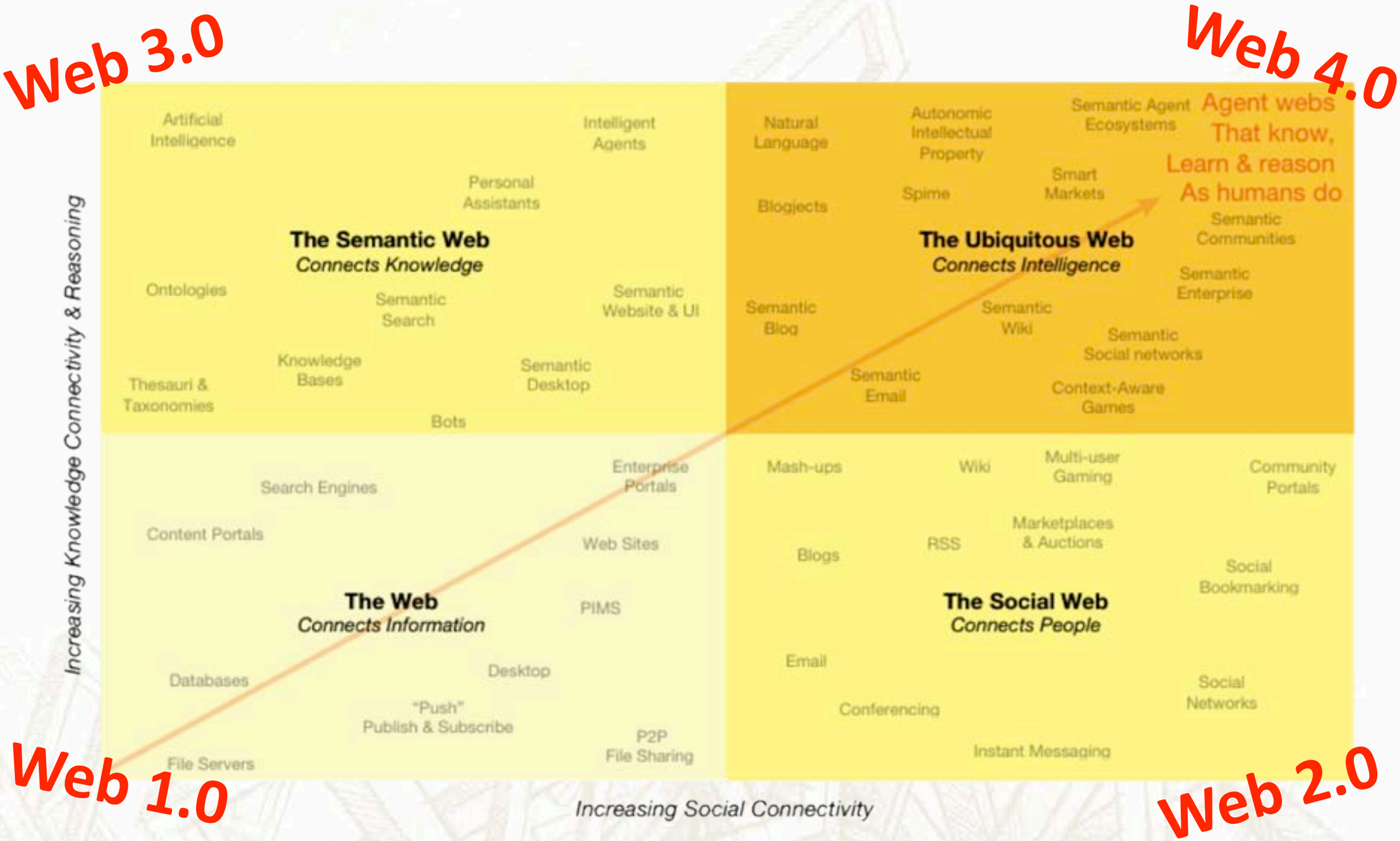
Some design patterns for the ontology-lexicon interface over the Web

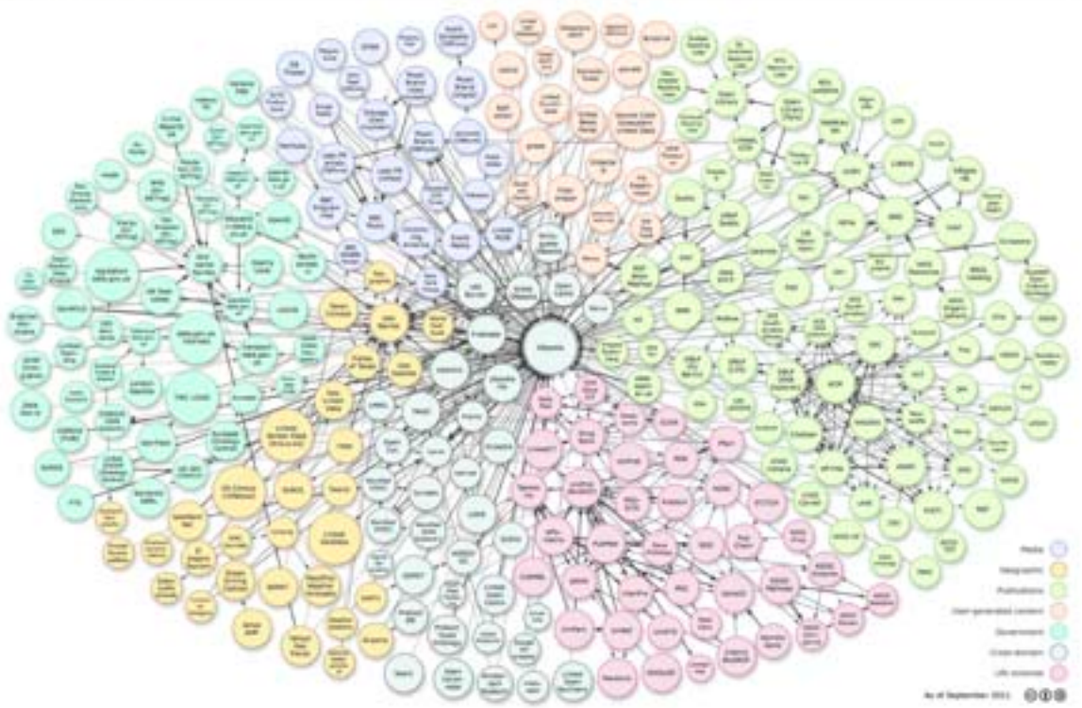
Aldo Gangemi
Semantic Technology Lab
ISTC-CNR, Roma
aldo.gangemi@cnr.it

Web evolution?



Courtesy of Nova Spivack

Overview on semantic technologies



DBpedia
Geodata
Bio_x
...
New York Times
Thomson Reuters
BBC
Amazon
data.gov
data.gov.uk...

- Semantic technologies make it emerge a lot of information from the Web and local information systems (*data silos*). ➢ *Also social networks like FaceBook are data silos ...*
- **Data** integration, through reengineering (e.g. *triplify*), or querying (e.g. *D2R*)
- Linking of heterogeneous data sources, either at schema or instance level
- Extraction of new data (*machine learning, NLP*) from **textual documents**, and their semantic representation
- Reasoning on those data, extracting more implicit information
- Presentation of data on the simplest platform: the Web, and so enabling sharing, collaborative editing, customization, etc. (URI-based data integration)

Linked data principles

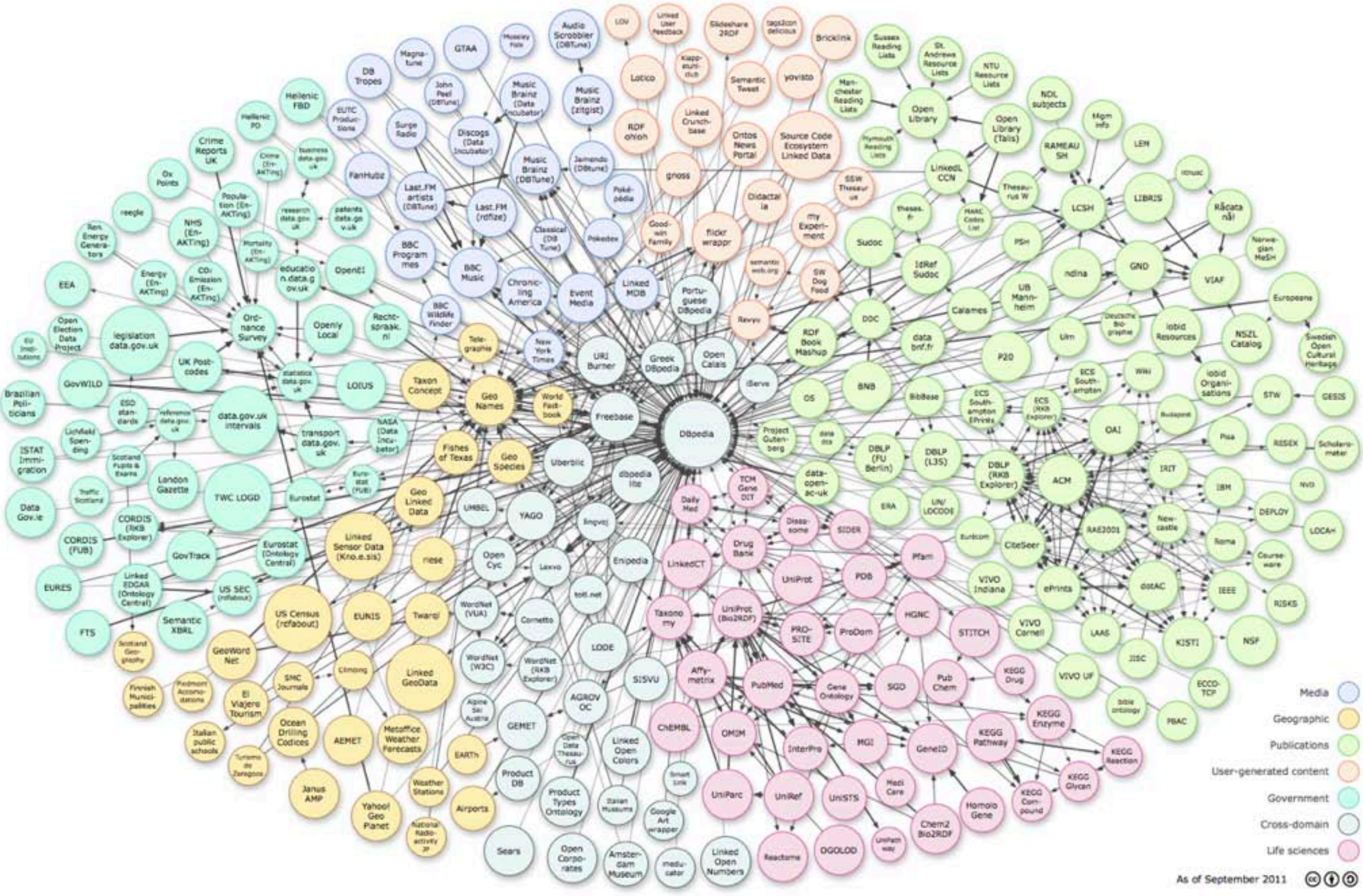
from tim berners-lee's linked data principles

- Use **URIs** as names for things
- Use **HTTP** URIs so that people can look up those names
  - Slash URIs are preferred
- When someone **looks up** a URI, provide useful information
  - Browsable graphs: a graph G is browsable if, for the URI of any node in G, if I look up that URI I will be returned information which describes the node, where describing a node means:
    - Returning all statements where the node is a subject or object
    - Describing all blank nodes attached to the node by one arc
- Include **links** to other URIs so that they can discover more things

Webbiness in the Semantic Web

opposed to design practices in lexical, NLP, or DB data

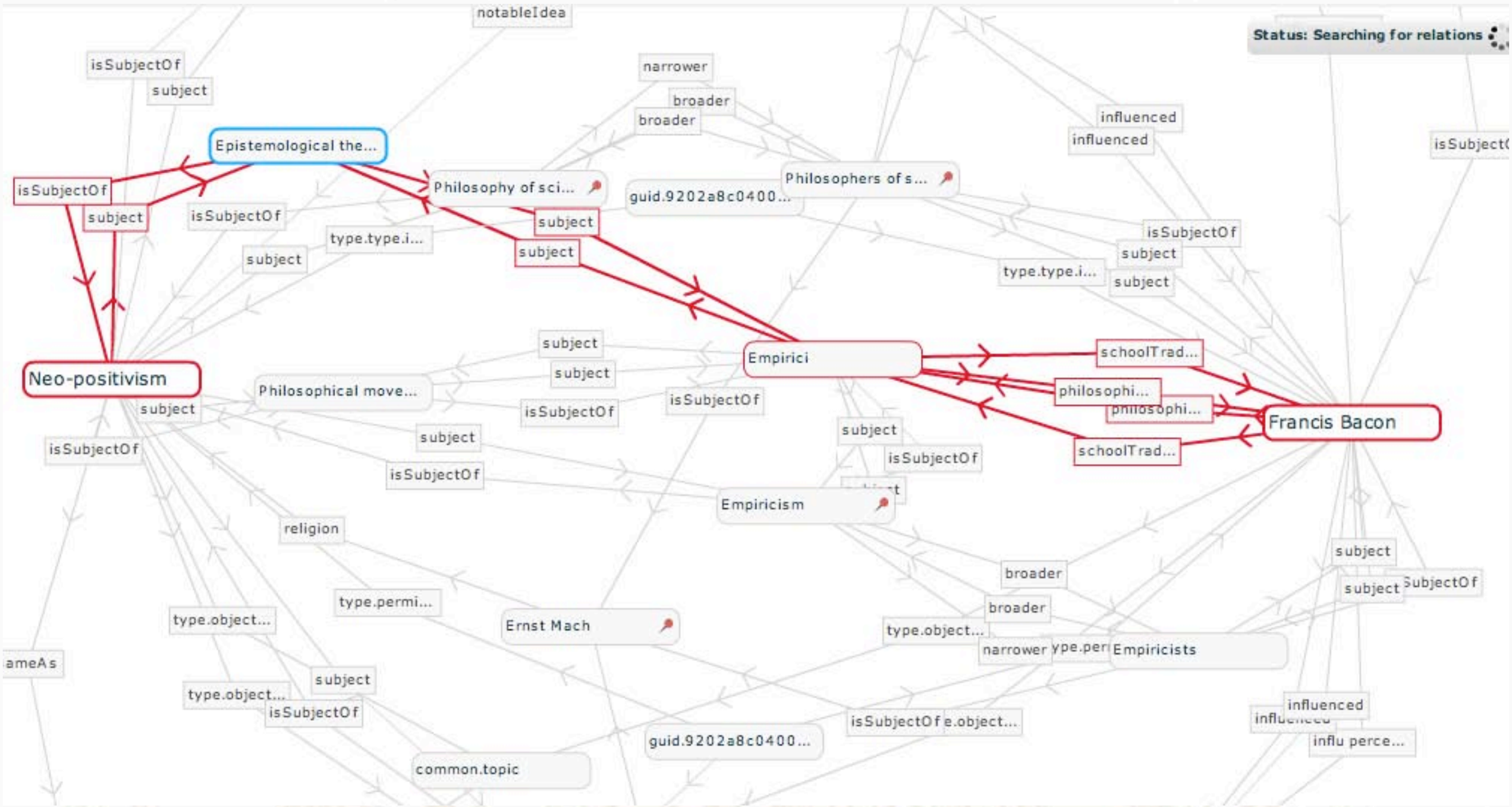
- Identity is based on locations (IRIs) and should be operationally addressed (dereferencing)
- Identity of web/non-web resources is a core problem
- Identity resolution performed through linking; loose usage of linking relations like *owl:sameAs* (logical identity)
- Open-world as default semantics: incomplete knowledge of individuals (domain), assertional (ABox), and terminological (TBox) axioms has to be assumed
- Ontologies are actually structured vocabularies, with minimal, incomplete, possibly incoherent or inconsistent axioms; linking and querying data are central, relations are central (rule-based favorite approach); hardly any anonymous concept construction, cf. *owl:Restriction*)
- Strong commitment on the object v. literal disjointness


