Our Roadmap

Web 2.0

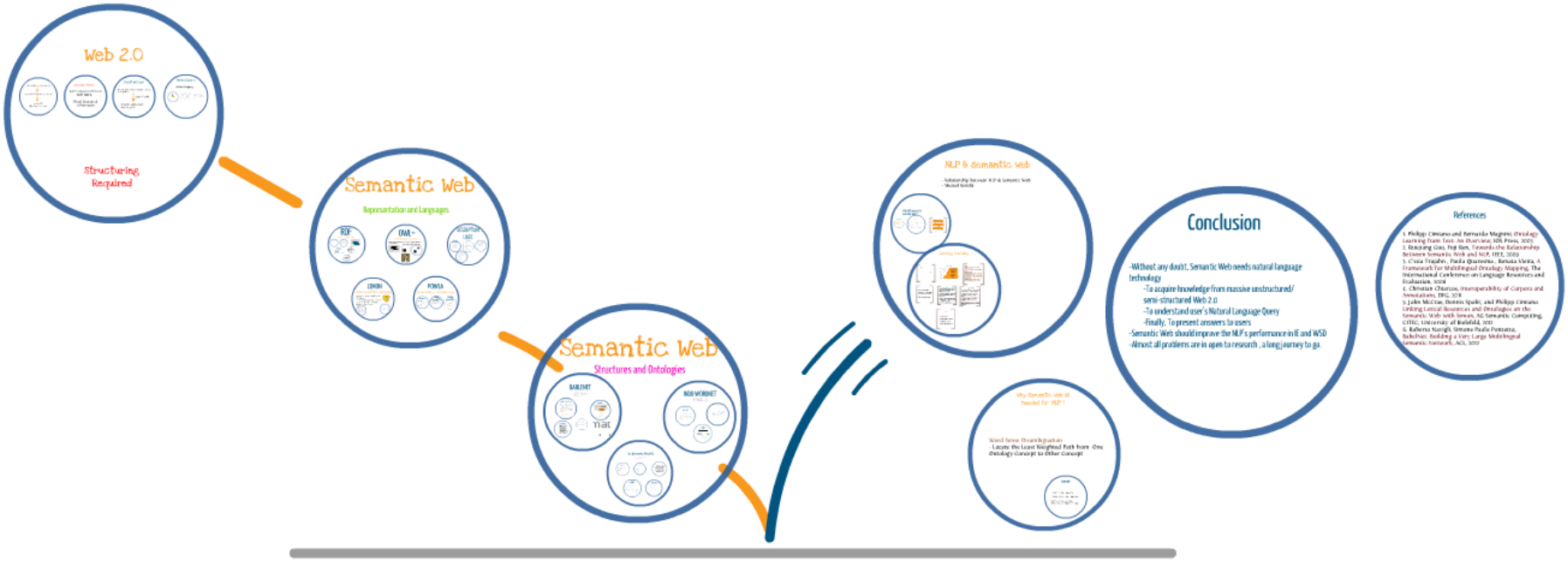
Semantic Web

Representation and Languages

Structures and Ontologies

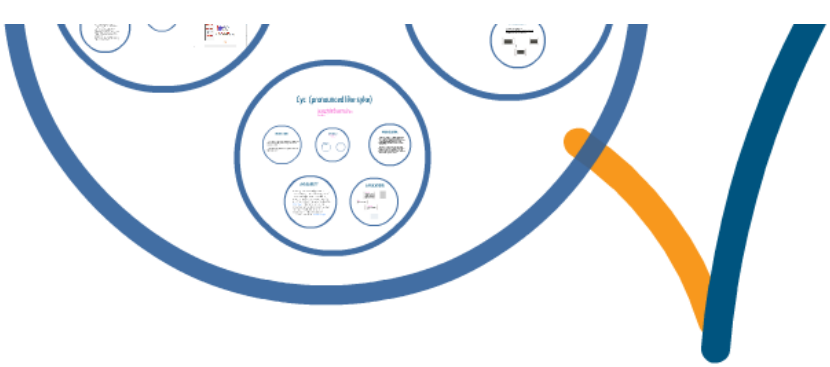
NLP and Semantic Web

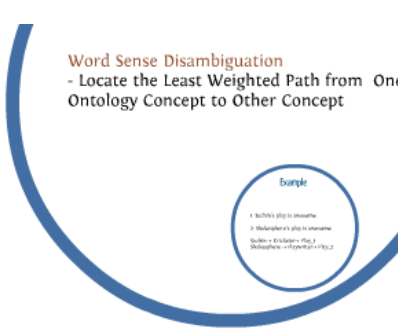
Conclusion



Semantic Web

Making the Web more readable for machines





Semantic Web

Making the Web more readable for machines

Road Map


India Road Network
National and State Highways
N
W
E
S
Srinagar
Uri
Leh
Kishtwar
Jammu
Dalhousie
Pathankot
Atari
Jalandhar
Kalpa
Shimla
Chandigarh
Badrinath
Ambala
Dehradun
Hisar
Meerut
Bikaner
DELHI
Gurgaon
Bareilly
Agra
Jaisalmer
Jodhpur
Jaipur
Lucknow
Gonda
Gorakhpur
Barmer
Ajmer
Gwalior
Kanpur
Varnasi
Shivpuri
Allahabad
Kota
Mount Abu
Udaipur
Mangawan
Radhanpur
Bhuj
Kandla
Ahmedabad
Bhopal
Jabalpur
Indore
Porbandar
Rajkot
Vadodara
Surat
Dhule
Nagpur
Raipur
Gir
Akola
Nasik
Bhavanipatna
Adilabad
Mumbai
Nanded
Jagdalpur
Pune
Nizamabad
Warangal
Sholapur
Ratnigiri
Gulbarga
Hyderabad
Kolhapur
Bijapur
Machilipatnam
Belgaum
Kurnool
Panaji
Hospet
Gooty
Chitradurga
Nellore
Thirupati
Bangalore
Chennai
Mangalore
Mysore
Kannur
Pondicherry
Salem
Coimbatore
Tanjavoor
