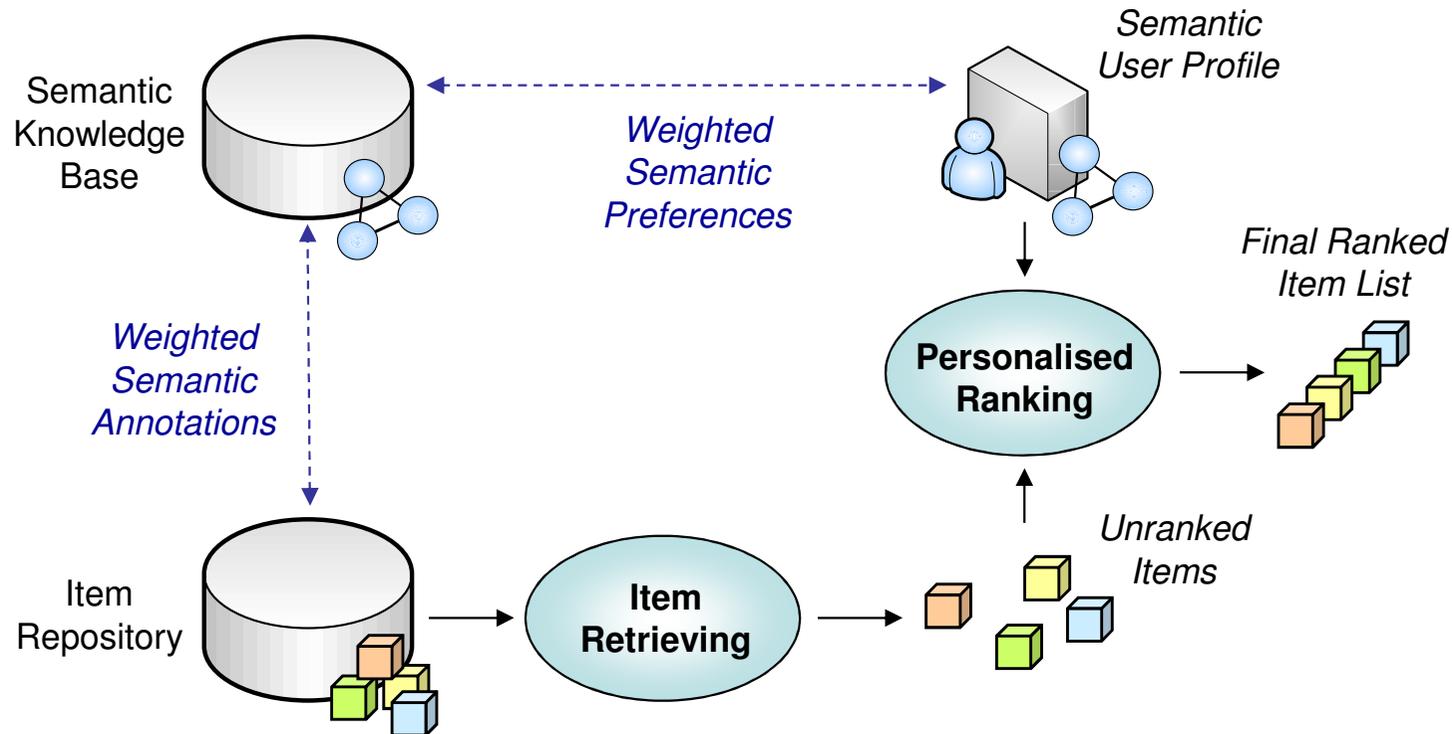
Ontology	Attributes			Average population accuracies	
	#classes	#instances	Avg. #instances/class	Class instantiation	Ontology instantiation
<i>Arts, culture, entertainment</i>	87	33,278	383	78.7	93.3
<i>Crime, law, justice</i>	22	971	44	62.7	73.3
<i>Disasters, accidents</i>	16	287	18	74.7	84.0
<i>Economy, business, finance</i>	161	25,345	157	69.3	80.0
<i>Education</i>	20	3,542	177	57.5	76.7
<i>Environmental issues</i>	41	20,581	502	72.0	85.3
<i>Health</i>	26	1,078	41	65.3	89.3
<i>Human interests</i>	6	576	96	64.0	84.0
<i>Labour</i>	6	133	22	70.7	78.7
<i>Lifestyle, leisure</i>	29	4,895	169	72.0	90.7
<i>Politics</i>	54	3,206	59	60.0	81.3
<i>Religion, belief</i>	31	3,248	105	84.0	90.7
<i>Science, technology</i>	50	7,869	157	68.0	86.7
<i>Social issues</i>	39	8,673	222	70.7	85.3
<i>Sports</i>	124	5,567	45	72.0	86.7
<i>Unrests, conflicts, wars</i>	23	1,820	79	61.3	80.0
<i>Weather</i>	9	66	7	69.7	89.5
	744	121,135	163 (avg.)	69.9	84.4

- **Semantic annotation of news items**

- 9,698 news items (gathered during two months)
- 66,378 generated annotations
- 7 annotations per item (average)
- 74.8% of accuracy (20 people evaluating 1,600 items)

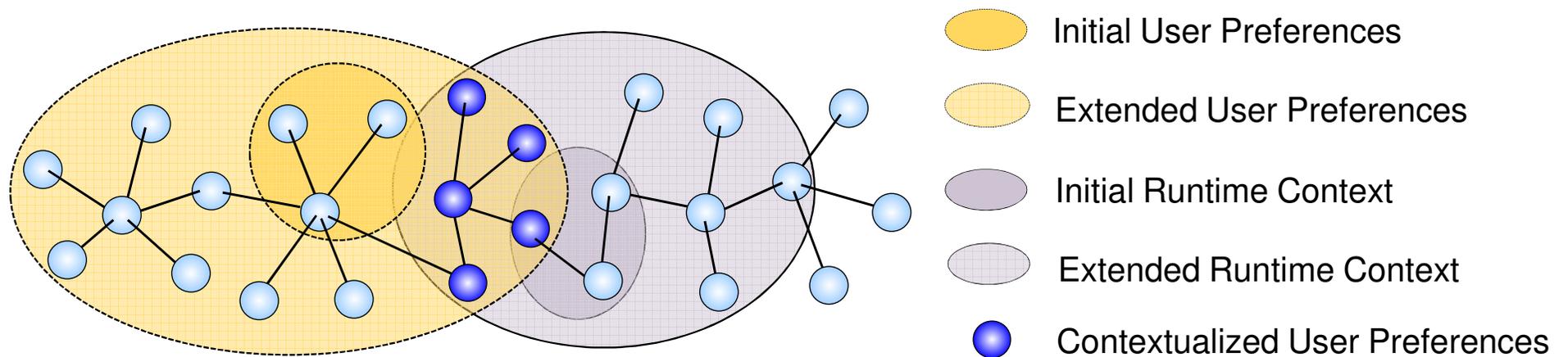
News section	Retrieved/generated data			
	#news items	#annotations	Avg. #annotations/item	Avg. accuracy
<i>Headlines</i>	2,660	18,210	7	71.4
<i>World</i>	2,200	17,767	8	72.7
<i>Business</i>	1,739	13,090	8	79.2
<i>Technology</i>	303	2,154	7	76.3
<i>Science</i>	346	2,487	7	74.1
<i>Health</i>	803	4,874	6	73.1
<i>Sports</i>	603	2,453	4	75.8
<i>Entertainment</i>	1,044	5,343	5	76.0
	9,698	66,378	7	74.8