DB Tropes
Magna-tune
GTAA
Moseley Folk
Audio Scrobbler (DBTune)
LOV
Linked User Feedback
Slideshare 2RDF
tags2con delicious
Bricklink
Sussex Reading Lists
St. Andrews Resource Lists
NTU Resource Lists
Hellenic FBD
Hellenic PD
EUTC Productions
John Peel (DBTune)
Surge Radio
Music Brainz (Data Incubator)
Music Brainz (zitgist)
Lotico
Klappstuhl-club
Semantic Tweet
yovisto
Manchester Reading Lists
Open Library
Open Library (Talis)
NDL subjects
Crime Reports UK
Discogs (Data Incubator)
RDF ohloh
Linked Crunchbase
Ontos News Portal
Source Code Ecosystem Linked Data
Plymouth Reading Lists
RAMEAU SH
Mgm info
LEM
Ox Points
Crime (En-AKTing)
business data.gov.uk
FanHubz
Last.FM artists (DBTune)
Music Brainz (DBTune)
Jamendo (DBtune)
gnoss
SSW Thesaur us
LinkedL CCN
Thesaurus W
LIBRIS
Råcata nål
reegle
Population (En-AKTing)
research data.gov.uk
patents data.gov.uk
Last.FM (rdfize)
Poképédia
Didactal ia
my Experiment
theses.fr
MARC Codes List
LCSH
Ren. Energy Generators
NHS (En-AKTing)
Mortality (En-AKTing)
BBC Programmes
Classical (DB Tune)
Pokedex
Good-win Family
flickr wrappr
semantic web.org
Sudoc
PSH
Norwegian MeSH
EEA
Energy (En-AKTing)
CO2 Emission (En-AKTing)
education.data.gov.uk
OpenEI
BBC Music
Chronicling America
Event Media
Linked MDB
Portuguese DBpedia
SW Dog Food
IdRef Sudoc
ndlna
GND
VIAF
Open Election Data Project
legislation data.gov.uk
Ordnance Survey
Openly Local
Rechtspraak.nl
BBC Wildlife Finder
Telegraphis
Revyu
DDC
Calames
UB Mannheim
Deutsche Biographie
lobid Resources
NSZL Catalog
Europeana
EU Institutions
UK Postcodes
statistics data.gov.uk
New York Times
URI Burner
Greek DBpedia
Open Calais
RDF Book Mashup
data bnf.fr
Ulm
P20
Swedish Open Cultural Heritage
GovWILD
LOIUS
Taxon Concept
Geo Names
World Factbook
Freebase
iServe
OS
BNB
BibBase
ECS Southampton
Wiki
lobid Organisations
STW
GESIS
Brazilian Politicians
ESD standards
reference data.gov.uk
data.gov.uk intervals
DBpedia
ECS Southampton EPrints
ECS (RKB Explorer)
Budapest
OAI
Pisa
RESEX
Scholarometer
Lichfield Spending
ISTAT Immigration
Scotland Pupils & Exams
transport data.gov.uk
NASA (Data Incubator)
Fishes of Texas
Geo Species
Uberblic
dbpedia lite
Project Gutenberg
data dcs
DBLP (FU Berlin)
DBLP (L3S)
DBLP (RKB Explorer)
ACM
IRIT
IBM
NVD
Traffic Scotland
London Gazette
TWC LOGD
Eurostat
Eurostat (FUB)
Geo Linked Data
UMBEL
YAGO
lingvoj
TCM Gene DIT
Daily Med
data-open-ac-uk
ERA
UN/LOCODE
Data Gov.ie
CORDIS (RKB Explorer)
CORDIS (FUB)
GovTrack
Eurostat (Ontology Central)
Linked Sensor Data (Kno.e.sis)
riese
Open Cyc
Lexvo
Enipedia
LinkedCT
Drug Bank
Diseasome
SIDER
Pfam
Eurécom
CiteSeer
RAE2001
Newcastle
Roma
DEPLOY
LOCAH
Courseware
EURES
Linked EDGAR (Ontology Central)
US SEC (rdfabout)
Semantic XBRL
FTS
US Census (rdfabout)
EUNIS
Twarql
WordNet (VUA)
Cornetto
totl.net
Taxonomy
UniProt
UniProt (Bio2RDF)
PDB
PRO-SITE
ProDom
HGNC
STITCH
VIVO Indiana
ePrints
dotAC
IEEE
RISKS
VIVO Cornell
LAAS
JISC
KISTI
NSF
Scotland Geography
GeoWord Net
SMC Journals
Climbing
Linked GeoData
WordNet (W3C)
WordNet (RKB Explorer)
LODE
SISVU
Affymetrix
PubMed
Gene Ontology
SGD
Pub Chem
KEGG Drug
VIVO UF
ECCO-TCP
bible ontology
PBAC
Finnish Municipalities
Piedmont Accomodations
El Viajero Tourism
Italian public schools
Turismo de Zaragoza
Ocean Drilling Codices
AEMET
Metoffice Weather Forecasts
Alpine Ski Austria
GEMET
AGROVOC
Open Data Thesaurus
ChEMBL
OMIM
InterPro
MGI
GeneID
KEGG Enzyme
KEGG Pathway
KEGG Reaction
Janus AMP
EARTh
Weather Stations
Product DB
Linked Open Colors
Smart Link
UniParc
UniRef
UniSTS
Medi Care
KEGG Compound
KEGG Glycan
Homolo Gene
Chem2 Bio2RDF
Yahoo! Geo Planet
National Radioactivity JP
Airports
Product Types Ontology
Italian Museums
Google Art wrapper
UniPath way
Sears
Open Corporates
Amsterdam Museum
meducator
Linked Open Numbers
Reactome
OGOLOD
Media
Geographic
Publications
User-generated content
Government
Cross-domain
Life sciences
As of September 2011

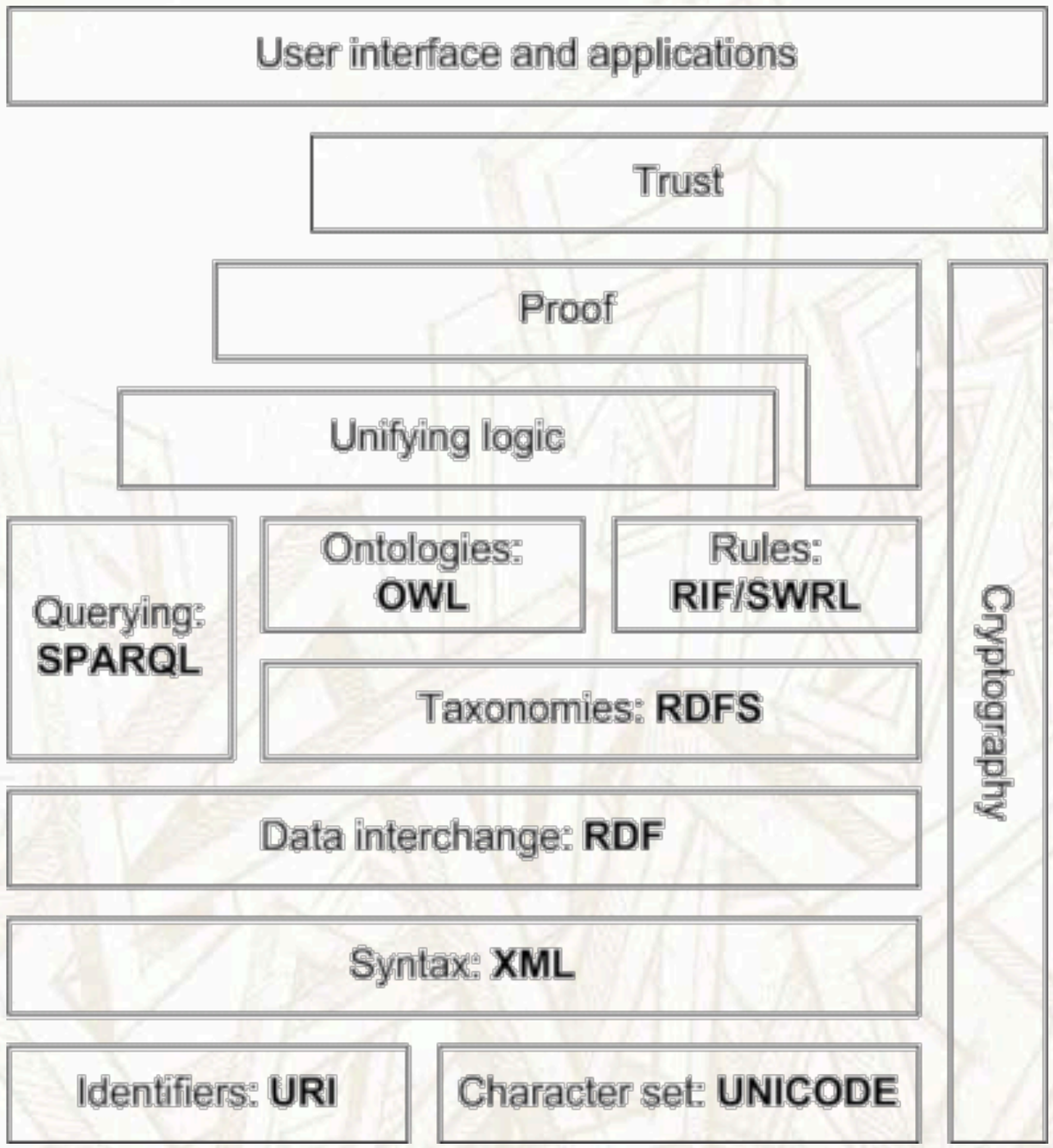
Methodological shift for science?

- As Royal Society president Martin Rees has written (Prospect, November 2010), **big data** will allow us to mine and mash our way to unexpected discoveries and insights. It allows us to ask new questions, ones that we couldn't have asked when science depended on the work of a few people in a single lab working in a limited area of knowledge with just a few gigabytes of processing power
- Some people say that big data also changes the way that we ask questions. Gone are the days of hypothesis-driven science as we know it. Nowadays, it's all about **pattern recognition**.


URL
examples
Kant
clear
Find Relations
relations: (94/100)
lass
link
conne...
num
vi
40/40
54/60
Kant
en
a.org
Status: Idle
isSubjectOf
subject
influenced
influences
Hannah Arendt
influencedBy
Carl Popper
influence.i...
F. W. Nietzsche
Immanuel Kant
Aristotle
Emanuel Swedenborg
Eric Voeg
Philosophers of law
Lisa Zinov'yevna R...
Jean-Jacques Rous
Doctor Angelicus
Empiricists
Leonardo Polo
influencias
influencia


Status: Searching for relations
notableIdea
isSubjectOf
subject
narrower
broader
broader
influenced
influenced
Epistemological the...
isSubjectOf
subject
isSubjectOf
Philosophy of sci...
guid.9202a8c0400...
Philosophers of s...
isSubjectOf
subject
type.type.i...
subject
subject
subject
type.type.i...
subject
subject
Empirici
schoolTrad...
Neo-positivism
Philosophical move...
subject
subject
isSubjectOf
isSubjectOf
philosophi...
philosophi...
Francis Bacon
subject
isSubjectOf
subject
isSubjectOf
isSubjectOf
schoolTrad...
Empiricism
religion
subject
broader
type.permi...
subject
broader
type.object...
Ernst Mach
type.object...
narrower
Empiricists
subject
type.object...
isSubjectOf
influenced
isSubjectOf
guid.9202a8c0400...
influ perce...
common.topic

Semantic Web Layers



Computational ontologies

- "Ontology" originally a philosophical name (but also "semantic web")
- (Computational) Ontologies as (software) components, expressed and managed in standard W3C languages like RDF, OWL, RIF, SPARQL. Also called vocabularies, schemata, conceptual models
- Ontology design is one of the core aspects of semantic technologies
- Design patterns needed for reengineering, data storage, schema construction, mash-up, integration, etc.

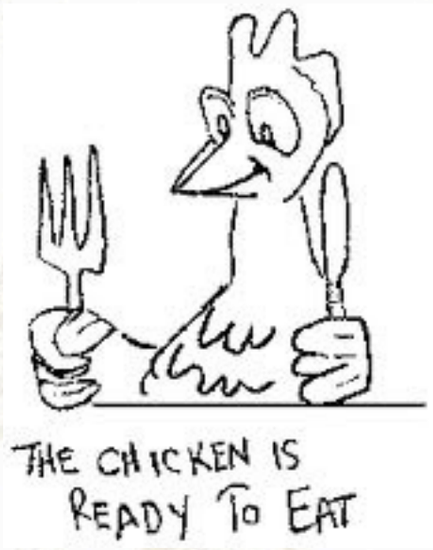
10

Natural and Formal Languages

- Formal interpretation is not (only :)) an academic game
- It gives us a precise way to establish what we are talking about, and therefore to provide reliable automated inferences when needed
- Natural language is able to describe very different types of facts with the same structure, for example ...

… different types of facts …

- Wile buys from ACME [*ground fact*]
- ACME has been reported for abusive discharge [*reported fact*]
- Wile is blonde [*attributive fact*]
- Wile is a coyote [*classification fact*]
- To discharge is to release someone from a job [*meaning fact*]
- To discharge is a transitive verb [*terminological fact*]
- Discharge is nine characters long [*information fact*]
- Discharge is a class [*formal fact*]
- Discharge can be abusive in Italy [*contextual fact*]
- Discharge is subject to obligations and duties [*attitude fact*]
- Discharge represents a failure [*interpretive fact*]

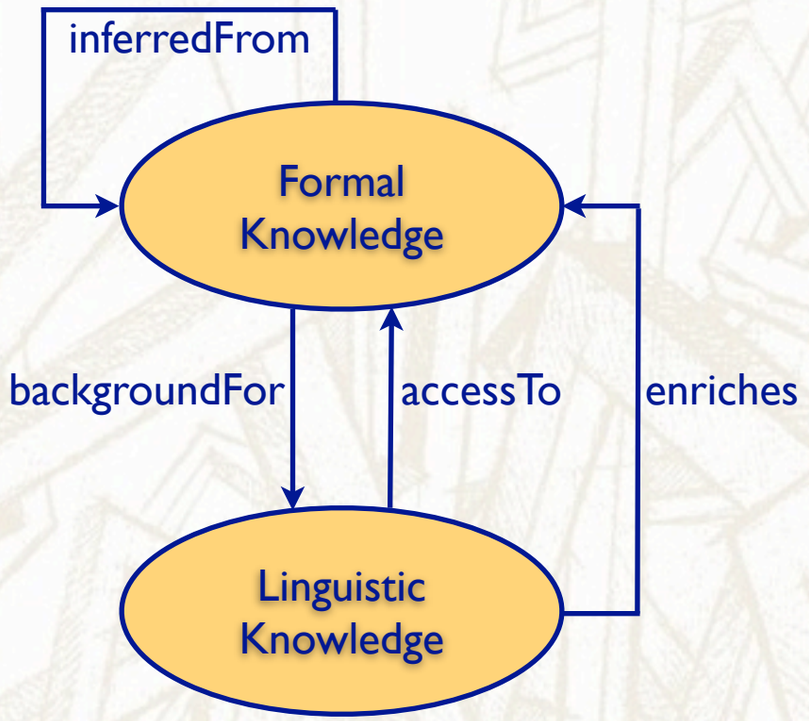


What semantics? (1/2)

- Linguistic semantics
  - From NLP/IR technologies: advanced search, information extraction, automatic tagging, text classification, …
  - Mainly informal background knowledge
  - VSM matrices are the predominant technology
- Formal semantics
  - From SW/AI technologies: data modelling, reengineering, linking, automated reasoning via query, rules, classification, …
  - Ontologies as background knowledge
  - Reasoning and graph traversal are the predominant technologies
- Some attempt to make formal semantics of language
  - E.g. DRT; scalability issues, non-trivial model-theoretical morphing to e.g. OWL
- Also other dimensions, e.g. data engineering, interaction

What semantics? (2/2)

- Trend towards hybridization
  - SW used as background knowledge in NLP/IR
  - Linguistic semantics extracted by NLP/IR tools (NER, TE, RE, FD, etc.), and “reconstructed” in an enriched SW
  - Advanced reasoning performed in the enriched SW

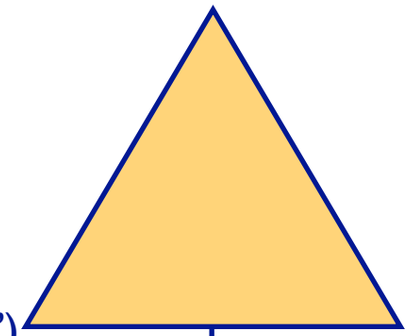


- Two problems: *knowledge soup* and *knowledge boundary*

Good ol’ semiotics

{*ontology classes, properties, schema-level axioms*}
{*lexical concepts, senses, synsets, frames*}

meaning (“interpretant”)



expression (“symbol”)

{*ontology IDs, labels, comments*}
{*texts, bags of words, terms*}

reference (“object”)

{*ontology individuals, factual axioms*}
{*named entities, extracted relations*}

interpreter

{*ontology designer, expert, crowd, learning algorithm*}
{*speaker, hearer, reader, crowd, lexicographer, learning algorithm*}

OWL

- Web Ontology Language
- Formal semantics over RDF
- Large use of small vocabularies
  - FOAF, SIOC, SKOS, VoID, content patterns from large ones (e.g. Dublin Core, DBpedia)
- OWL *representation and reasoning* is the next step at a web scale, but already a reality for intrawebs and small domains
  - Automatic concept classification, consistency and coherence checking, materialization of knowledge
  - E.g. what are the sources that contain sentences about *stipulatio*, cite Ulpian, and have commentaries produced in the last 10 years?
  - E.g. what cases in Common Law systems contain interpretations related to contracts analogous to *stipulatio*?