Thrissur
Dindigul
Kochi
Madurai
Rameswaram
Thiruvananthapuram
Palayamkottai
Kanyakumari
Ziro
Along
Bomdila
Dibrugarh
Jorhat
Gangtok
Guwahati
Motihari
Dalkota
Kooch Bihar
Kohima
Patna
Bhagalpur
Imphal
Ingraj Bazar
Agartala
Aizwal
Ambikapur
Haora
Bangaon
Raurkela
Kolkata
Sambalpur
Baleswar
Cuttack
Paradweep
Bhubaneswar
Puri
Koraput
Bhiminipuram
Vishakhapatanam
BAY OF BENGAL
ARABIAN SEA
INDIA OCEAN
National Highways with Number
Other Main Road Network
Major Cities
Map Not To Scale
© Probharat.com

Our Roadmap

Web 2.0

Semantic Web

- Representation and Languages
- Structures and Ontologies

NLP and Semantic Web

Conclusion

Web 2.0

Type something on a search engine

↓

It just matches the keywords in your query

↓

Lot of work !!!!

- 1 million hits on an avg query

Limitation of HTML

Used to represent information

NOT DATA

Visual Structured Information

Crux of the Problem

Machine didn't understand what the user actually wants

↓ even if it did

It Didn't understand how to get it

Example Scenario

Online Shopping

Structuring Required

Type something on a search engine

It just matches the keywords in your query

Lot of work !!!!