- (Some) Recommendation algorithms
 - Content-based recommender
 - Context-aware recommender
 - Multi-layer hybrid recommender



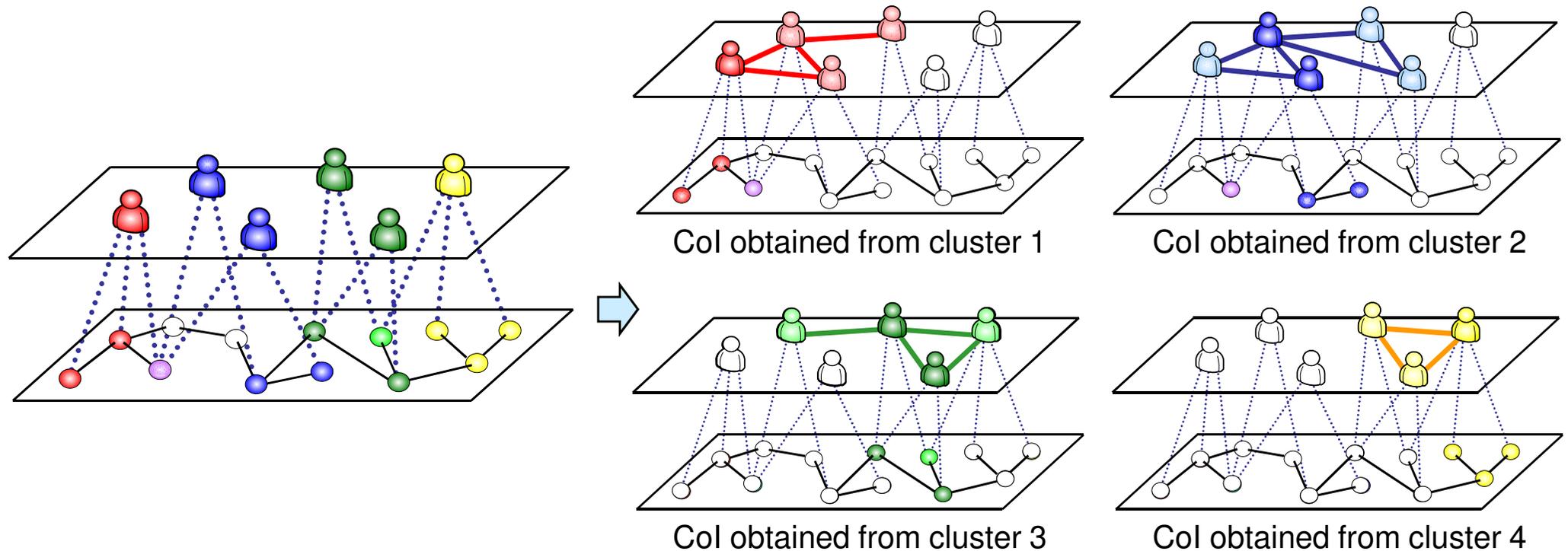
Cantador, I., Castells, P., Bellogín, A. 2011. *An Enhanced Semantic Layer for Hybrid Recommender Systems: Application to News Recommendation*. *International Journal on Semantic Web and Information Systems* 7(1), pp. 44-77.

- (Some) Recommendation algorithms
 - Content-based recommender
 - Context-aware recommender
 - Multi-layer hybrid recommender



Cantador, I., Castells, P., Bellogín, A. 2011. **An Enhanced Semantic Layer for Hybrid Recommender Systems: Application to News Recommendation**. *International Journal on Semantic Web and Information Systems* 7(1), pp. 44-77.

- (Some) Recommendation algorithms
 - Content-based recommender
 - Context-aware recommender
 - Multi-layer hybrid recommender



Cantador, I., Castells, P., Bellogín, A. 2011. *An Enhanced Semantic Layer for Hybrid Recommender Systems: Application to News Recommendation*. *International Journal on Semantic Web and Information Systems* 7(1), pp. 44-77.

- (Many) Recommendation results...
- A summary:
 - Good results
 - **Semantic expansion of ontology-based user preferences**
 - Combination of personalized and contextualized approaches
 - Hybrid approach in cold-start and high-sparsity situations
 - Open issues
 - **Improving disambiguation**
 - Addressing scalability
 - User preference learning (e.g. synonym) and recommendation

- The Social Semantic Web
 - Web 2.0: The Social Web
 - Web 3.0: The Semantic Web
- Ontology-based User Modeling
 - Ontology-based User Preferences
 - Mapping Social Tags to Ontology Concepts
 - Mapping Facebook Likes to DBpedia Entities
- **Ontology-based Recommendation**
 - An Example of Ontology-based Recommender System
 - **A Linked Data-based Cross-domain Recommendation Approach**
- Conclusions

- Inter-domain semantic relatedness

Vienna State Opera, Vienna, Austria
http://en.wikipedia.org/wiki/Vienna_State_Opera



The Vienna State Opera (Wiener Staatsoper) is an opera house - and opera company - with a history dating back to the mid-19th century. It is located in the centre of Vienna, Austria. It was originally called the Vienna Court Opera (Wiener Hofoper). In 1920, with the replacement of the members of the Vienna Philharmonic are

First Austrian Republic, it was renamed the members of the Vienna Philharmonic are

Matching places of interest with musicians Completed tasks (8 out of 10) ●●●●●●●●○○ Logout

Vienna State Opera



City: Vienna, Austria
Date: 1869

Architecture categories: Opera houses, Theatres

Description: The Vienna State Opera (Wiener Staatsoper) is an opera house – and opera company – with a history dating back to the mid-19th century. It is located in the centre of Vienna, Austria. It was originally called the Vienna Court Opera (Wiener Hofoper). In 1920, with the replacement of the Habsburg Monarchy by the First Austrian Republic, it was renamed the Vienna State Opera. The members of the Vienna Philharmonic are recruited from its orchestra. ...
[More]
http://en.wikipedia.org/wiki/Vienna_State_Opera

Arnold Schoenberg



Birth/origin city: Vienna, Austria
Death city: Los Angeles, USA
Birth/origin date: 1874
Death date: 1951

Music categories: 20th-century classical composers, American music, Ballet composers, Classical music, Jewish classical musicians, Modernist composers, Opera composers

Description: Arnold Schoenberg (13 September 1874 – 13 July 1951) was an Austrian and later American composer, associated with the expressionist movement in German poetry and art, and leader of the Second Viennese School. He used the spelling Schönberg until after his move to the United States in 1934 (Steinberg 1995, 463), whereupon he altered it to Schoenberg "in deference to American practice" (Foss 1951, 401), though one writer claims he made the change...
[More]
http://en.wikipedia.org/wiki/Arnold_Schoenberg

In your opinion, how related is Arnold Schoenberg to the place Vienna State Opera?