From NL to FL

- This functionality of natural language cannot be easily reproduced for machine interpretation
- The formal interpretation of OWL can do something with (set-theoretical) *extensional semantics*, cf. basic Aristotelian logic (genus et differentia specifica, inheritance of properties, etc.)
- More is provided by good practices in modelling ontologies and data
- Partly formalizing text elements and corpus processing are often needed
- Some arbitrariness is unavoidable
  - *Requirements, requirements, requirements*

Example: semantic enrichment of annotations

Parallel text annotation (XML)

Text tags

<SPEAKER ID=6 LANGUAGE="EN" NAME="Doyle"> Signor Presidente, come l'onorevole De Rossa anch'io metto in dubbio la ricevibilitv† degli emendamenti. I problemi dell'Irlanda del Nord sono troppo importanti per essere degradati a meschino pretesto di speculazione propagandistica sul Fondo internazionale per l'Irlanda. <P> Vorrei chiedere di considerare il nostro modo di procedere con estrema attenzione, visto che non sarv† possibile avere una discussione sul tema. Anche i miei colleghi, gli onorevoli De Rossa e McCartin, condividono le mie preoccupazioni. <P> Raytheon v® stata accolta a Derry addirittura da due premi Nobel per la pace: John Hume - un nostro collega - e David Trimble. Raytheon sarv† finanziata dall'ente per lo sviluppo industriale dell'Irlanda del Nord, senza che un euro o una sterlina irlandese vengano spesi per questo progetto dal Fondo internazionale per l'Irlanda. Gli emendamenti sono del tutto inopportuni. <P>	<SPEAKER ID=6 NAME="Doyle"> Mr President, like Mr De Rossa I also question the fact of the amendments being in order. The sensitivities of Northern Ireland are too important for any ill-informed bandwagoning on the International Fund for Ireland. <P> I would ask you to consider very carefully how we proceed here, particularly as we are not in a position to have a debate. I share this concern with my colleagues, Mr De Rossa and Mr McCartin. <P> Raytheon has been welcomed to Derry by no less than Nobel Peace Prize winners, John Hume - one of our own colleagues, and David Trimble. Raytheon will be funded by the Industrial Development Board in Northern Ireland. Not one euro nor one Irish pound from the International Fund for Ireland is going to Raytheon. The amendments are totally inappropriate. <P>

Automatic extraction of entities (*w/identity*)

Italian ○ English ⊙ Parameter 1.5 Single sentence false

Insert here your text

Extract

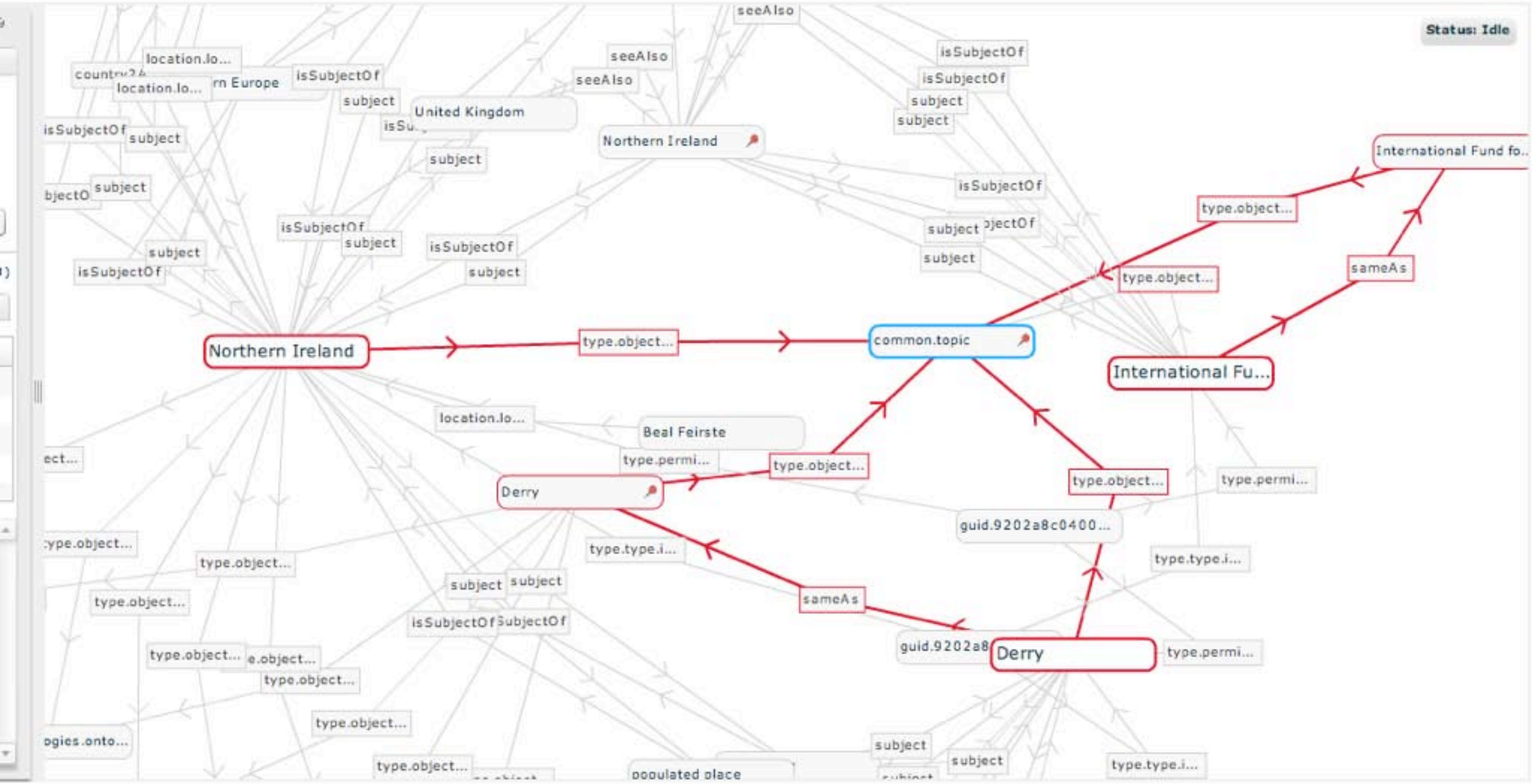
Text
Mr President, like Mr De Rossa I also question the fact of the amendments being in order. The sensitivities of Northern Ireland are too important for any ill-informed bandwagoning on the International Fund for Ireland. I would ask you to consider very carefully how we proceed here, particularly as we are not in a position to have a debate. I share this concern with my colleagues, Mr De Rossa and Mr McCartin. Raytheon has been welcomed to Derry by no less than Nobel Peace Prize winners, John Hume - one of our own colleagues, and David Trimble. Raytheon will be funded by the Industrial Development Board in Northern Ireland. Not one euro nor one Irish pound from the International Fund for Ireland is going to Raytheon. The amendments are totally inappropriate.

Key Words	URI First Choice	URI Second Choice
ireland	http://en.wikipedia.org/wiki/Ireland;	http://en.wikipedia.org/wiki/Northern_Ireland;
irish pound	http://en.wikipedia.org/wiki/Irish_pound;	none
derry	http://en.wikipedia.org/wiki/Derry;	http://en.wikipedia.org/wiki/County_Londonderry;
international fund	http://en.wikipedia.org/wiki/International_Fund_for_Ireland;	none
david trimble	http://en.wikipedia.org/wiki/David_Trimble%2C_Baron_Trimble;	http://en.wikipedia.org/wiki/David_Trimble_%28congressman%29;
mr president	http://en.wikipedia.org/wiki/Mr._President_%28musical%29;	http://en.wikipedia.org/wiki/Mr._President_%28band%29;
nobel peace prize	http://en.wikipedia.org/wiki/Nobel_Peace_Prize;	http://en.wikipedia.org/wiki/Nobel_Peace_Prize_Concert;
bandwagoning	http://en.wikipedia.org/wiki/Bandwagon_effect;	http://en.wikipedia.org/wiki/Johnny_Johnson_and_the_Bandwagon;
northern ireland	http://en.wikipedia.org/wiki/Northern_Ireland;	http://en.wikipedia.org/wiki/BBC_Northern_Ireland;
john hume	http://en.wikipedia.org/wiki/John_Hume;	http://en.wikipedia.org/wiki/John_Hume_%28bishop%29;
raytheon	http://en.wikipedia.org/wiki/Raytheon;	http://en.wikipedia.org/wiki/Raytheon_Sentinel;

Parallel text annotation (XML+RDFa)

Semantic tags

Signor Presidente, come l'onorevole De Rossa anch'io metto in dubbio la ricevibilitv† degli emendamenti. I problemi dell' Irlanda del Nord sono troppo importanti per essere degradati a meschino pretesto di speculazione propagandistica sul Fondo internazionale per l'Irlanda. Vorrei chiedere di considerare il nostro modo di procedere con estrema attenzione, visto che non sarv† possibile avere una discussione sul tema. Anche i miei colleghi, gli onorevoli De Rossa e McCartin, condividono le mie preoccupazioni. Raytheon v® stata accolta a Derry addirittura da due premi Nobel per la pace: John Hume - un nostro collega - e David Trimble. Raytheon. sarv† finanziata dall'ente per lo sviluppo industriale dell'Irlanda del Nord, senza che un euro o una sterlina irlandese vengano spesi per questo progetto dal Fondo internazionale per l'Irlanda. Gli emendamenti sono del tutto inopportuni.	Mr President, like Mr De Rossa I also question the fact of the amendments being in order. The sensitivities of Northern Ireland are too important for any ill-informed bandwagoning on the International Fund for Ireland. I would ask you to consider very carefully how we proceed here, particularly as we are not in a position to have a debate. I share this concern with my colleagues, Mr De Rossa and Mr McCartin. Raytheon has been welcomed to Derry by no less than Nobel Peace Prize winners, John Hume - one of our own colleagues, and David Trimble. Raytheon will be funded by the Industrial Development Board in Northern Ireland. Not one euro nor one Irish pound from the International Fund for Ireland is going to Raytheon. The amendments are totally inappropriate.


Status: Idle
Northern Ireland
type.object...
common.topic
International Fund fo...
International Fu...
sameAs
Derry
Beal Feirste
United Kingdom
location.lo...
isSubjectOf
subject
seeAlso
guid.9202a8c0400...
guid.9202a8
type.permi...
type.type.i...
populated place

More enhancement

- Information enrichment
  - Extraction (see previous example)
  - RDF-ization of unstructured data
  - Materialization (*in praesentia* inference)
  - Classification (*in absentia* inference)
- Advanced search paradigms
  - Semantic search
  - Semantic exploration

A knowledge pattern science



Patterns* in general

- “Invariances across observed data or objects”
- They exist in natural, social, cognitive, or abstract worlds
- Mathematical pattern science is about symbols, i.e. non-interpreted information objects
- Objects of knowledge engineering are interpreted (either formally or cognitively)
- Mutual support/dependencies

* A **frame** will be here defined as the type of objects that are used for interpreting **patterns**

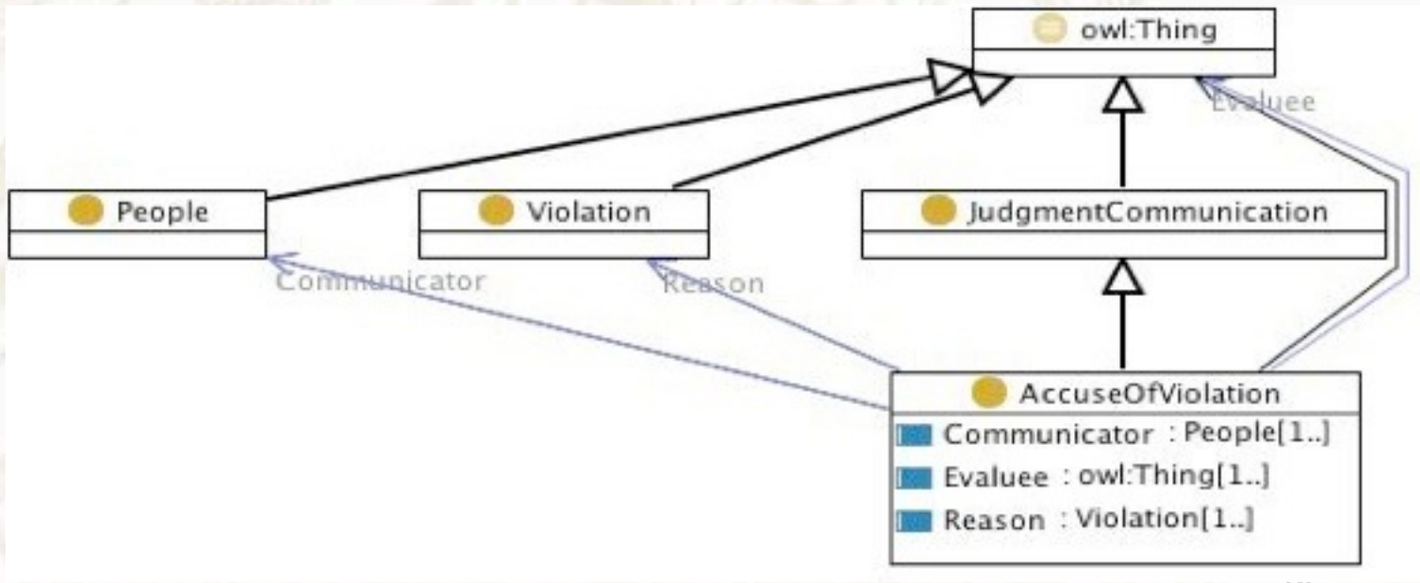
The frame hypothesis (1/2)

- Most realistic tasks require knowledge shaped in a way that allows to navigate through the data meaningfully
  - i.e. similarly to how a human interprets and aggregates them
- The typical shape of meaningful knowledge is that of a *frame (or situation, context, schema)*
- Minsky 1974:
  - *there would be large advantages in having mechanisms that could use these same structures both for thinking and for communicating*



The frame hypothesis (2/2)

- If we do not take that “similar” too seriously, a plausible hypothesis is that the *unit of meaning* in Semantic Web technologies is the *frame*, as opposed to scattered classes or properties (binary relations)
- Agents/reasoners must be able to recognize such *frames* and reason on them
- Current applications, either from DB of SW technologies, try to catch the “perfectly fit” frame ... just look at Facebook
- In ontology design, frames are called *knowledge patterns*, a special kind of *design patterns*

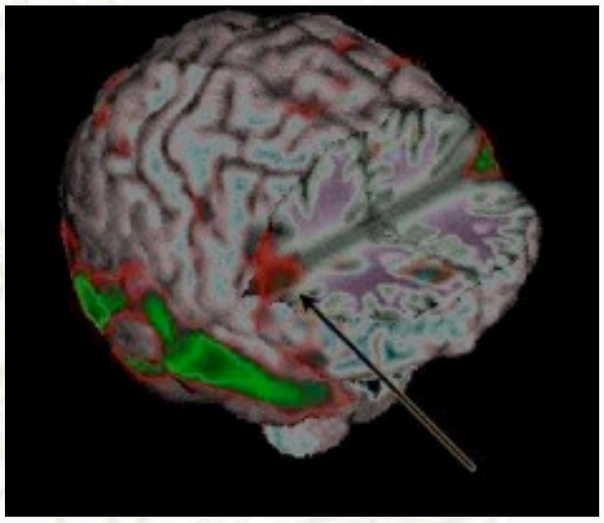


Cognitive foundations

- A search for the *relevant* units of meaning (not the primitives)
- Agents' understanding is based on [*patterns*] abstracted from previously occurred situations involving those agents, which are adapted and recombined on-the-fly in novel situations
  - Bartlett's experiments on schemata (1932), Neisser (1967)
  - Piaget's experiments on schemata (1954)
  - Fillmore's frame semantics (1968)
  - Minsky's frames (1974)
  - Schank's scripts (1977)
  - Gibson's affordances (1977)
  - Biederman's experiments on scene recognition (1982)
  - Barsalou's ad-hoc goal-oriented categories (1983)
  - Lakoff's conceptual metaphors (1987), Langacker's compositional paths (1987)
  - Barsalou's simulators (1999)
  - Bar's associative-analogical network of frames for anticipation (2007)
  - Rizzolatti, Iacoboni, Gallese, etc. results and plausibility of mirror-neuron-system-based frames (2008)