~ 1 million hits on an avg query

Limitation of HTML

Used to represent information

NOT DATA

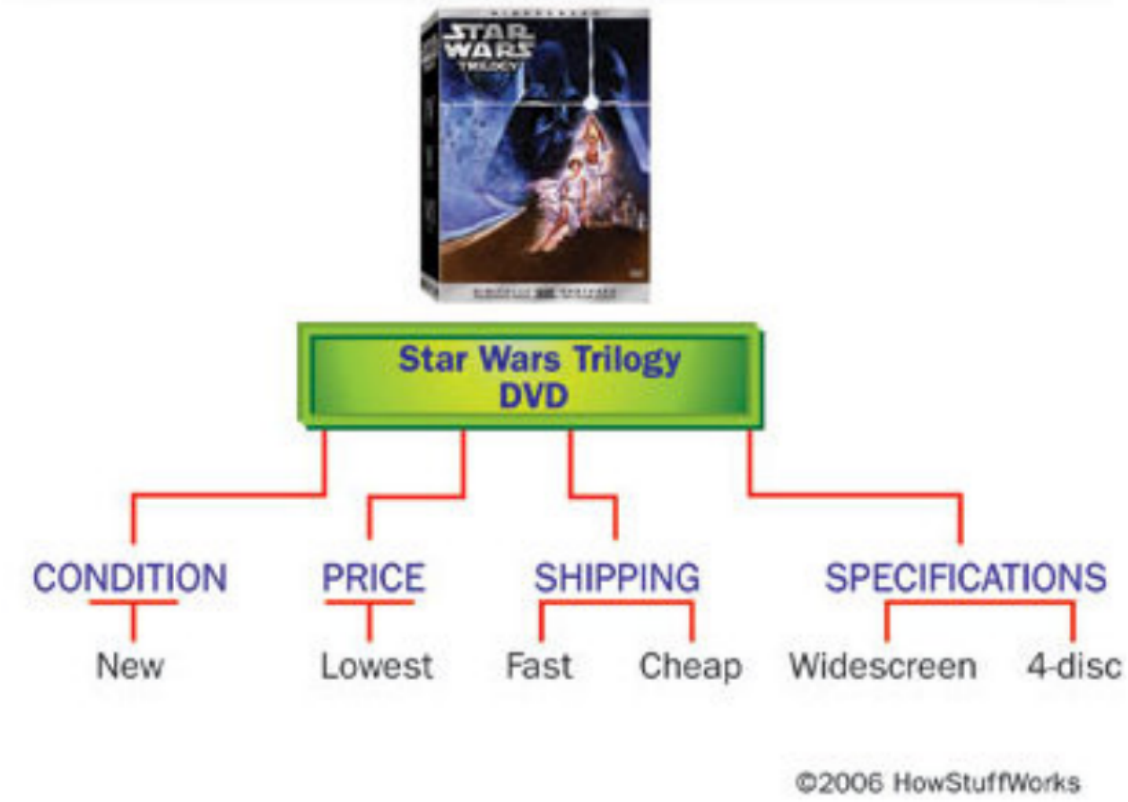
Visual Structured Information

Example Scenario

Online Shopping

At this point in the evolution of the Web, your best bet would be to look at different retailers' web pages, comparing prices and shipping times and rates. You could also look for a site that will compare prices and shipping options from several retailers all at once. Either way, you have to do most of the virtual legwork, then make your buying decision and place your order yourself.

With the Semantic Web, you'd have another option. You could enter your preferences into a computerized agent, which would search the Web, find the best option for you, and place your order.


STAR WARS
TRILOGY
Star Wars Trilogy DVD
CONDITION
PRICE
SHIPPING
SPECIFICATIONS
New
Lowest
Fast
Cheap
Widescreen
4-disc
©2006 HowStuffWorks

At this point in the evolution of the Web, your best bet would be to look at different retailers' web pages, comparing prices and shipping times and rates. You could also look for a site that will compare prices and shipping options from several retailers all at once. Either way, you have to do most of the virtual legwork, then make your buying decision and place your order yourself.

With the Semantic Web, you'd have another option. You could enter your preferences into a computerized agent, which would search the Web, find the best option for you, and place your order.

Crux of the Problem

Machine didn't understand what the user actually wants

even if it did

It Didn't understand how to get it

Structuring Required

Semantic web

Representation and Languages

RDF

Resource Description Format

OWL (WOL)

Web Ontology Language

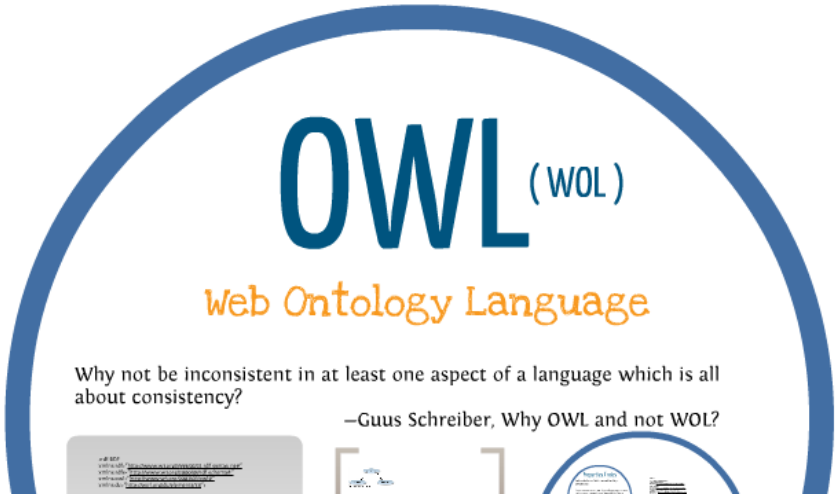
Why not be inconsistent in at least one aspect of a language which is all about consistency?

—Guus Schreiber, Why OWL and not WOL?

DESCRIPTION LOGIC

Semantic We

Representation and Languages



RDF

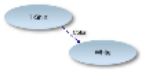
ReSource DeScription Format

Graph Database

Syntax

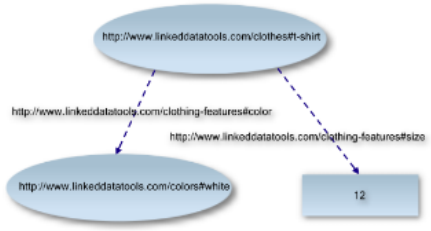
Defines data in the form of a <Subject,Predicate,Object> triple

```
<rdf:Description rdf:about="subject">
    <predicate rdf:resource="object" />
    <predicate>literal value</predicate>
<rdf:Description>
```



Example

```
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
    xmlns:feature="http://www.linkeddatatools.com/clothing-features#">
    <rdf:Description rdf:about="http://www.linkeddatatools.com/clothes#t-shirt">
        <feature:size>12</feature:size>
        <feature:color rdf:resource="http://www.linkeddatatools.com/colors#white"/>
    </rdf:Description>
</rdf:RDF>
```