Very related Related Poorly related Not related

If you think they are related, justify your response by clicking in the boxes associated to the information you consider as relevant

	Very relevant	Relevant	Not relevant
Birth/origin place: Vienna, Austria	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
Death place: Los Angeles, USA	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
Birth/origin date: 1874	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
Death date: 1951	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
Music categories: 20th-century classical composers, American music, Ballet composers, Classical music, Jewish classical musicians, Modernist composers, Opera composers, Second Viennese school	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
Arnold Schoenberg is a Opera composers musician/band. Opera composers is related with Opera houses in Austria, which is an architecture category of Vienna State Opera.	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
Arnold Schoenberg is a Classical music musician/band. Classical music has a subcategory called Opera. Opera is related with Opera houses in Austria, which is an architecture category of Vienna State Opera.	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>

How interesting is the suggested match between Arnold Schoenberg and Vienna State Opera?

Very interesting Interesting Poorly interesting Not interesting

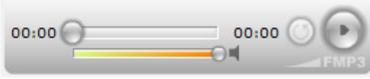
Is the suggested match between Arnold Schoenberg and Vienna State Opera obvious?

Yes No

Send responses

Select the tracks that in your opinion are the most relevant for the described location:

Jean-Baptiste Arban - Carnival of Venice
http://en.wikipedia.org/wiki/Jean-Baptiste_Arban



Alban Berg - Lyric Suite for String Quartet
http://en.wikipedia.org/wiki/Alban_Berg