Sample frame-like structures

- Lexical frames
- Microformats, Infoboxes
- Contractor's structural signatures
- Query types
- Competency questions
- ≥2-ary relations
- OWL/RDFS classes with proper (locally complete?) sets of restrictions or properties
- Content ontology design patterns
- Data and conceptual models (if modularized according to requirements)
- Some HTML structures
- Some XML schemata

FrameNet *Cure* frame

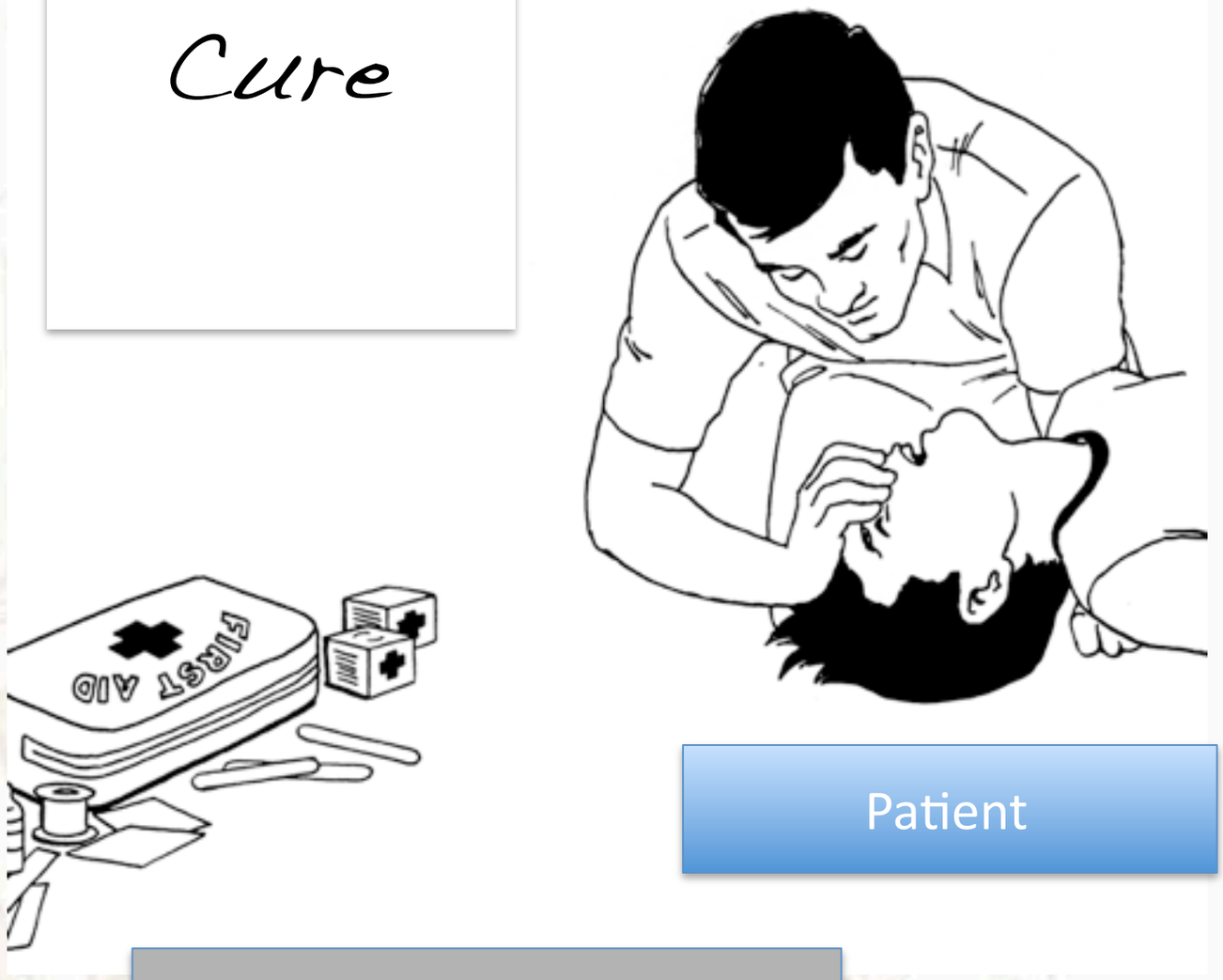
This frame deals with a Healer treating and curing an Affliction (the injuries, disease, or pain) of the Patient, sometimes also mentioning the use of a particular Treatment or Medication.

Medication [Med] **Semantic Type** Physical_entity	The injested, applied, injected, etc. substance designed to cure the Patient. He needs prolonged TREATMENT with anibiotics. Note the tight relationship between Treatment and Medicine.

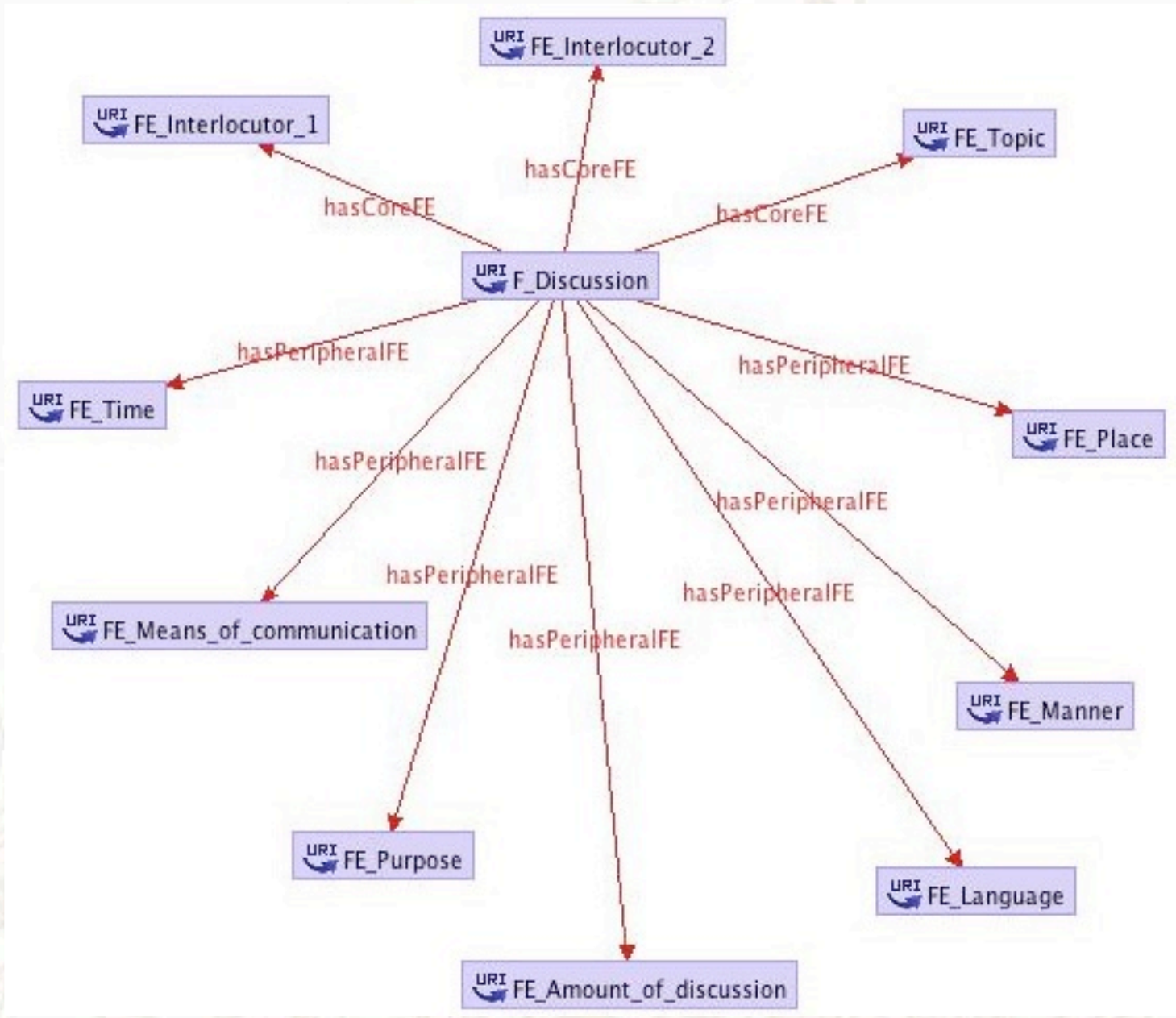
Healer [Hlr] **Semantic Type** Sentient	The Healer, anyone who treats or cures the Patient, occurs as the External Argument of verbs: Doctors ALLEVIATED his suffering.

http://framenet.icsi.berkeley.edu/

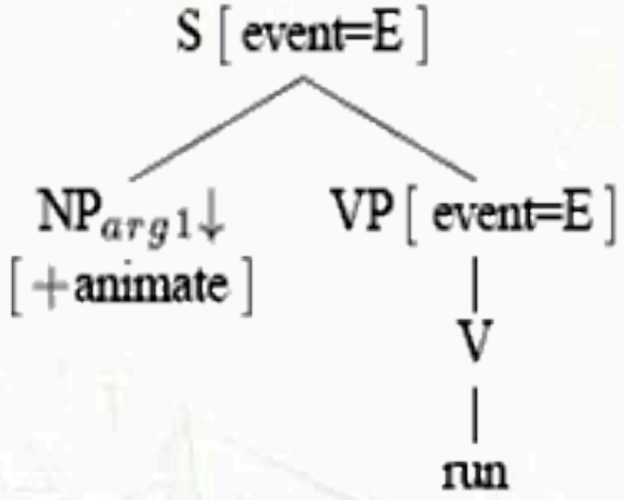
Cure

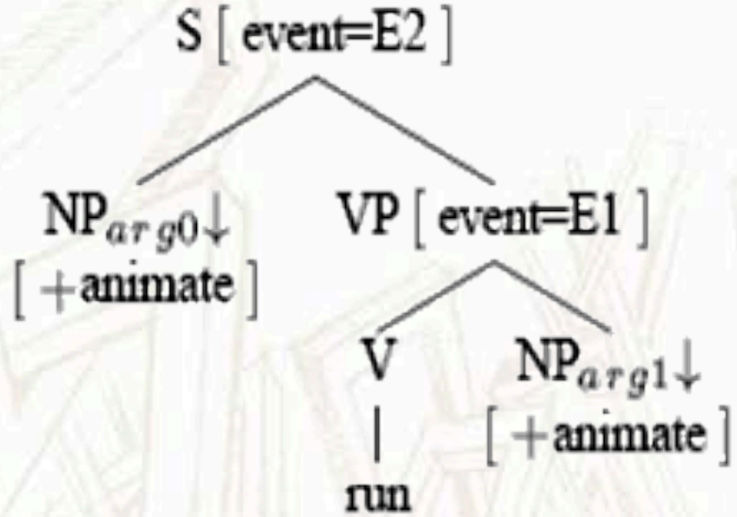


FrameNet *Discussion* frame



VerbNet *Motion* verb class





$$motion(during(E), X_{arg1})$$

$$cause(during(E2), X_{arg0}, E1) \wedge motion(during(E1), X_{arg1})$$

"He kicked the door open."	cause(Agent, E) manner(during(E), directedmotion, Agent) not(contact(during(E), Agent, Patient)) manner(end(E), forceful, Agent) contact(end(E), Agent, Patient) Pred(result(E), Patient)	
"He hit the door open with his foot."	cause(Agent, E) manner(during(E), directedmotion, Instrument) not(contact(during(E), Instrument, Patient)) manner(end(E), forceful, Instrument) contact(end(E), Instrument, Patient) Pred(result(E), Patient)	
"He hit the window to pieces."	cause(Agent, E) manner(during(E), directedmotion, Agent) not(contact(during(E), Agent, Patient)) manner(end(E), forceful, Agent) contact(end(E), Agent, Patient) Pred(result(E), Patient)	

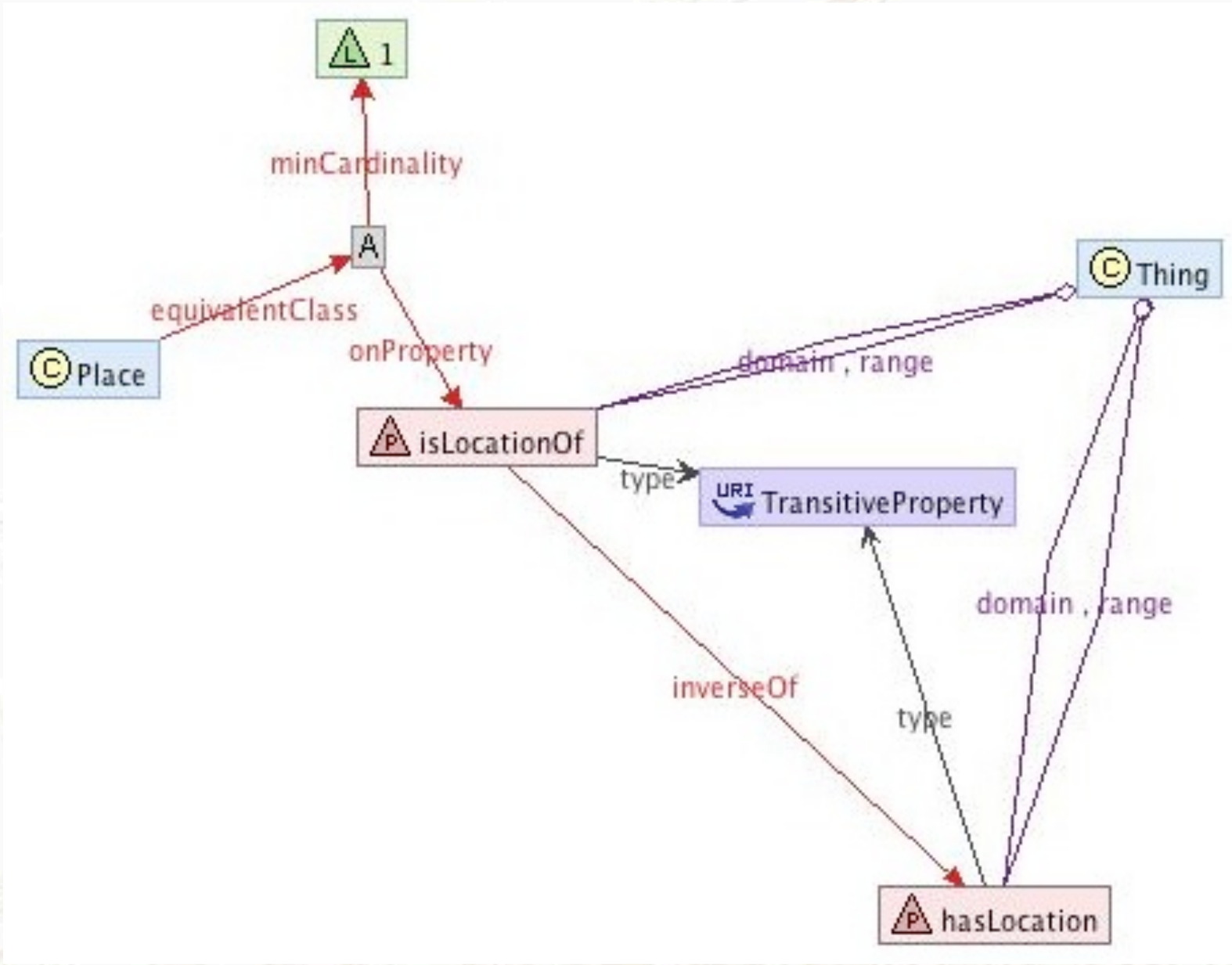
CLib *Attach* component

```
(Attach has
    (superclasses (Action))
    (required-slot (object base))
    (primary-slot (agent)) )

(every Attach has
    (object ((exactly 1 Tangible-Entity) (a Tangible-Entity)))
    (base ((exactly 1 Tangible-Entity) (a Tangible-Entity)))

(every Attach has
    (preparatory-event ((:default
    (a Make-Contact with
            (object ((the object of Self)))
            (base ((the base of Self))))
    (a Detach with
            (object ((the object of Self)))
            (base ((the base of Self)))) ))))
```

ODP *Place* content pattern



Situations and frames

- *Situations* are experienced as complex, Gestalt phenomena rather than collections of events and objects
- The cognitive operation of unifying a situation is principled (*principium individuationis*) (cf. L. Barsalou)
- This unifying cognitive act is the application of a *frame* to entities (available through perceptual – or machine-available – data)
- The process of applying a frame is carried out by a cognitive agent with some *semiotic system*

Meaning and Relevance

- Being meaningful is usually associated with *relevance in context*, i.e. having a clear *boundary* in order to matter to someone
- How much state-of-art applications in the data/semantic web address relevance in context?