Reification

Involves representation of factual assertions that are represented by some other assertions

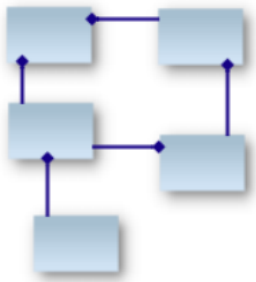
To compare logical assertions from different witnesses in order to determine their credibility

Drawback

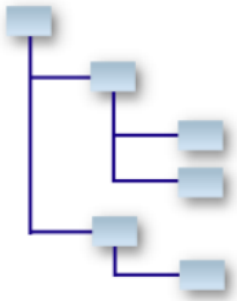
RDF, whilst the foundation of defining data structures for the semantic web, does not in itself describe the semantics, or meaning, behind the data

for that we need
Schema or Ontology

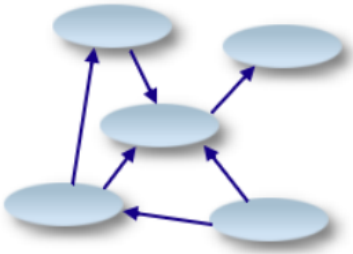
Graph Database



Relational DB
Tables Related By
Primary Key



Hierarchical DB
Parent Nodes Have More
Intrinsic Importance

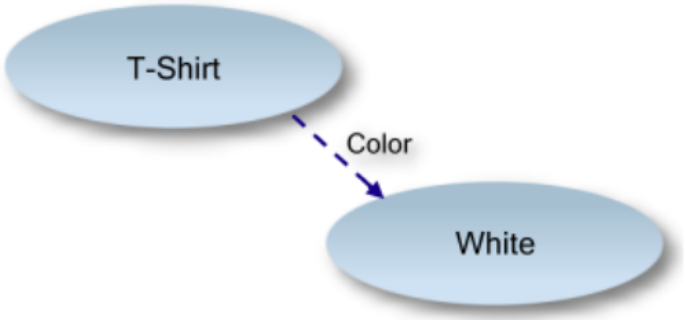


Graph DB
Arbitrary Object Relations
No Intrinsic Importance

Syntax

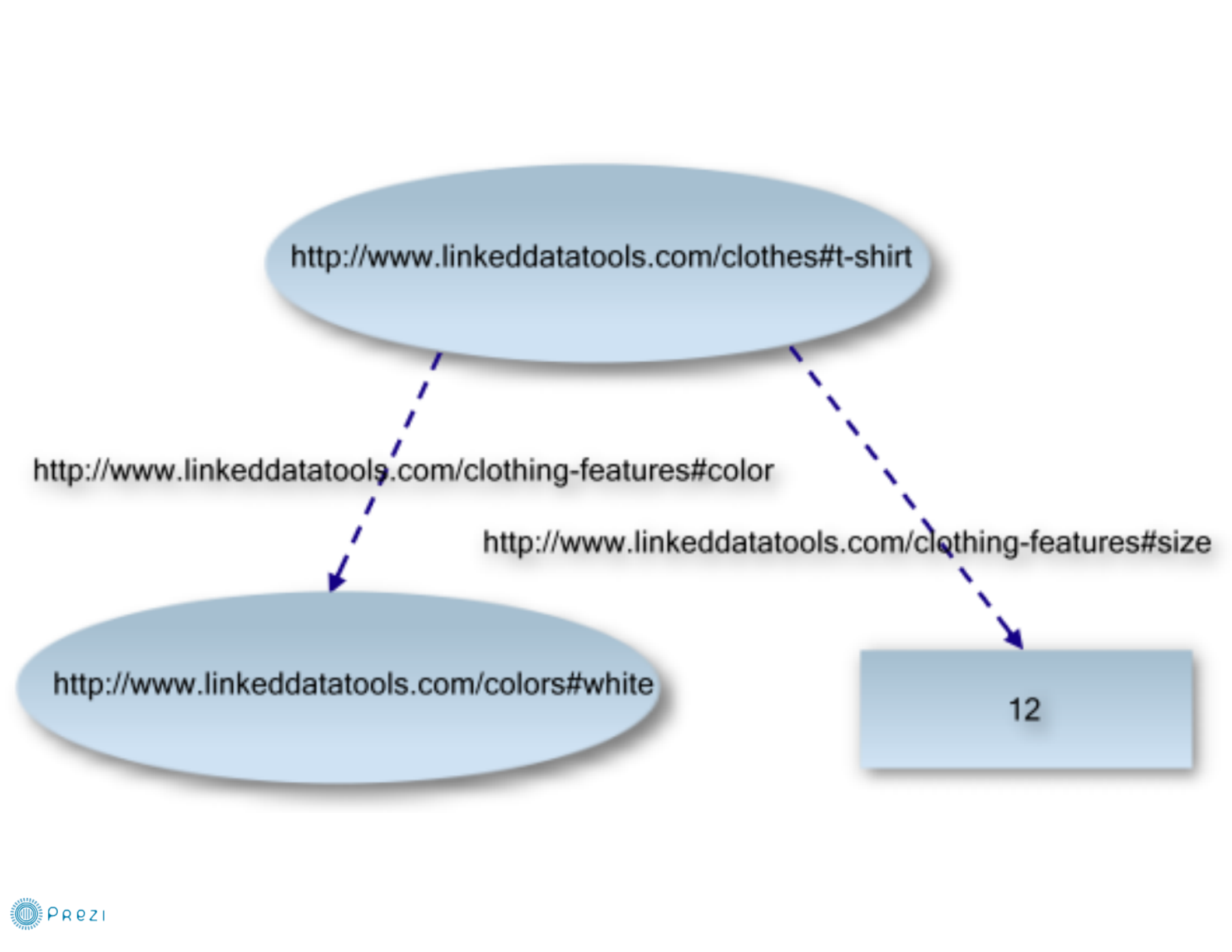
Defines data in the form of a <Subject,Predicate,Object> triple

```
<rdf:Description rdf:about="subject">
      <predicate rdf:resource="object" />
      <predicate>literal value</predicate>
<rdf:Description>
```