Heinrich Proch - Das Alpenhorn op.18
http://en.wikipedia.org/wiki/Heinrich_Proch

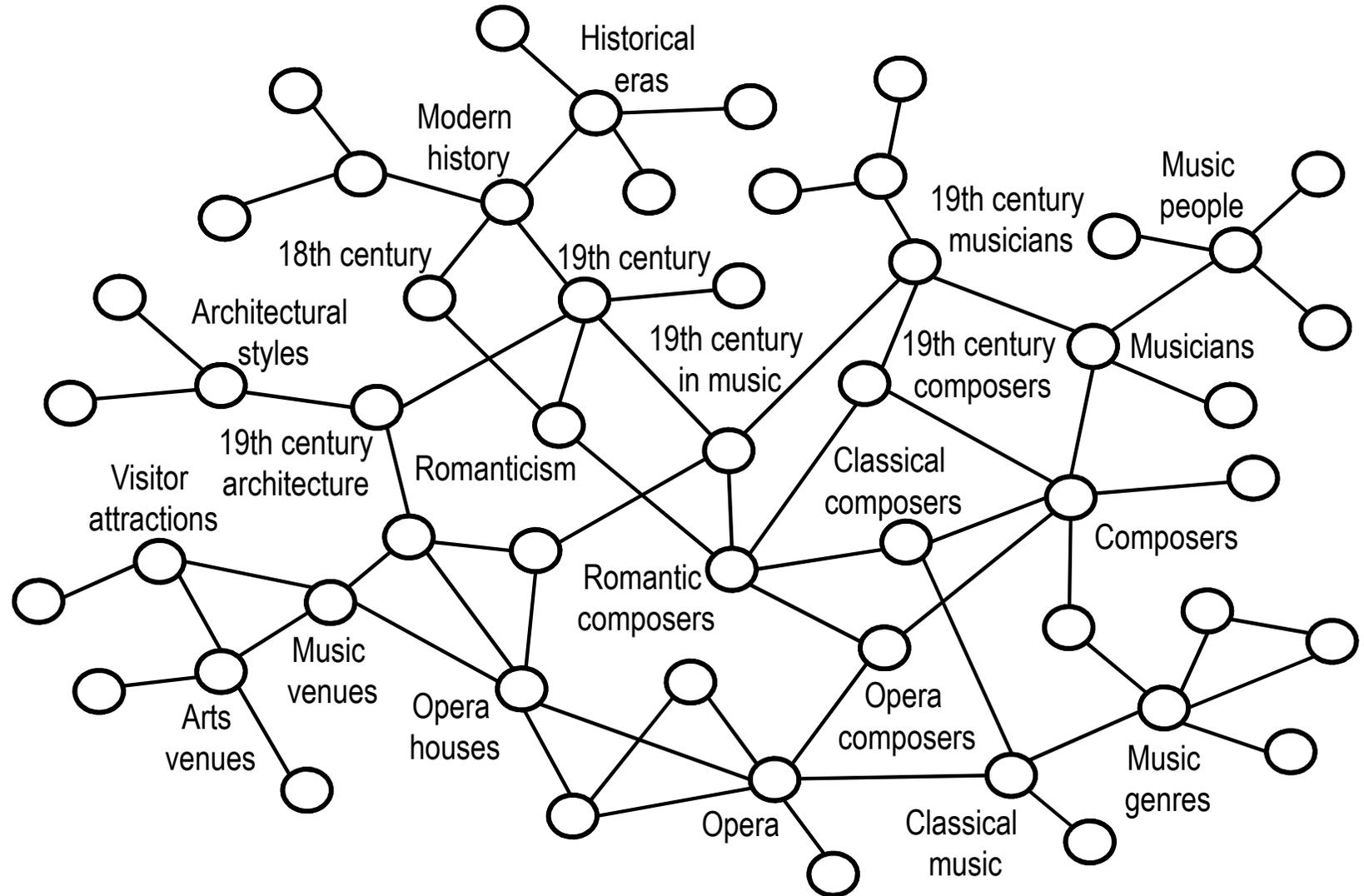


None of the above tracks goes well with the POI

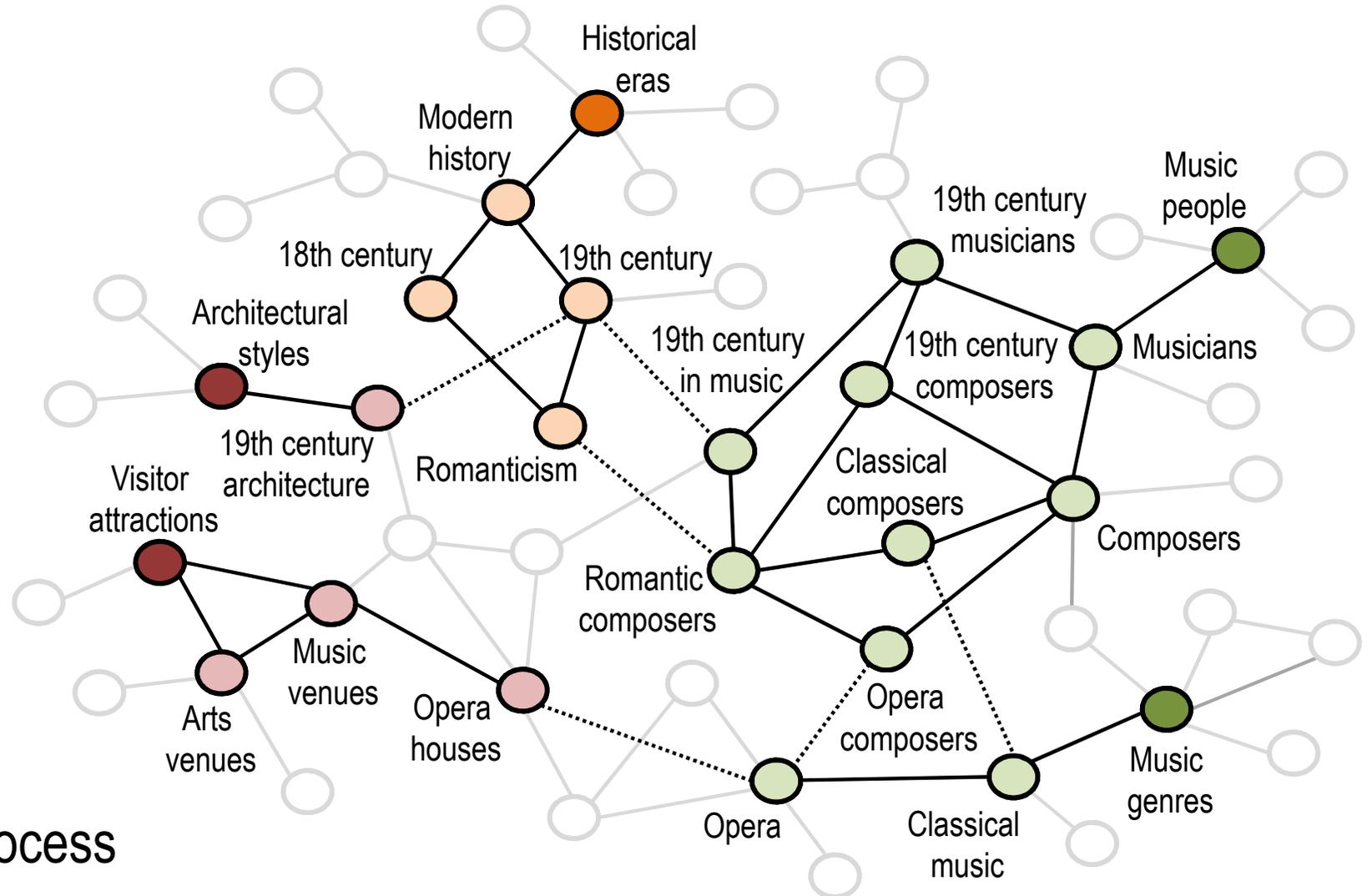
Submit

Kaminskas, M., Fernández-Tobías, I., Ricci, F., Cantador, I. 2014. *Knowledge-based Identification of Music Suited for Places of Interest*. *Journal of Information Technology and Tourism* 14(1), pp. 73-95.