Applications (I)

www.zemanta.com

Your content enhanced!

Aldo Gangemi graduated in philosophy, and then worked in medical informatics, knowledge engineering, and semantic technologies, including the Semantic Web. He's senior researcher at the CNR Institute of Cognitive Sciences and Technology in Rome, and head of the Semantic Technology Lab. He also co-founded the Laboratory for Applied Ontology. His work on ontology engineering dates back to 1992, when pioneering the application of ontologies to knowledge organization systems in biomedicine. His current research focuses on pattern-based ontology design, semantic social networks, collaborative modeling, hybridation of NLP, lexical and semantic resources, and metamodels for heterogeneous knowledge integration. He has published more than 100 refereed articles in proceedings of international conferences, journals, and books, and has been involved in several projects funded by the European Union, governmental organizations, and industrial companies (either as research partner or consultant). STLab is responsible for CNR in the EU projects NeOn, IKS, and BONy, and has worked in the seminal ontology-related projects Galen, WonderWeb, OntoWeb, Metokis, and the Eureka project IKF. He has organized several conferences and workshops, and regularly tutors in PhD courses and international schools. Notable research products include the DOLCE ontology, its OWL versions and extensions within the DnS paradigm, e.g. DOLCE+DnS Ultralite, the Ontology Design Patterns (ODP) portal, the Semantic Scout for CNR researchers, etc. In legally-related fields, he has organized workshops on regulatory and legal ontologies, and has made active research on reusable components for legal ontology design, including the Core Legal Ontology. He is member of the advisory boards of ECDC and SEMIC.EU.



Image via Wikipedia

Related articles by Zemanta

Thank the Lord it's Friday! (blindscientist.genedrift.org)
Knowledge Acquisition from Social Awareness Streams (slideshare.net)
Brown Bag Lunch: Methods for Semantic Discovery, Annotation and Mediation (mkbergman.com)



Content recommendations



Somehow related

Unrelated

Zemanta automatic enrichment of open domain text:

Great enrichment!

Relevant enrichment?

Your content enhanced!

Stefano Nolfi is research director at the Institute of Cognitive Science and Tecnology and head of the Laboratory of Artificial Life and Robotics. Stefano coordinated the European project ECAgents (Embodied and Communicating Agents) and is team leader of the European Projects I-TALK (Transfer of Action and Language Knowledge in Robots) and Swarmanoid (Towards Humanoid Robotic Swarms Humanoids).



Image by Getty Images via @daylife

Stefano has been a fellow of the: Centre for Research in Language, University of California, San Diego, USA; Laboratory of Microcomputing, Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland; SONY Computer Science Laboratory, Tokyo, Japan; Institute of Advanced Studies of Berlin, Germany; University of New South Wales, Canberra, Australia.

Stefano's research activities focus on Adaptive Behaviours, Autonomous Robots, Embodied Cognition, Collective and Swarm Intelligence, Evolutionary Robotics, Artificial Life, Evolution of Language, Complex Systems.

The main themes underlying my work are: (a) that behavioural strategies and neural mechanisms are understood better when an organism (living or artificial) is caught in the act, that is when one considers situated and embodied agents in their interaction with the environment; (b) that to understand how natural agents behave and to build useful artificial agents one should study how living organisms change, phylogenetically and ontogenetically as they adapt to their environment.

Related articles by Zemanta

Public is relaxed about artificial life, says study (independent.co.uk)
Robotics project titles, 2009 2010 ieee ncct final year projects (slideshare.net)
So, how real are today's robots? (cnn.com)

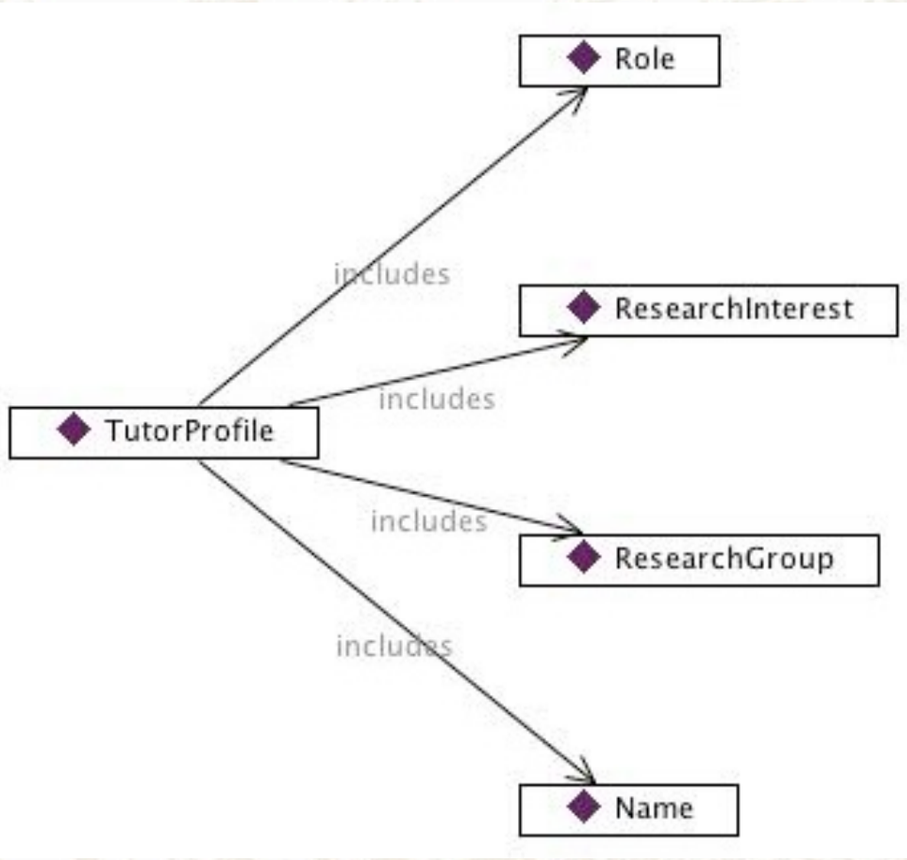


Content recommendations



Tutor info: *bounded*

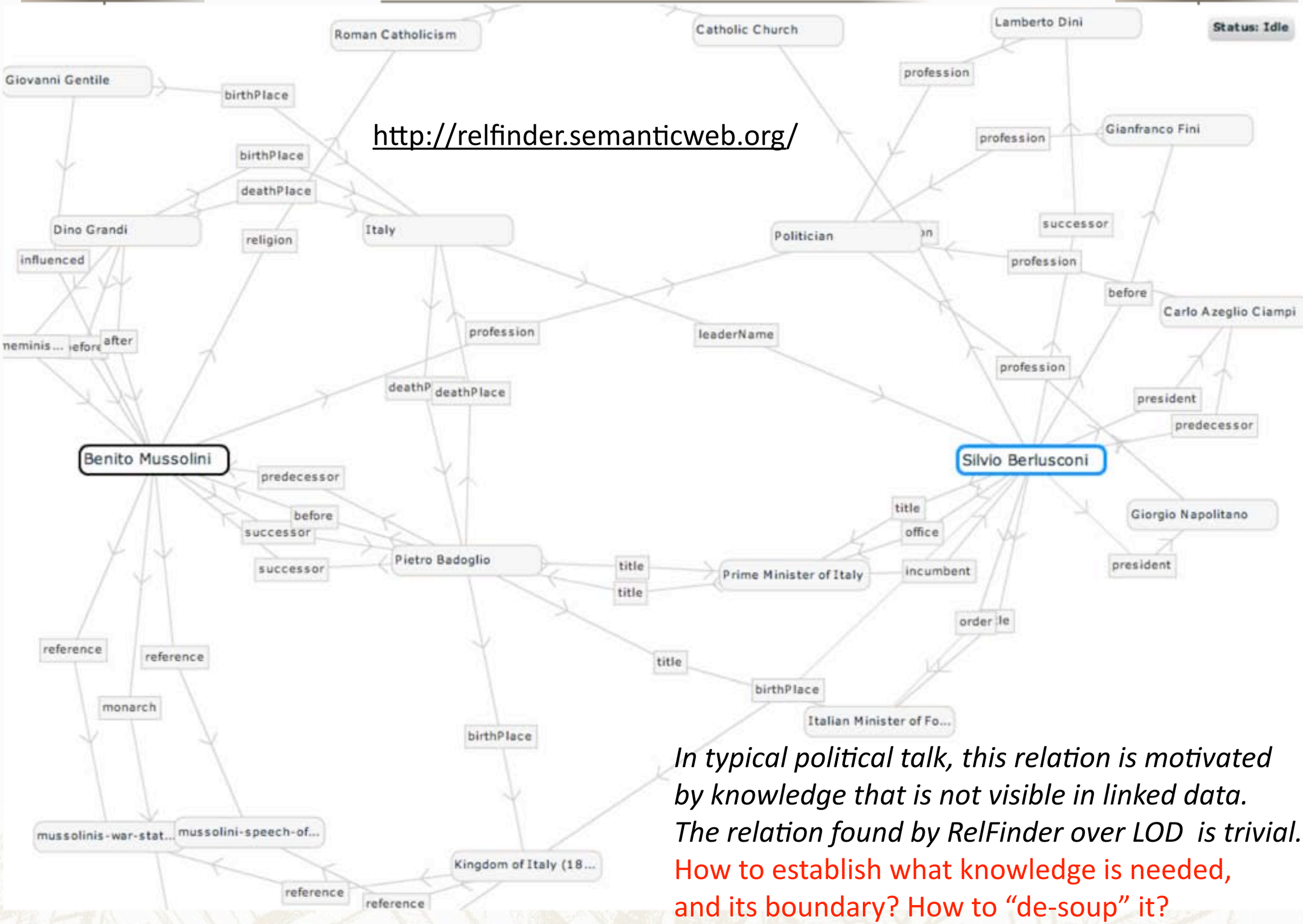
- *Name*: Aldo Gangemi
- *Institute*: ISTC-CNR, Rome, Italy
- *Research Group*: Semantic Technology Laboratory (STLab)
- *Role*: Senior Researcher
- *Research interest*: ontology design, KE+NLP, Enterprise 3.0, ...



Tutor info: *unbounded*

- *Name*: Aldo Gangemi
- *Institute*: ISTC-CNR, Rome, Italy
- *Research Group*: Semantic Technology Laboratory (STLab)
- *Role*: Senior Researcher
- *Research interest*: ontology design, KE+NLP, Enterprise 3.0, ...
- *Knows*:
- *Likes*:
- *Author of*:
- *Plays*:
- *Birthday*:
- *Car*:
- *Zodiacal sign*:
- ...

How to establish the boundary?

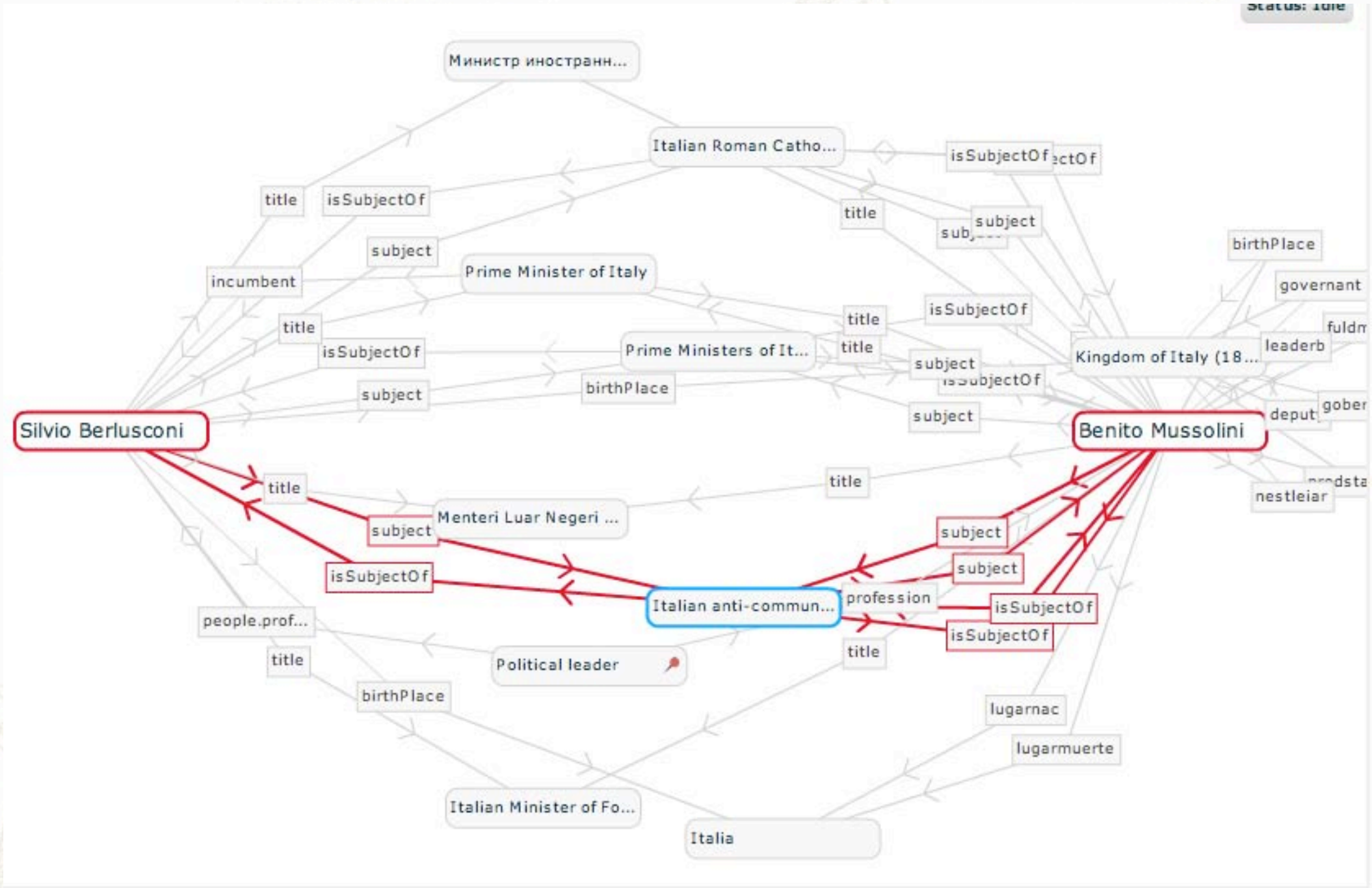


http://relfinder.semanticweb.org/

In typical political talk, this relation is motivated by knowledge that is not visible in linked data. The relation found by RelFinder over LOD is trivial.

How to establish what knowledge is needed, and its boundary? How to “de-soup” it?

After some mishmashing, something emerges, but not yet something specific ...



Isn't KR enough?

- Knowledge representation has progressed a lot since the seventies
- Its original motivations, as e.g. described by Minsky and Brachman, have been drastically simplified
  - concentration on representation and complexity
- A lot of different KR languages can represent frames: description logics, frame logic, first-order logic, conceptual graphs, etc.
  - but: which frames should be used, for what purpose, and with what pros and cons
- We need some *design*

How many frames?