Example

```
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
      xmlns:feature="http://www.linkeddatatools.com/clothing-features#">
      <rdf:Description rdf:about="http://www.linkeddatatools.com/clothes#t-shirt">
            <feature:size>12</feature:size>
            <feature:color rdf:resource="http://www.linkeddatatools.com/colors#white"/>
      </rdf:Description>
</rdf:RDF>
```


http://www.linkeddatatools.com/clothes#t-shirt
http://www.linkeddatatools.com/clothing-features#color
http://www.linkeddatatools.com/clothing-features#size
http://www.linkeddatatools.com/colors#white
12

Reification

Involves representation of factual assertions that are representated by some other assertions

To compare logical assertions from different witnesses in order to determine their credibility

Drawback

RDF, whilst the foundation of defining data structures for the semantic web, does not in itself describe the semantics, or meaning, behind the data

for that we need
Schema or Ontology

DESCRIPTION LOGIC

Motivation

Conceptual Models are required in

- Artificial Intelligence
- Database Design
- Software Engineering
- Information Integration

Fundamental Ontology

Conceptual model is populated by
- Individuals
- related by binary relationships (called roles & features)
- grouped into classes (concepts)

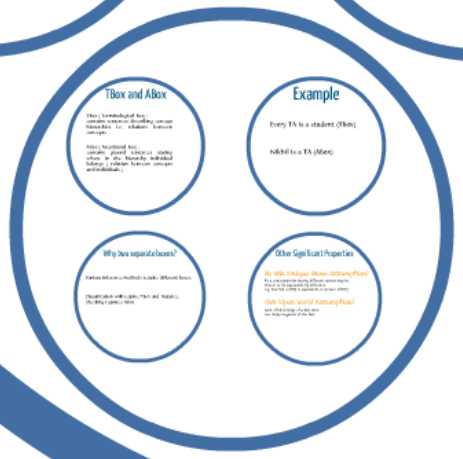
So we need ability to describe concepts, relationships and individuals

FOL cannot be used as it is not decidable

Description logic is characterized by a set of constructors that allow to build complex concepts and roles from atomic ones.

concept - class / set of objects

role - binary relations on objects



Motivation

Conceptual Models are required in

- Artificial Intelligence
- Database Design
- Software Engineering
- Information Integration

Fundamental Ontology

Conceptual model is populated by

- Individuals
- related by binary relationships (called roles & features)
- grouped into classes (concepts)

So we need ability to describe concepts, relationships and individuals

FOL cannot be used as it is not decidable

Description logic is characterized by a set of constructors that allow to build complex concepts and roles from atomic ones.

concept - class / set of objects

role - binary relations on objects

TBox and ABox

TBox (Terminological Box) :
contains sentences describing concept hierarchies i.e; relations between concepts

ABox (Assertional Box) :
contains ground sentences stating where in the hierarchy individual belongs (relation between concepts and individuals)

Example

Every TA is a student (TBox)

Nikhil is a TA (ABox)

Why two separate boxes?

Various inference methods require different boxes

Classification will require TBox and Instance checking requires ABox.

Other Significant Properties

No UNA (Unique Name Assumption)

Two concepts/roles having different names may be shown to be equivalent by inference
e.g; married_to(X,Y) is equivalent to spouse_of(X,Y)

OWA (Open World Assumption)

Lack of knowledge of a fact does not imply negation of the fact

TBox and ABox

TBox (Terminological Box) :
contains sentences describing concept hierarchies i.e; relations between concepts

ABox (Assertional Box) :
contains ground sentences stating where in the hierarchy individual belongs (relation between concepts and individuals)

Example

Every TA is a student (TBox)

Nikhil is a TA (ABox)

Why two separate boxes?