- **Wikipedia category graph**



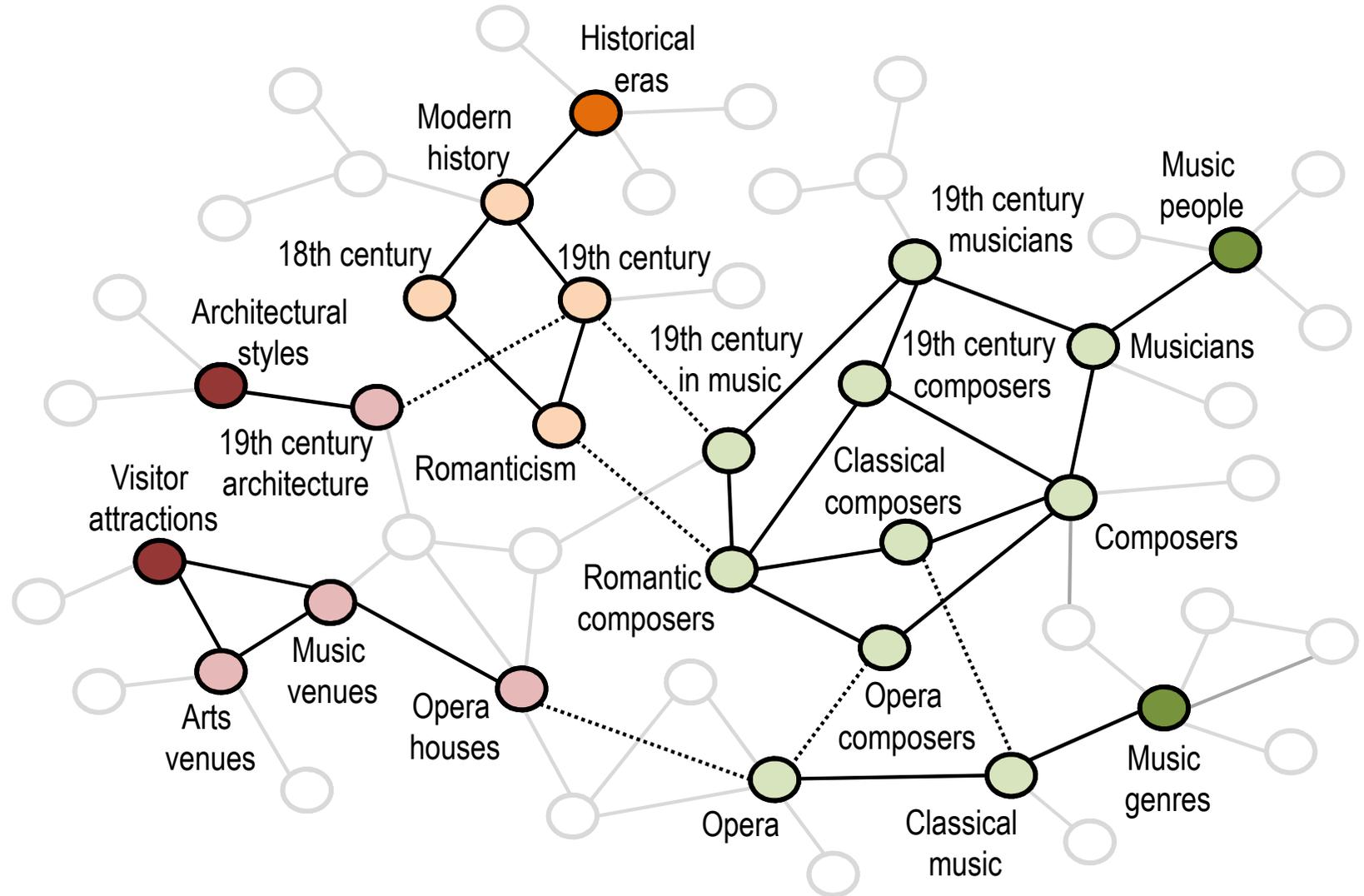
- Identifying domain-specific graphs



- Generic process

- Setting domain starting / **root concepts**
- Establishing **allowed and forbidden relations** between concepts for each domain