- STLab is collecting, reengineering, and aligning frames from different knowledge formats
  - Linguistics: FrameNet, VerbNet, NLP-extracted selectional restrictions, etc.
  - CLib and other AI repositories
  - Web data: microformats, microdata, HTML templates
  - Linked data: induced schemas, frame discovery
  - (Semantic web) ontologies: extracted modules

Main issues

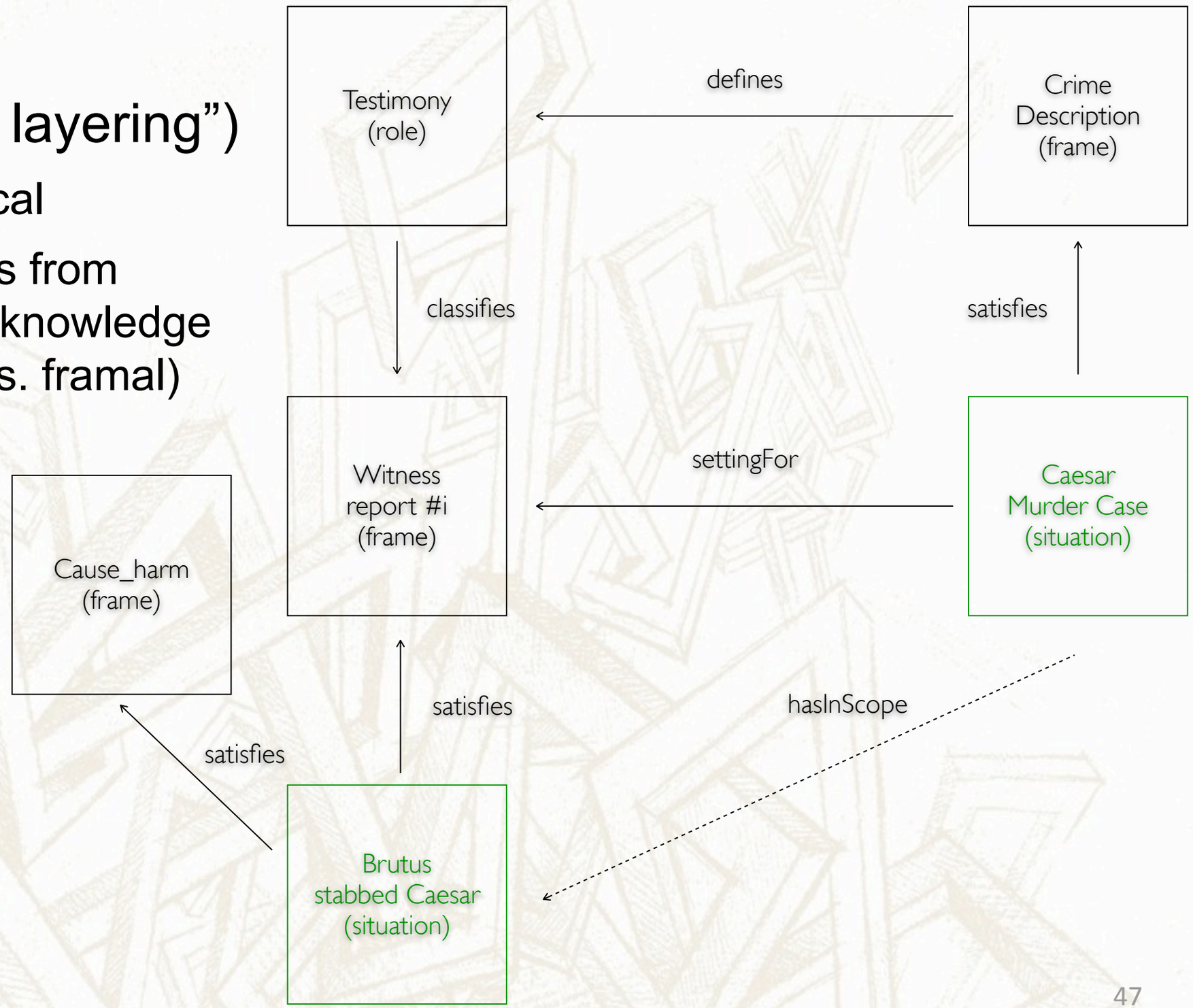
- The invariance, or boundary problem
  - what is invariant in each knowledge format? what are its boundaries?
- The category-invariance problem
  - how to establish that invariances from different knowledge formats are similar?
- Early results
  - Frames from Wikipedia crowd-sourced links
  - Frames from linked data
  - Ongoing experiments with deep parsing from text and non-structure-preserving morphing to known or novel frames
  - A relaxed FOL with meta-level sugar as a common representation
  - But how to distinguish heterogeneous knowledge?

Frame semantics 1/3

- Minimal
  - n-ary polymorphic relation:
    - give(*?donor, ?receiver, ?gift, ?time, ?space, ...)*
- More flexible
  - neo-davidsonian (situation, token-boundary):
    - give(*?s*); donor(*?s,?d*); receiver(*?s,?r*); ...
    - requires identification constraint (e.g. *owl:hasKey*)
  - duper (frame, type-boundary):
    - Frame(*Giving*); satisfies(*?s, Giving*)
  - super-duper (role structure for frames):
    - defines(*Giving, {Donor; Receiver; Object; ...}*)
    - classifies(*Donor, ?d*) ; ...

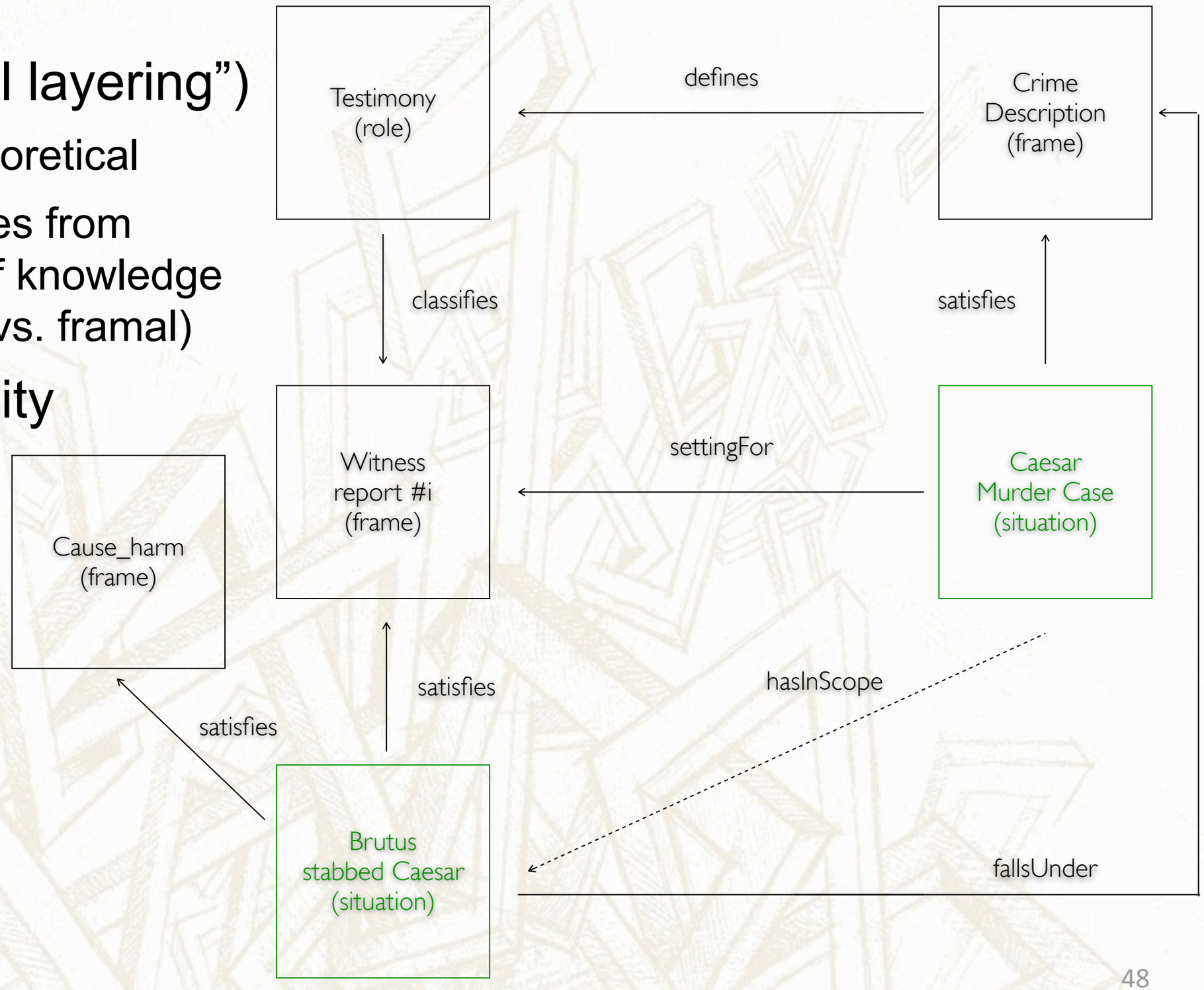
Frame semantics (2/3)

- Diagonality ("epistemological layering")
  - Not type-theoretical
  - First-order entities from different types of knowledge (e.g. situational vs. framal)



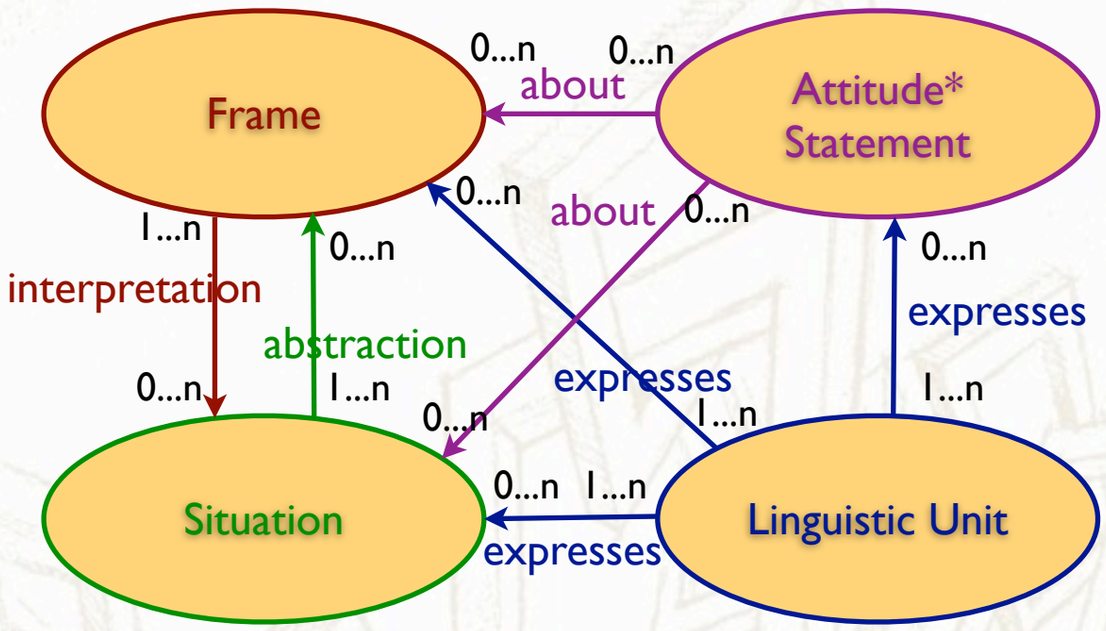
Frame semantics (3/3)

- Diagonality ("epistemological layering")
  - Not just type-theoretical
  - First-order entities from different types of knowledge (e.g. situational vs. framal)
- Indirect diagonality

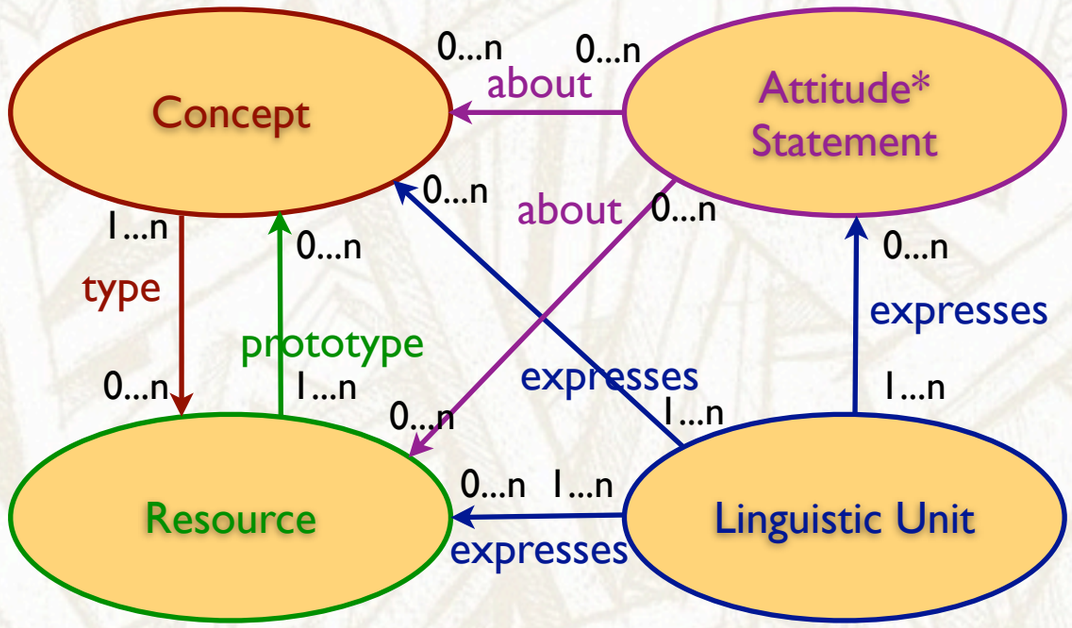


Four types of context

with boundaries (super-duper)



without boundaries



*"attitude" in a broad sense:
obligations, duties, desires, competences, authorship, provenance, trust, etc.

Possible links to “assignment semantics”

- *type* and *subject* relations easily map to *assignment*
- compatibility works well across similar typing + authorship/provenance
- less clear how compatibility works for binary relations
- framing (interpretation) needs *multiple assignments* (a bundle) within a boundary
- what about language?

Lexicon in context (1/2)

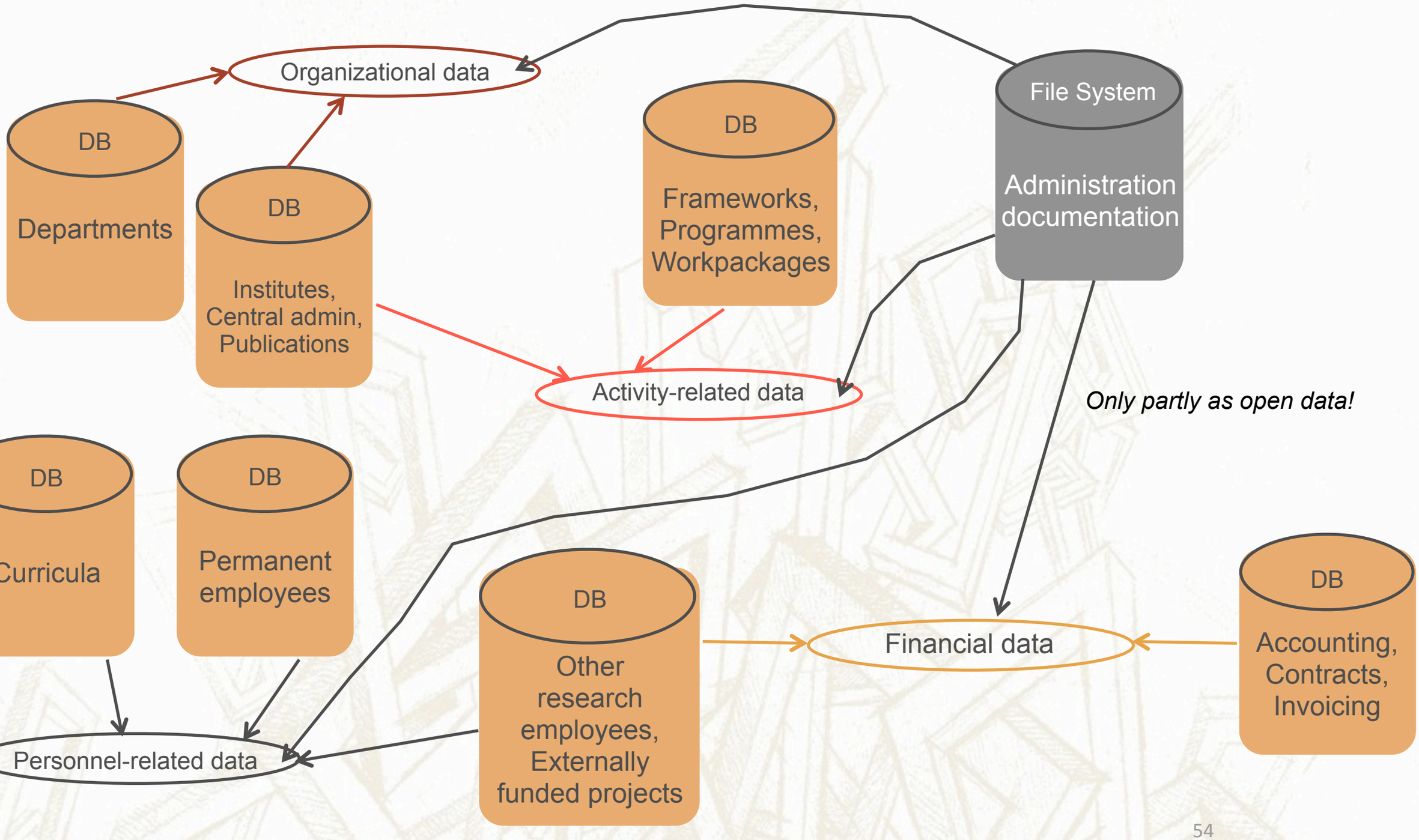
- In a semiotic perspective, language expresses any entity from any context
- Semiotics takes meaning as “interpretants” i.e. any reformulation, paraphrase or other linguistic clue
- Pragmatics takes attitudes as linguistic (speech or social) acts
- Situations depend on framing (and attitudes)
- Can we safely assume a semiotic cage, so reducing the problem of the ontology-lexicon interface to nothing?