Various inference methods require different boxes

Classification will require TBox and Instance checking requires ABox.

Other Significant Properties

No UNA (Unique Name Assumption)

Two concepts/roles having different names may be shown to be equivalent by inference
e.g; married_to(X,Y) is equivalent to spouse_of(X,Y)

OWA (Open World Assumption)

Lack of knowledge of a fact does not imply negation of the fact

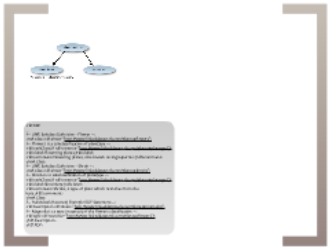
OWL (WOL)

Web Ontology Language

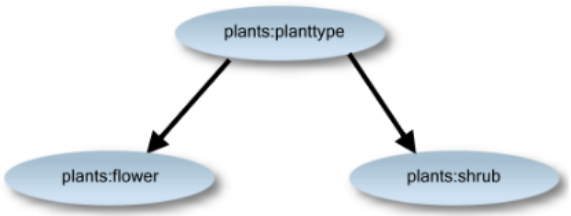
Why not be inconsistent in at least one aspect of a language which is all about consistency?

—Guus Schreiber, Why OWL and not WOL?

```
<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#"
xmlns:dc="http://purl.org/dc/elements/1.1/">

<!-- OWL Header Example -->
<owl:Ontology rdf:about="http://www.linkeddatatools.com/plants">
<dc:title>The LinkedDataTools.com Example Plant Ontology</dc:title>
<dc:description>An example ontology written for the
LinkedDataTools.com RDFS & OWL introduction
tutorial</dc:description>
</owl:Ontology>

<!-- Remainder Of Document Omitted For Brevity... -->

</rdf:RDF>
```

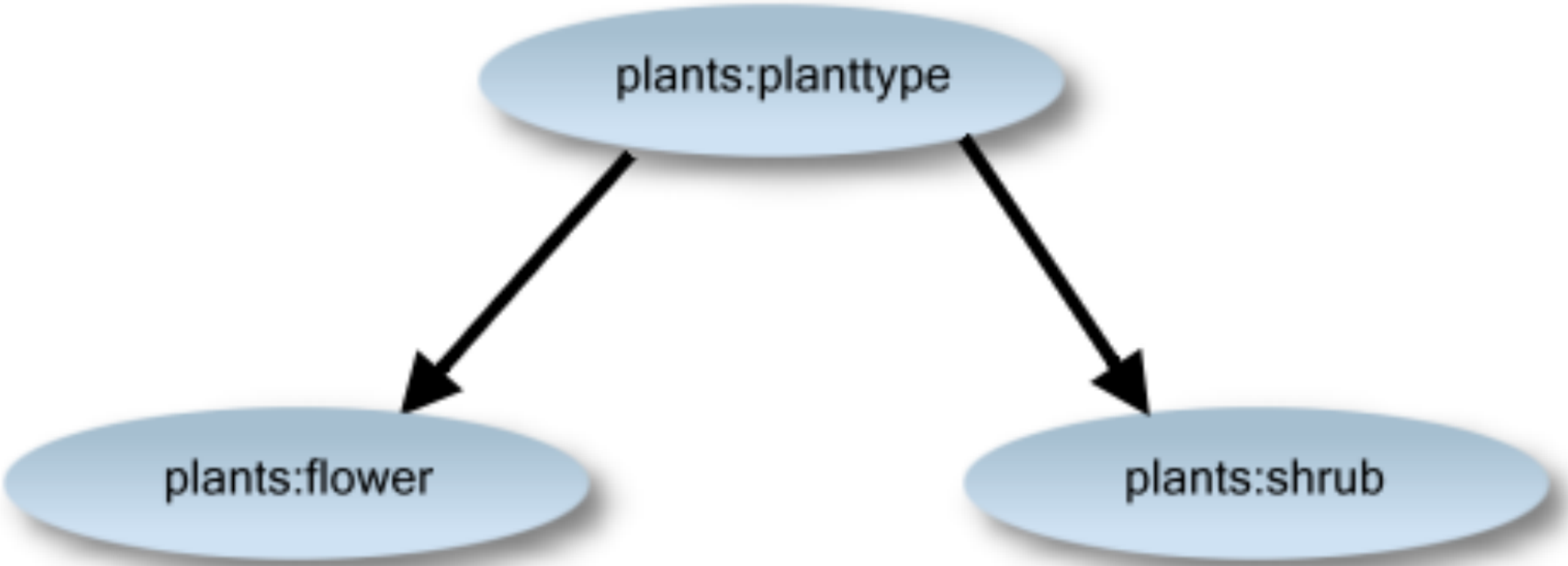


```
<rdf:RDF
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
xmlns:owl="http://www.w3.org/2002/07/owl#"
xmlns:dc="http://purl.org/dc/elements/1.1/">

<!-- OWL Header Example -->
<owl:Ontology rdf:about="http://www.linkeddatatools.com/plants">
<dc:title>The LinkedDataTools.com Example Plant Ontology</dc:title>
<dc:description>An example ontology written for the
LinkedDataTools.com RDFS & OWL introduction
tutorial</dc:description>
</owl:Ontology>

<!-- Remainder Of Document Omitted For Brevity... -->

</rdf:RDF>
```