- Generating **sub-graphs** from starting concepts through allowed relations

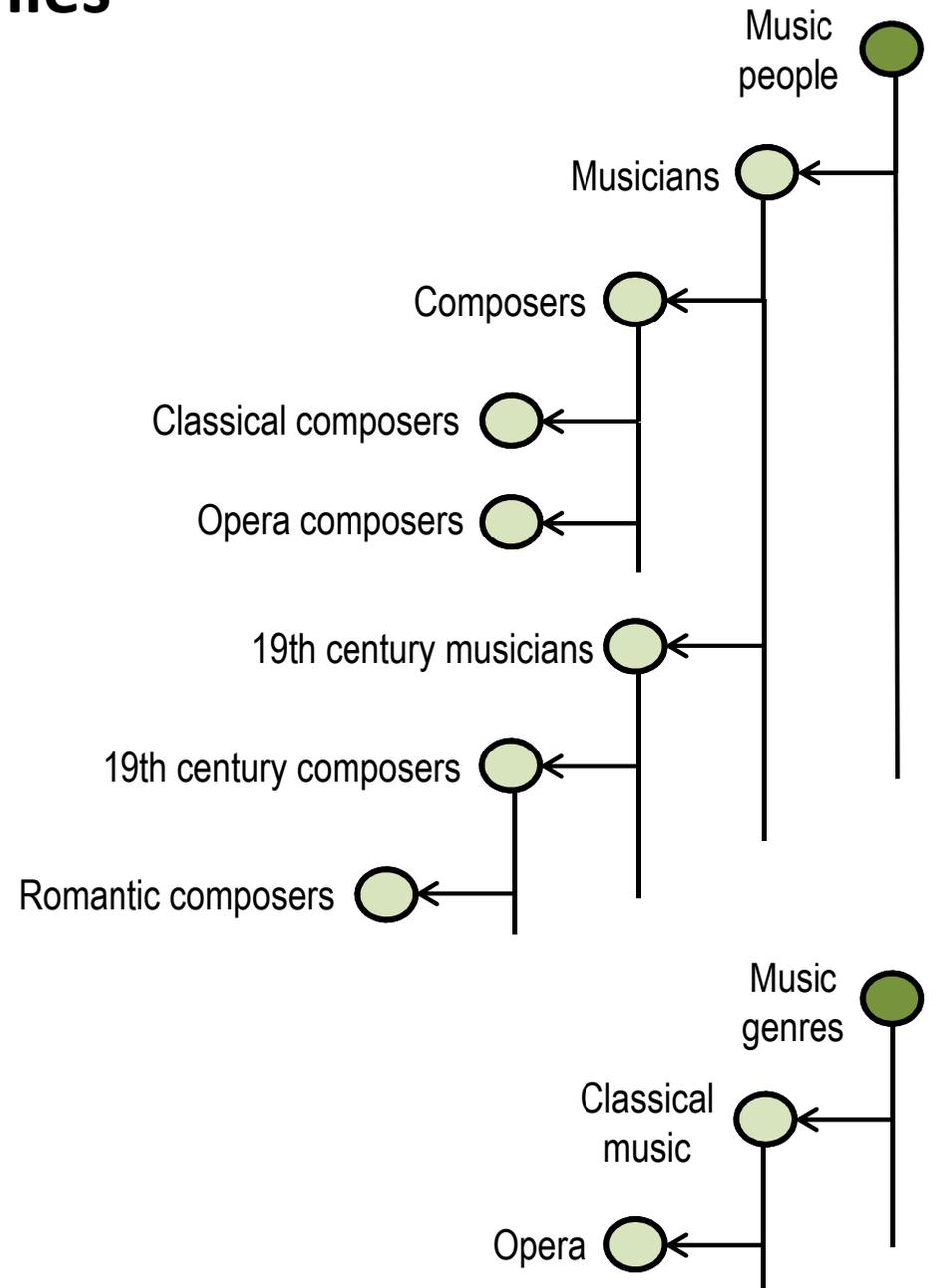
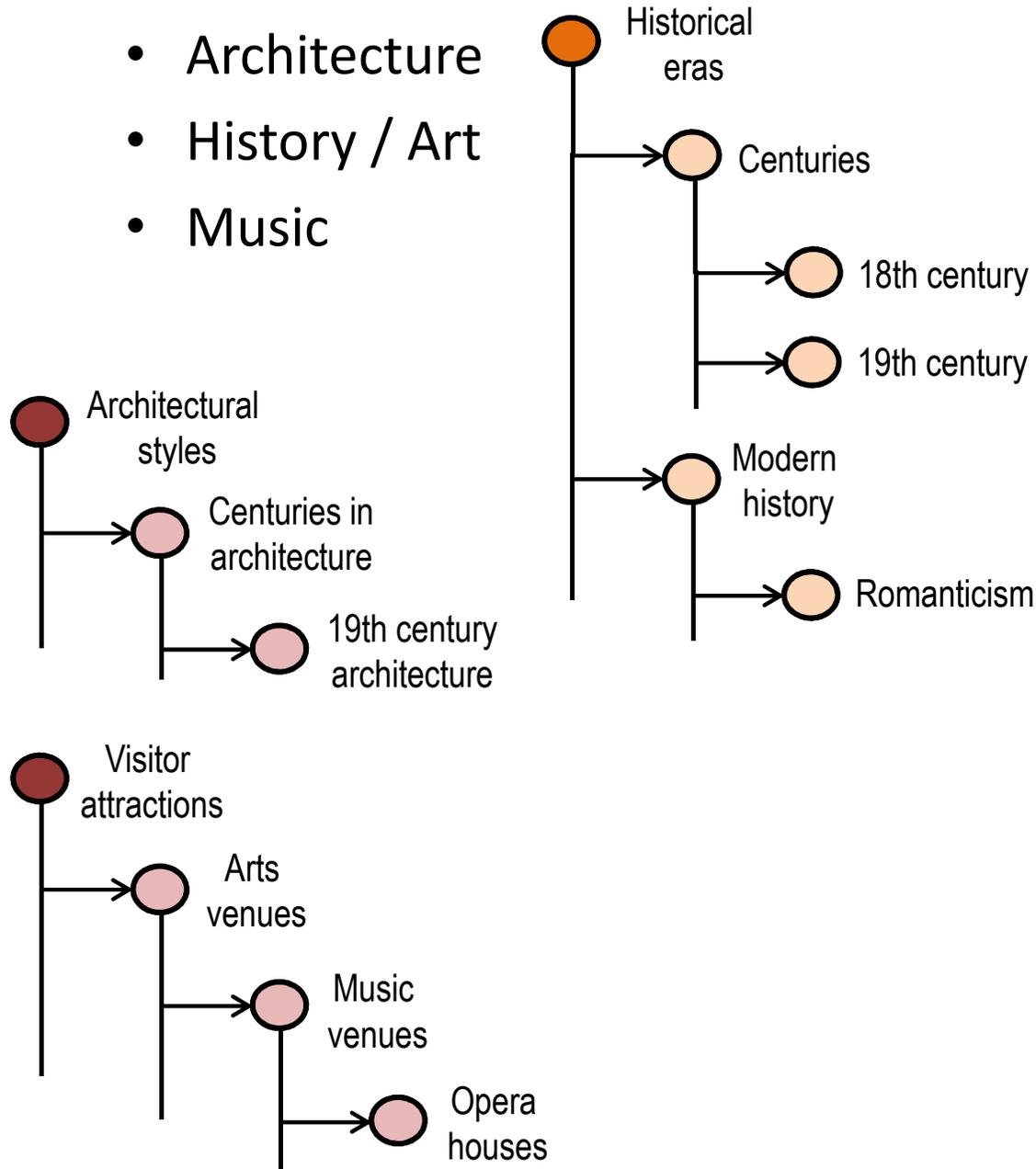
- Identifying domain-specific graphs