Lexicon in context (2/2)

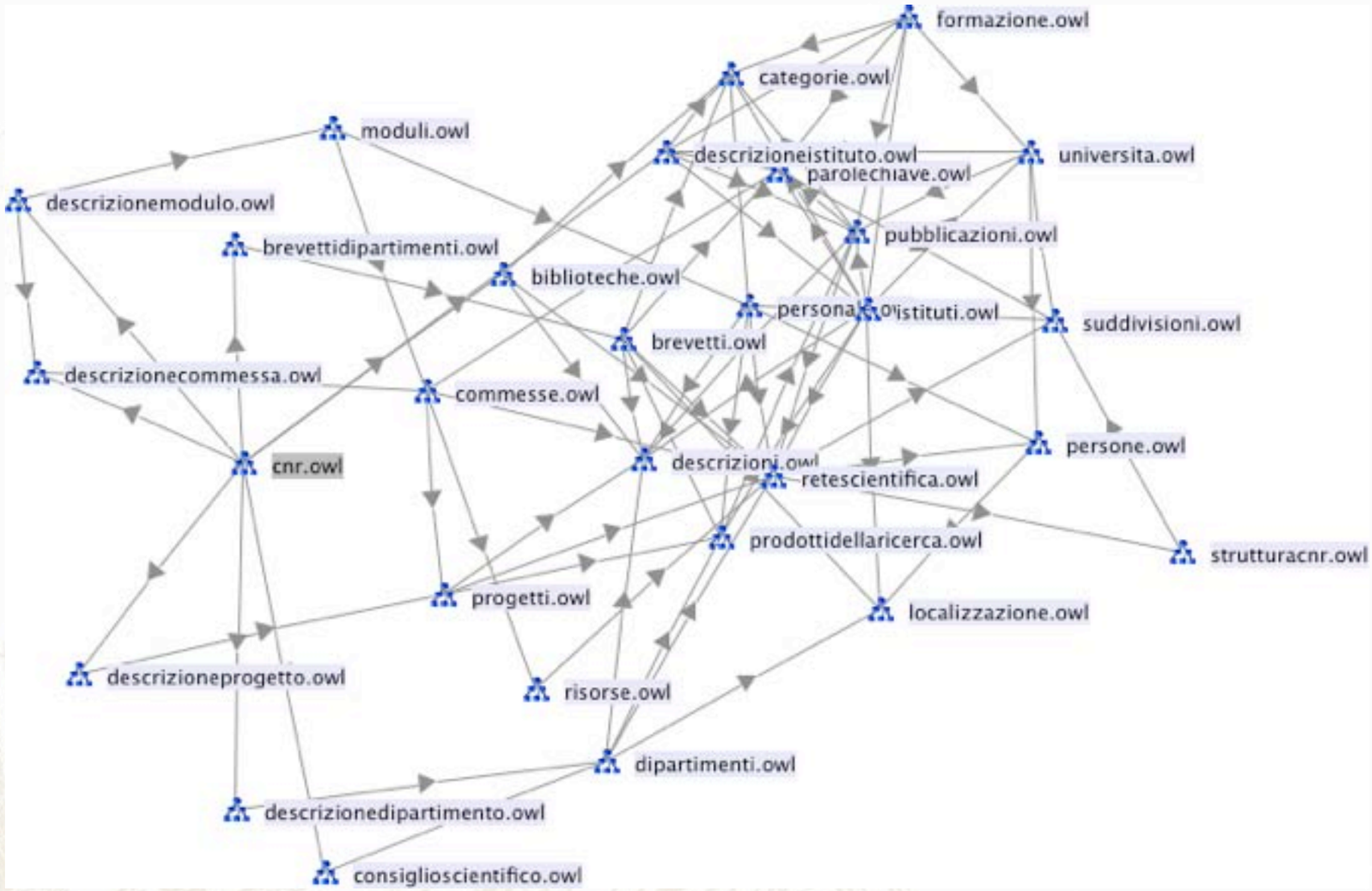
- Theoretically speaking, yes
- Practically, no
  - at least on the Semantic Web, the difference between data knowledge and linguistic knowledge is still clear and broad
  - attempts to reduce one into the other are probably to fail
  - hybridisation of techniques, assumptions, and modeling choices seems promising

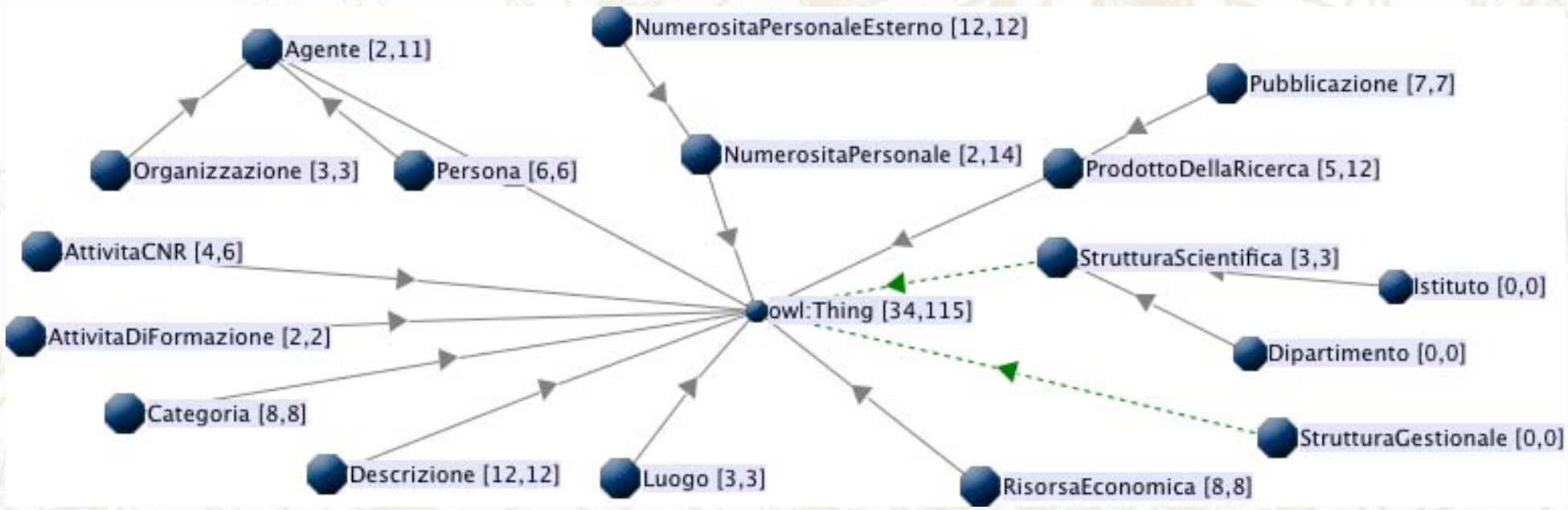
Example: research data integration @ http://data.cnr.it

The CNR data sources



The CNR ontology





Consiglio Nazionale delle Ricerche

Home | English version | Mappa | Commenti | Sondaggio | Staff | Contattaci | RSS

Cerca nel sito

Semantic Scouting

kant — Search

Type: base latent Lang: IT EN

Elenco completo | Persone | Attività

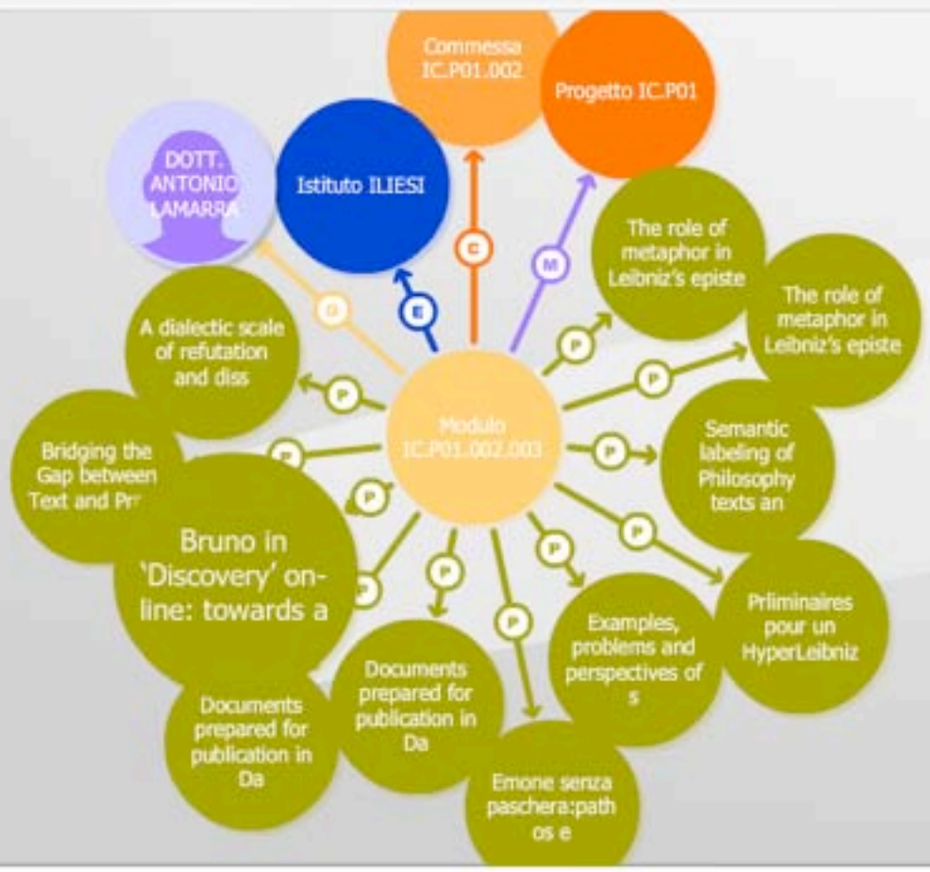
- Hansmichael Hohenegger (Unità di personale esterno)
- Storia delle idee e della terminologia di cultura nell'età moderna (IC.P01.002.001) (Modulo)
- Formazione alla lessicografia filosofica e scientifica (IC.P01.006.001) (Modulo)
- Formazione alla lessicografia filosofica e scientifica (IC.P01.006) (Commessa)
- Discovery - Digital Corpora for Virtual Research in Philosophy (IC.P01.002.003) (Modulo)
- Storia delle idee e della terminologia di cultura nell'età moderna (IC.P01.002) (Commessa)

Home | Il CNR | I servizi | News | Eventi | Istituti | Focus

webtemp.src.cnr.it/semanticscouting/semar

CNR Semantico

Bruno in 'Discovery' on-line: b



Tipo
Abstract

Label
Bruno in 'Discovery' on-line: towards a semantic enrichment of Bruno's Latin works

URL
http://www.cnr.it/ontology/cnr/individuo/pubblicazioni-abstract/ID20873

Graph Explorer

Consiglio Nazionale delle Ricerche

Home | English version | Mappa | Commenti | Sondaggio | Staff | Contattaci | RSS

Semantic Scouting

diritto romano Search

Type: base latent Lang: IT EN

Elenco completo | Persone | Attività | Strutture CNR

- Dott. Vincenzo Casolino (Unita di personale interno) sc:0.7418197
- Riccardo Cardilli (Unita di personale esterno) sc:0.5563648
- Sandro Schipani (Unita di personale esterno) sc:0.4946223
- Sull' "uso del diritto romano in Cina". Formazione del diritto cinese nell'ambito del sistema giuridico romanistico (IC.P07.005.001) (Modulo) sc:0.48894736
- Laura Formichella (Unita di personale esterno) sc:0.44192487
- Giuseppe Terracina (Unita di personale esterno) sc:0.43719235
- Dott.ssa Sabrina Lanni (Unita di personale interno) sc:0.4174379
- Dott. David Fabio Esborraz (Unita di personale interno) sc:0.3900252
- Sull' "uso del diritto romano" in Cina. Formazione del diritto cinese nell'ambito del sistema giuridico romanistico (IC.P07.005) (Commessa) sc:0.3864348
- Diritto romano e informatica (IC.P09.002.001) (Modulo) sc:0.37860563
- Formazione dei sottosistemi del sistema giuridico romanistico, processi di integrazione regionale, unificazione del diritto (IC.P07.002.001) (Modulo) sc:0.3505379
- Franco Vallocchia (Unita di personale esterno) sc:0.34591013
- Dott. Giancarlo Taddei Elmi (Unita di personale interno) sc:0.34128606
- Pierangelo Catalano (Unita di personale esterno) sc:0.31925353
- Nicola Palazzolo (Unita di personale esterno) sc:0.31869587
- Idea giuridica di Roma e diffusione del diritto romano (informazione e formazione anche attraverso strumenti informatici) (IC.P09.002) (Commessa) sc:0.31361654
- Dott. Natale Rampazzo (Unita di personale interno) sc:0.31248808
- Formazione dei sottosistemi del sistema giuridico romanistico, processi di integrazione regionale, unificazione del diritto (IC.P07.002) (Commessa) sc:0.2994452
- I diritti fondamentali come fattori problematici nei processi di sviluppo e di integrazione degli ordinamenti giuridici (IC.P08.007) (Commessa) sc:0.2967202