An example of a **taxonomy** hierarchy

```
<rdf:RDF
...
<!-- OWL Subclass Definition - Flower -->
<owl:Class rdf:about="http://www.linkeddatatools.com/plants#flowers">
<!-- Flowers is a subclassification of planttype -->
<rdfs:subClassOf rdf:resource="http://www.linkeddatatools.com/plants#planttype"/>
<rdfs:label>Flowering plants</rdfs:label>
<rdfs:comment>Flowering plants, also known as angiosperms.</rdfs:comment>
</owl:Class>
<!-- OWL Subclass Definition - Shrub -->
<owl:Class rdf:about="http://www.linkeddatatools.com/plants#shrubs">
<!-- Shrubs is a subclassification of planttype -->
<rdfs:subClassOf rdf:resource="http://www.linkeddatatools.com/plants#planttype"/>
<rdfs:label>Shrubbery</rdfs:label>
<rdfs:comment>Shrubs, a type of plant which branches from the
base.</rdfs:comment>
</owl:Class>
<!-- Individual (Instance) Example RDF Statement -->
<rdf:Description rdf:about="http://www.linkeddatatools.com/plants#magnolia">
<!-- Magnolia is a type (instance) of the flowers classification -->
<rdf:type rdf:resource="http://www.linkeddatatools.com/plants#flowers"/>
</rdf:Description>
</rdf:RDF>
```



An example of a **taxonomy** hierarchy

```
‹rdf:RDF
...
‹!-- OWL Subclass Definition - Flower --›
‹owl:Class rdf:about="http://www.linkeddatatools.com/plants#flowers"›
‹!-- Flowers is a subclassification of planttype --›
‹rdfs:subClassOf rdf:resource="http://www.linkeddatatools.com/plants#planttype"/›
‹rdfs:label›Flowering plants‹/rdfs:label›
‹rdfs:comment›Flowering plants, also known as angiosperms.‹/rdfs:comment›
‹/owl:Class›
‹!-- OWL Subclass Definition - Shrub --›
‹owl:Class rdf:about="http://www.linkeddatatools.com/plants#shrubs"›
‹!-- Shrubs is a subclassification of planttype --›
‹rdfs:subClassOf rdf:resource="http://www.linkeddatatools.com/plants#planttype"/›
‹rdfs:label›Shrubbery‹/rdfs:label›
‹rdfs:comment›Shrubs, a type of plant which branches from the
base.‹/rdfs:comment›
‹/owl:Class›
‹!-- Individual (Instance) Example RDF Statement --›
‹rdf:Description rdf:about="http://www.linkeddatatools.com/plants#magnolia"›
‹!-- Magnolia is a type (instance) of the flowers classification --›
‹rdf:type rdf:resource="http://www.linkeddatatools.com/plants#flowers"/›
‹/rdf:Description›
‹/rdf:RDF›
```

Properties / roles

Individuals in OWL are related by properties

Object properties (owl:ObjectProperty) relates individuals (instances) of two OWL classes.

Datatype properties (owl:DatatypeProperty) relates individuals (instances) of OWL classes to literal values.

```
<rdf:RDF
...
<!-- Define the family property -->
<owl:DatatypeProperty
rdf:about="http://www.linkeddatatools.com/plants#family"/>

<rdf:Description
rdf:about="http://www.linkeddatatools.com/plants#magnolia">
<!-- Magnolia is a type (instance) of the flowers class -->
<rdf:type
rdf:resource="http://www.linkeddatatools.com/plants#flowers"/>
<!-- The magnolia is part of the 'Magnoliaceae' family -->
<plants:family>Magnoliaceae</plants:family>
</rdf:Description>

</rdf:RDF>
```

```
<rdf:RDF
<!-- Define the family property -->
<owl:DatatypeProperty rdf:about="http://www.linkeddatatools.com/plants#family"/>
<!-- Define the similarlyPopularTo property -->
<owl:ObjectProperty rdf:about="http://www.linkeddatatools.com/plants#similarlyPopularTo"/>
<!-- Define the Orchid class instance -->
<rdf:Description rdf:about="http://www.linkeddatatools.com/plants#orchid">
...
</rdf:Description>

<!-- Define the Magnolia class instance -->
<rdf:Description rdf:about="http://www.linkeddatatools.com/plants#magnolia">
<!-- Magnolia is an individual (instance) of the flowers class -->
<rdf:type rdf:resource="http://www.linkeddatatools.com/plants#flowers"/>
<!-- The magnolia is part of the 'Magnoliaceae' family -->
<plants:family>Magnoliaceae</plants:family>
<!-- The magnolia is similarly popular to the orchid -->
<plants:similarlyPopularTo rdf:resource="http://www.linkeddatatools.com/plants#orchid"/>
</rdf:Description>

</rdf:RDF>
```