Fernández-Tobías, I., Kaminskas, M., Cantador, I., Ricci, F. 2011. A Generic Semantic-based Framework for Cross-domain Recommendation. In Proc. of the 2nd Intl. Workshop on Information Heterogeneity and Fusion in Recommender Systems (HetRec 2011).

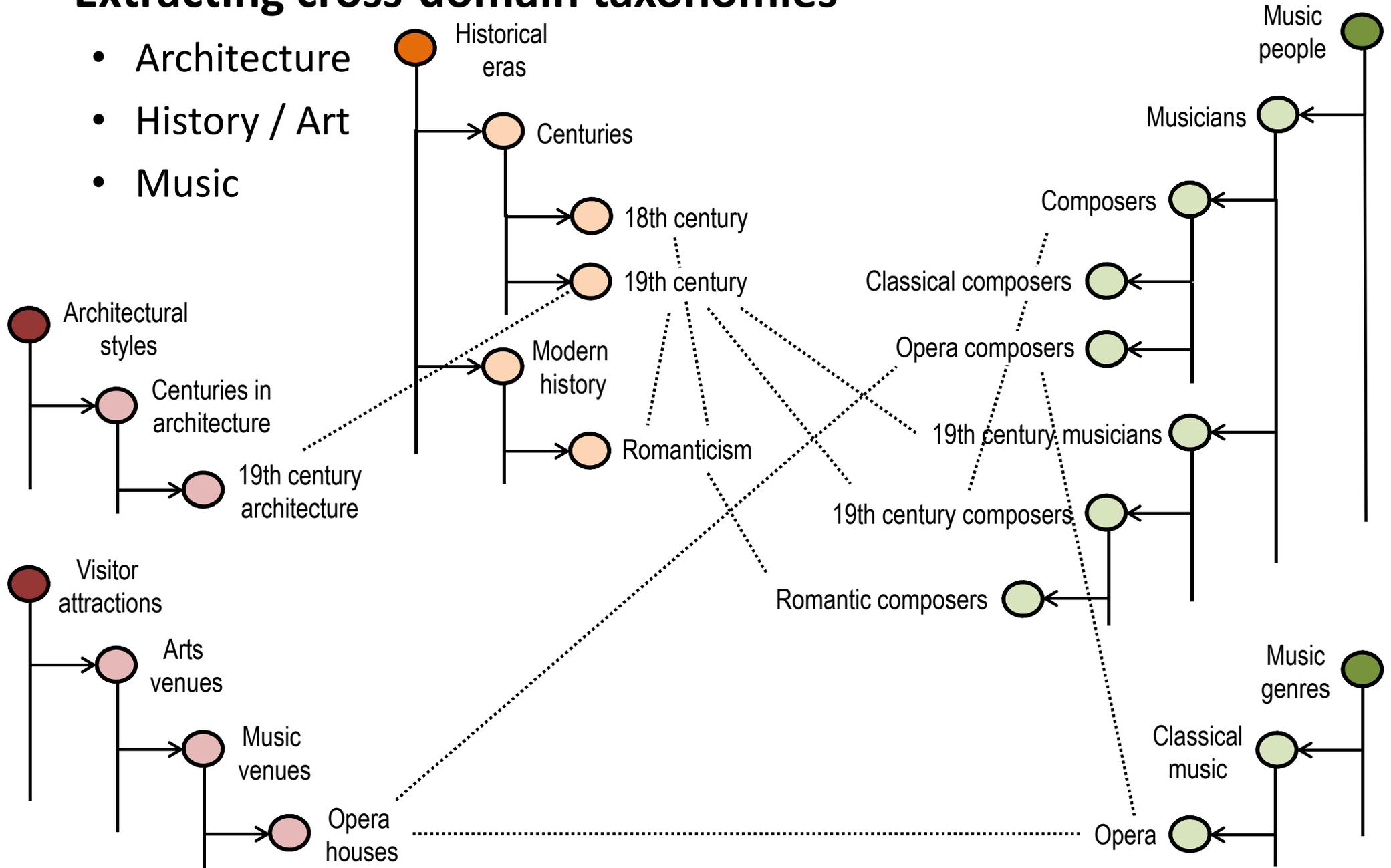
• Extracting cross-domain taxonomies

- Architecture
- History / Art
- Music

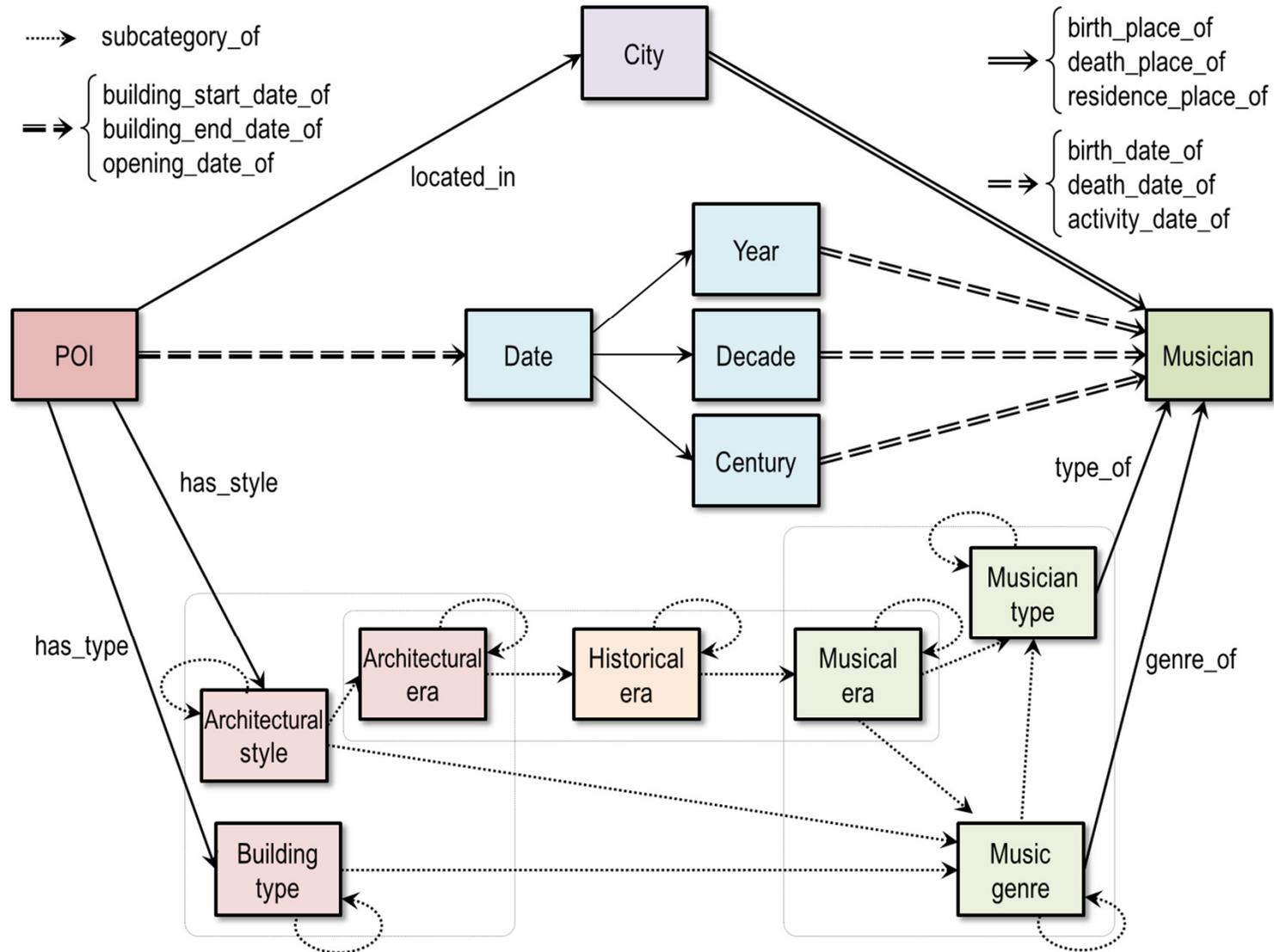


• Extracting cross-domain taxonomies

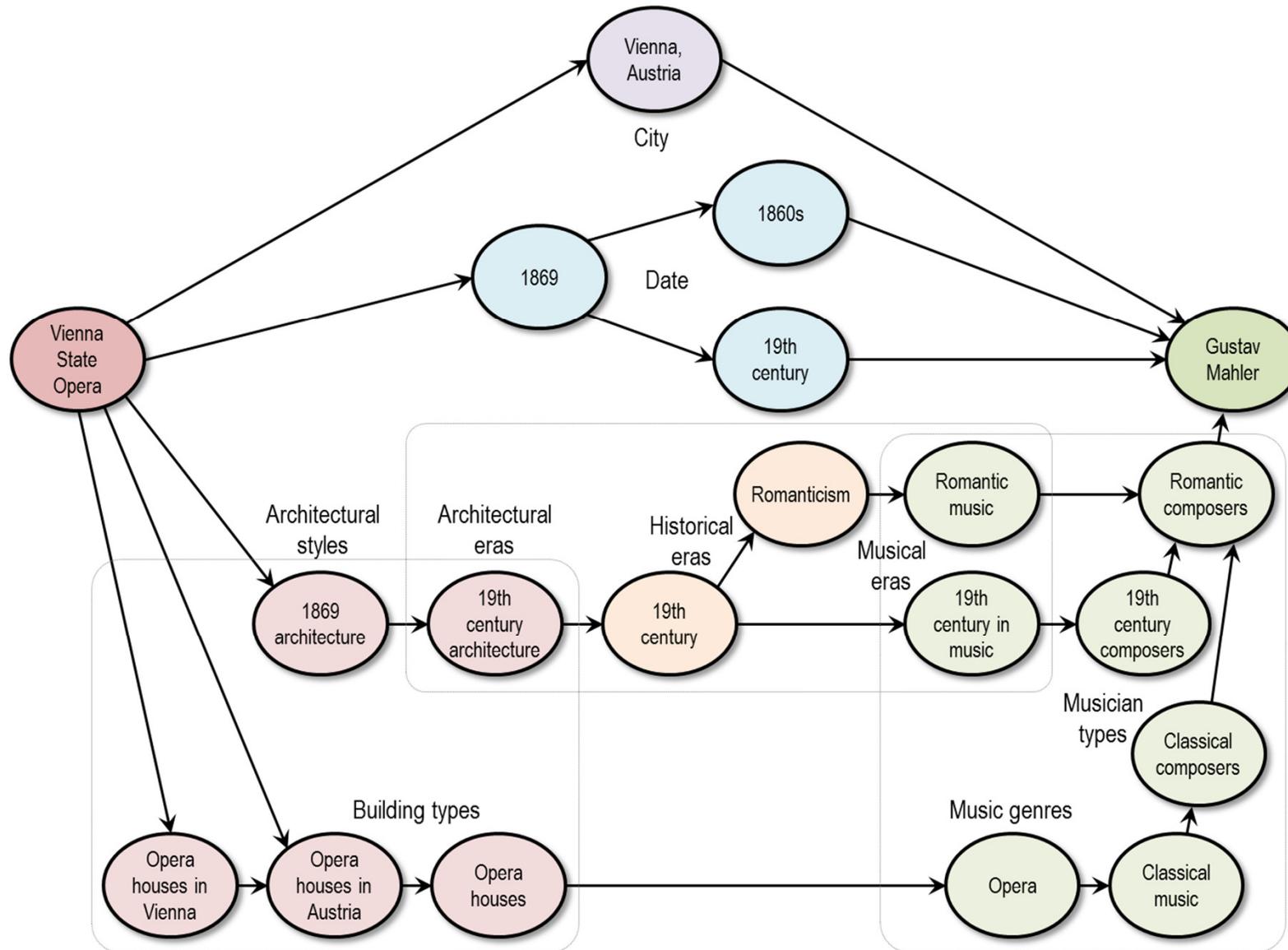
- Architecture
- History / Art
- Music



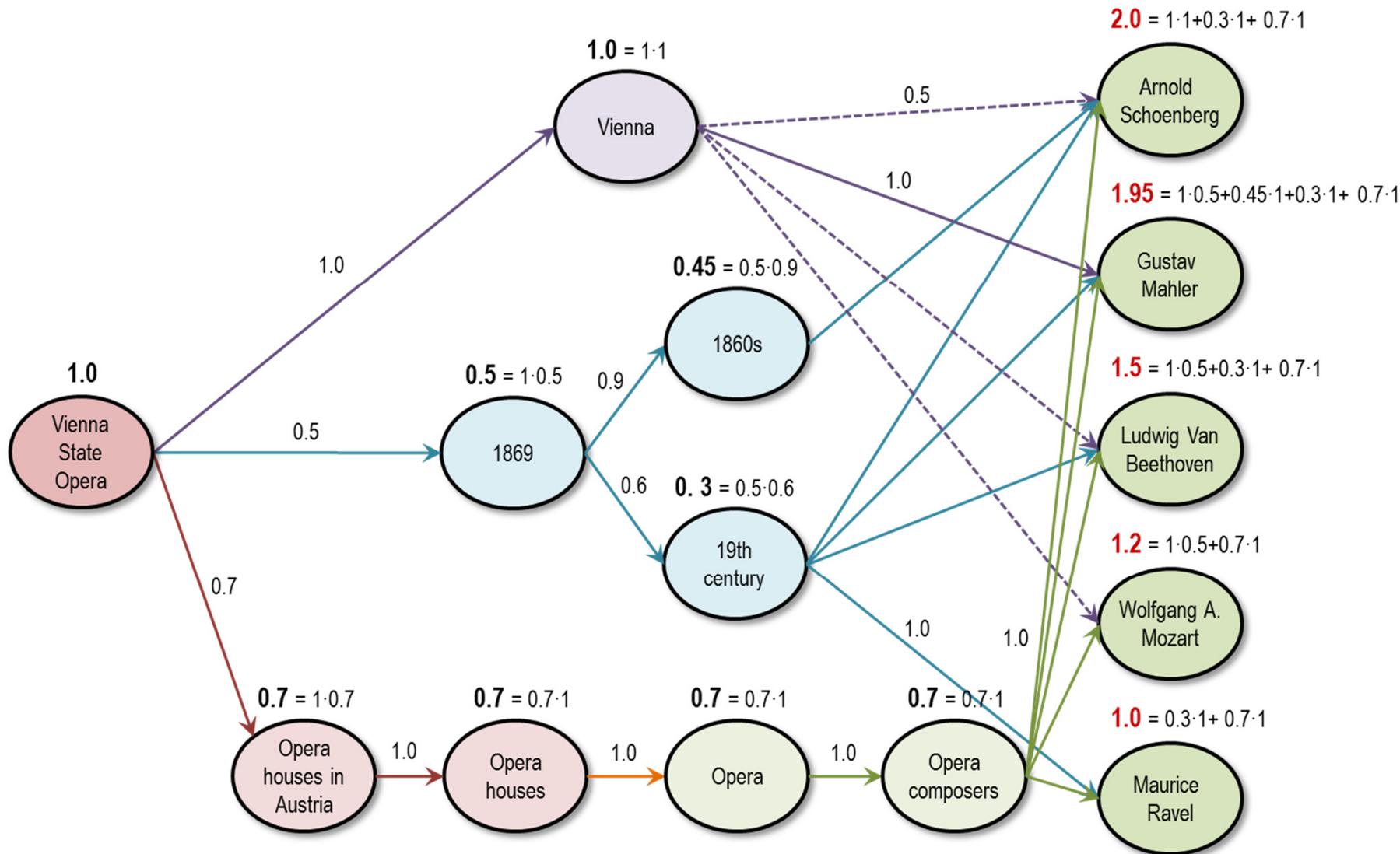
- Inter-domain Class Framework



- Inter-domain Instance Graph



- Graph-based Ranking Methods





Uhhmm! The idea seems interesting, but...

**Where are the instance-based
user preferences?**

- **ESWC Challenge on Linked Open Data-enabled Recommendation**

- 2014 edition, <http://challenges.2014.eswc-conferences.org/index.php/RecSys>
 - Organizers: Tommaso Di Noia, Iván Cantador, Vito Claudio Ostuni
 - **Dataset:** 75K LibraryThing ratings from 6K users for 6K books linked to DBpedia entities
 - **Related datasets:** MovieLens movie ratings and Last.fm listening records for music bands and musicians linked to DBpedia entities

<http://sisinflab.poliba.it/semanticweb/lod/recsys/datasets>

LibraryThing

movieLens