Frame discovery

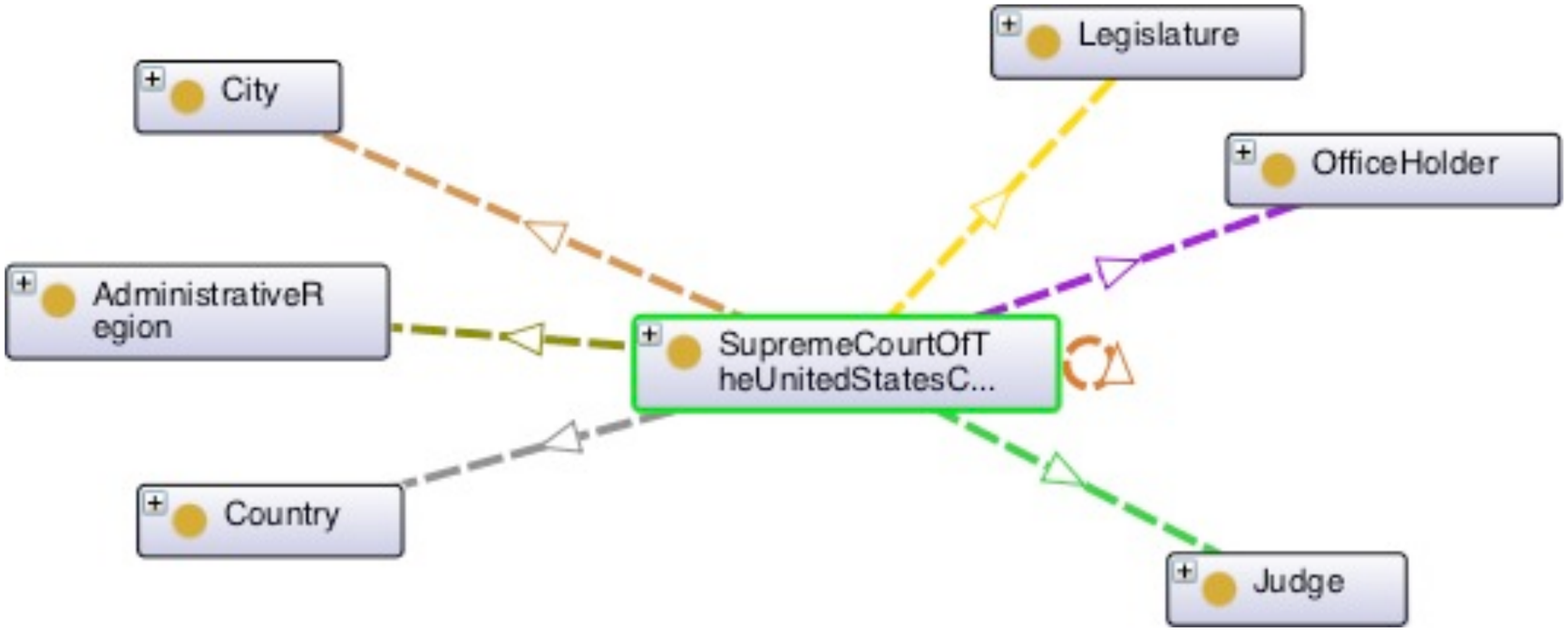
Encyclopedic knowledge patterns

- Crowd-sourced meaning in wikipedia frames: empirical knowledge engineering
- We have discovered invariances in Wikipedia links
  - ≈108M Wikipedia links
  - 272 types
  - “typical types” linking to resources of a certain type cluster above 11% of coverage
  - 7±2 typical types stand out (magical number!)
- We have evaluated them in a user study
  - good correlation between users and between users and standout types
- Cf. ISWC2011 paper

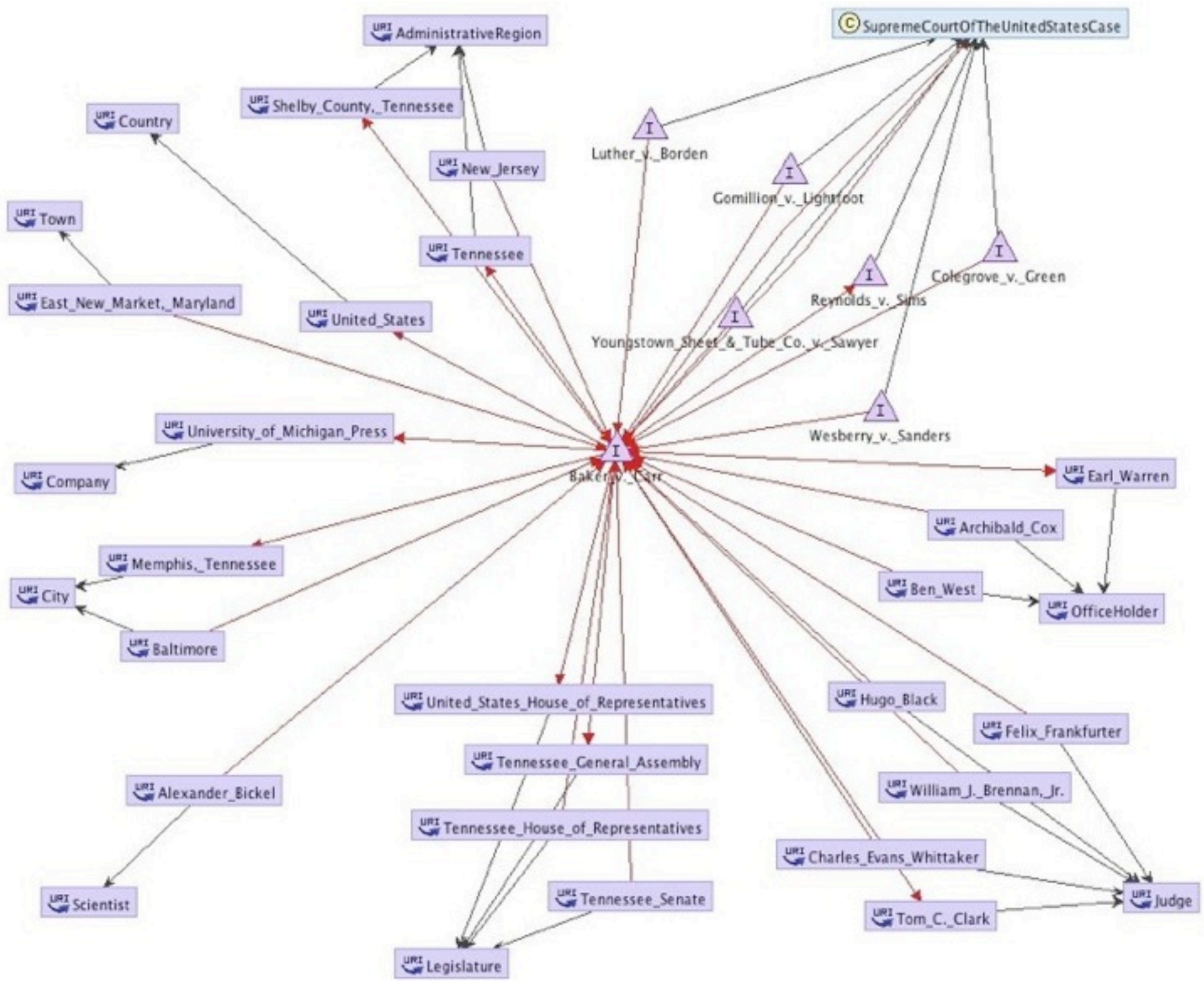
Empirical research: wikipedia link patterns for US supreme court cases

[SupremeCourtOfTheUnitedStatesCase, SupremeCourtOfTheUnitedStatesCase]	45.45
[SupremeCourtOfTheUnitedStatesCase, OfficeHolder]	33.91
[SupremeCourtOfTheUnitedStatesCase, AdministrativeRegion]	33.66
[SupremeCourtOfTheUnitedStatesCase, Judge]	30.40
[SupremeCourtOfTheUnitedStatesCase, City]	15.10
[SupremeCourtOfTheUnitedStatesCase, Country]	13.37
[SupremeCourtOfTheUnitedStatesCase, Legislature]	12.25
[SupremeCourtOfTheUnitedStatesCase, President]	4.88
[SupremeCourtOfTheUnitedStatesCase, Disease]	4.52
[SupremeCourtOfTheUnitedStatesCase, University]	3.86
[SupremeCourtOfTheUnitedStatesCase, EthnicGroup]	3.61
[SupremeCourtOfTheUnitedStatesCase, MilitaryConflict]	3.15
[SupremeCourtOfTheUnitedStatesCase, Senator]	1.88
[SupremeCourtOfTheUnitedStatesCase, Newspaper]	1.83
[SupremeCourtOfTheUnitedStatesCase, Drug]	1.78
[SupremeCourtOfTheUnitedStatesCase, HistoricPlace]	1.73
[SupremeCourtOfTheUnitedStatesCase, Scientist]	1.53
[SupremeCourtOfTheUnitedStatesCase, Governor]	1.53
[SupremeCourtOfTheUnitedStatesCase, School]	1.47
[SupremeCourtOfTheUnitedStatesCase, TradeUnion]	1.32
[SupremeCourtOfTheUnitedStatesCase, River]	1.27
[SupremeCourtOfTheUnitedStatesCase, MilitaryUnit]	1.27
[SupremeCourtOfTheUnitedStatesCase, Congressman]	1.02
[SupremeCourtOfTheUnitedStatesCase, Non-ProfitOrganisation]	0.92
[SupremeCourtOfTheUnitedStatesCase, Magazine]	0.92
[SupremeCourtOfTheUnitedStatesCase, Broadcast]	0.86
[SupremeCourtOfTheUnitedStatesCase, Election]	0.76
[SupremeCourtOfTheUnitedStatesCase, Writer]	0.76
[SupremeCourtOfTheUnitedStatesCase, Website]	0.71
[SupremeCourtOfTheUnitedStatesCase, Town]	0.71
[SupremeCourtOfTheUnitedStatesCase, Book]	0.66
[SupremeCourtOfTheUnitedStatesCase, Film]	0.61
[SupremeCourtOfTheUnitedStatesCase, Language]	0.51

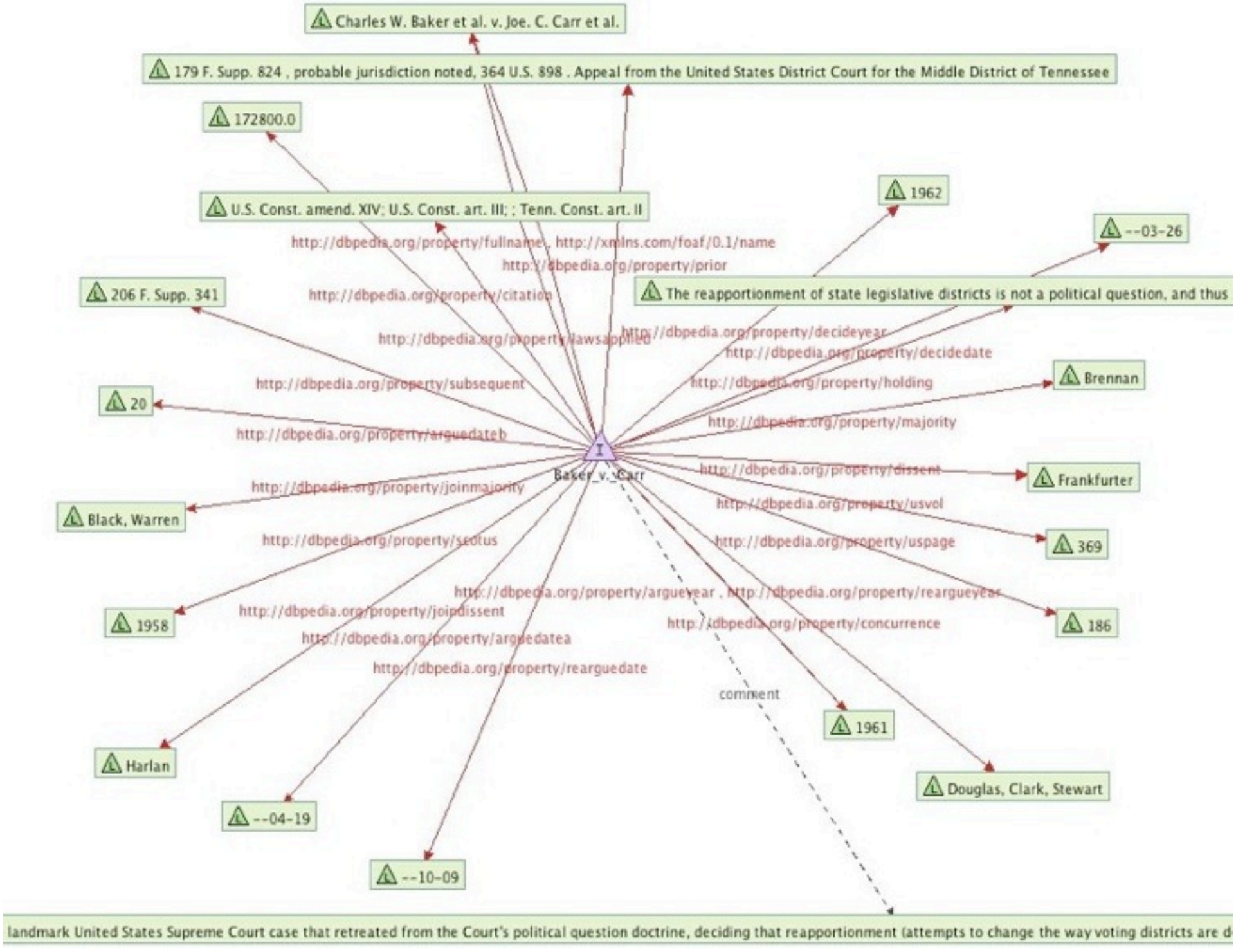
Emerging frame (“encyclopedic knowledge pattern”)



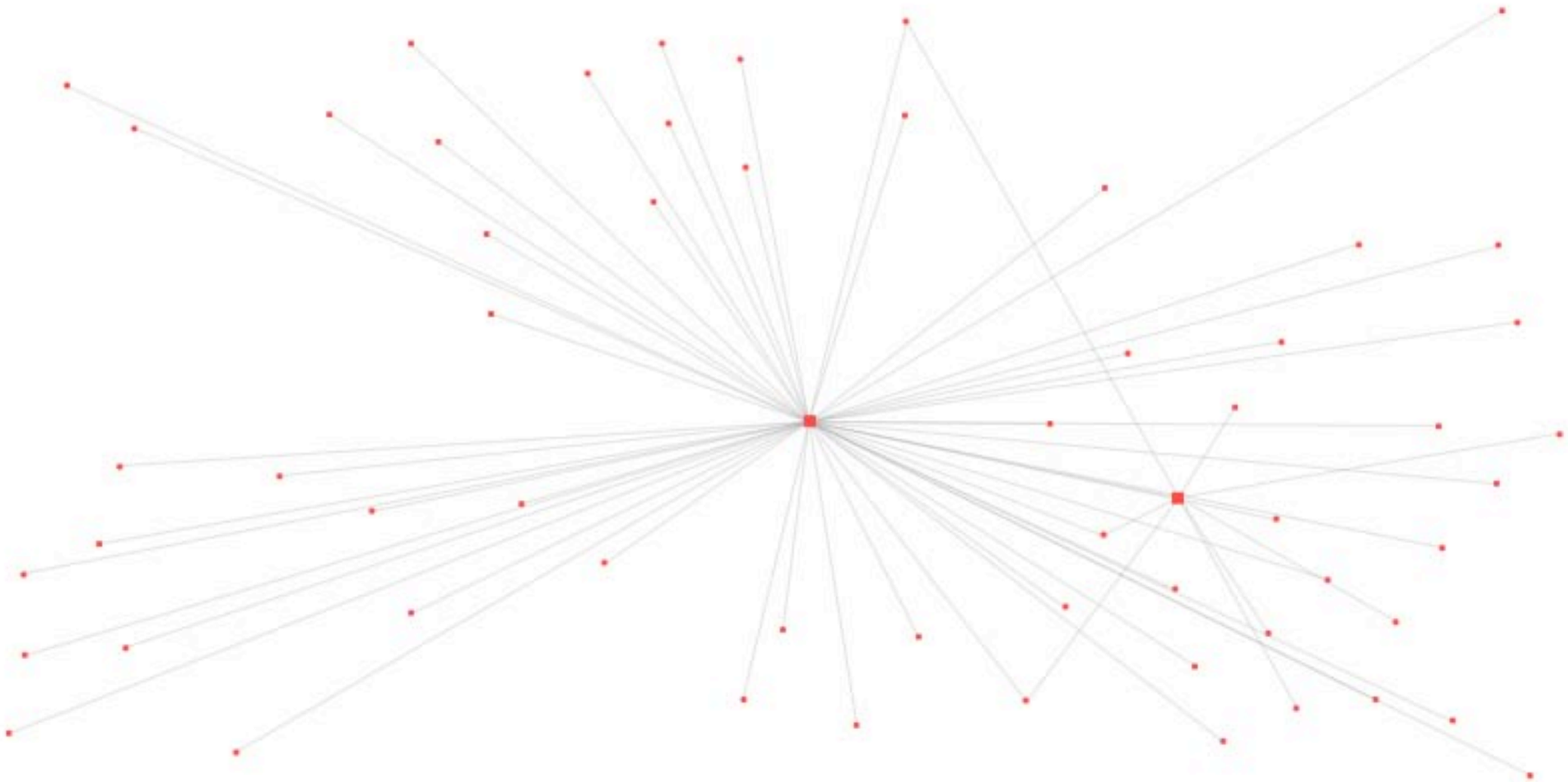
Baker vs. Carr voting district reapportionment case, 1962: conceptualising links



Baker vs. Carr voting district reapportionment case, 1962: conceptualising data



Baker vs. Carr case semantic social network (persons). Hubs are Warren (politician) and Clark (judge)



nodes: 60
edges: 62
drawing: 4 ms
calculation: 0 ms

Baker vs. Carr case semantic social network (persons) at degree 2: hubs include most prominent politicians from the 60s in US