last.fm

facebook

- 2015 edition, <http://2015.eswc-conferences.org>
 - **Coming soon!!!**
 - **Organizers:** Tommaso Di Noia, Iván Cantador, Vito Claudio Ostuni, Matthew Rowe
 - **Dataset:** 5.5M Facebook likes from 193K users for 15K movie, music and book items linked to Dbpedia entities

- The Social Semantic Web
 - Web 2.0: The Social Web
 - Web 3.0: The Semantic Web
- Ontology-based User Modeling
 - Ontology-based User Preferences
 - Mapping Social Tags to Ontology Concepts
 - Mapping Facebook Likes to DBpedia Entities
- Ontology-based Recommendation
 - An Example of Ontology-based Recommender System
 - A Linked Data-based Cross-domain Recommendation Approach
- **Conclusions**

- Ontology- and semantic-based user modeling and recommendation approaches
 - have shown benefits, e.g., addressing cold-start and sparsity, and understanding the users' preferences
 - they are of (a renovated) interest : workshops and challenges, research topics in conferences, new datasets
 - have to address with still open problems:
 - entity recognition, disambiguation and linking
 - relation extraction
 - ontology population
 - semantic annotation

Ontology-based User Modeling and Recommendation

Iván Cantador

ivan.cantador@uam.es

Information Retrieval Group

Universidad Autónoma de Madrid, Spain