Properties / roles

Individuals in OWL are related by properties

Object properties (owl:ObjectProperty) relates individuals (instances) of two OWL classes.

Datatype properties (owl:DatatypeProperty) relates individuals (instances) of OWL classes to literal values.

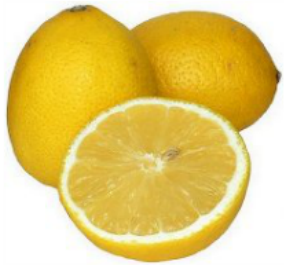
```
‹rdf:RDF
...
‹!-- Define the family property --›
‹owl:DatatypeProperty
rdf:about="http://www.linkeddatatools.com/plants#family"/›

‹rdf:Description
rdf:about="http://www.linkeddatatools.com/plants#magnolia"›
‹!-- Magnolia is a type (instance) of the flowers class --›
‹rdf:type
rdf:resource="http://www.linkeddatatools.com/plants#flowers"/›
‹!-- The magnolia is part of the 'Magnoliaceae' family --›
‹plants:family›Magnoliaceae‹/plants:family›
‹/rdf:Description›

‹/rdf:RDF›
```

```
‹rdf:RDF
‹!-- Define the family property --›
‹owl:DatatypeProperty rdf:about="http://www.linkeddatatools.com/plants#family"/›
‹!-- Define the similarlyPopularTo property --›
‹owl:ObjectProperty rdf:about="http://www.linkeddatatools.com/plants#similarlyPopularTo"/›
‹!-- Define the Orchid class instance --›
‹rdf:Description rdf:about="http://www.linkeddatatools.com/plants#orchid"›
...
‹/rdf:Description›

‹!-- Define the Magnolia class instance --›
‹rdf:Description rdf:about="http://www.linkeddatatools.com/plants#magnolia"›
‹!-- Magnolia is an individual (instance) of the flowers class --›
‹rdf:type rdf:resource="http://www.linkeddatatools.com/plants#flowers"/›
‹!-- The magnolia is part of the 'Magnoliaceae' family --›
‹plants:family›Magnoliaceae‹/plants:family›
‹!-- The magnolia is similarly popular to the orchid --›
‹plants:similarlyPopularTo rdf:resource="http://www.linkeddatatools.com/plants#orchid"/›
‹/rdf:Description›

‹/rdf:RDF›
```

LEMON

Lexicon model for ontologies

- Separate Lexicon models and Ontologies
- Linking them can aid in lot of NLP applications like Q/A, Machine Translation



Lexicon is the vocabulary of the language

Uses RDF for defining links

Meaning of a word given by reference

Reference (Ontology) capable of representing more complex semantic information

Lemon model provides a principled chain between the semantic representation and its linguistic realization

Lexicon is the vocabulary of the language

Uses RDF for defining links

Meaning of a word given by reference

Reference (Ontology) capable of representing more complex semantic information

Lemon model provides a principled chain between the semantic representation and its linguistic realization

POWLA

Interoperability of Linguistic Corpora

Motivation

Same text, Different annotations
POS tags, Parse trees, etc

A processing stage may require different annotations to the same text

Common representation which provides access to all the linguistic information conveyed in the annotations

Structural Interoperatibility

Use RDF to represent all the annotations of the corpus in an interoperable way, integrate their information without restrictions and query the information

Use OWL/DL to specify and verify formal constraints on the correct representation of linguistic corpora in RDF

Conceptual Interoperatibility

Heterogenous annotation schemes

↓

Terminological Reference repository

Provide interlingua that allow mapping form scheme A to scheme B

GOLD

Motivation

Same text, Different annotations

POS tags, Parse trees, etc

A processing stage may require different annotations to the same text

Common representation which provides access to all the linguistic information conveyed in the annotations

Structural Interoperatibility

Use RDF to represent all the annotations of the corpus in an interoperable way, integrate their information without restrictions and query the information

Use OWL/DL to specify and verify formal constraints on the correct representation of linguistic corpora in RDF

Conceptual Interoperatibility

Heterogenous annotation schemes

↓

Terminological Reference repository

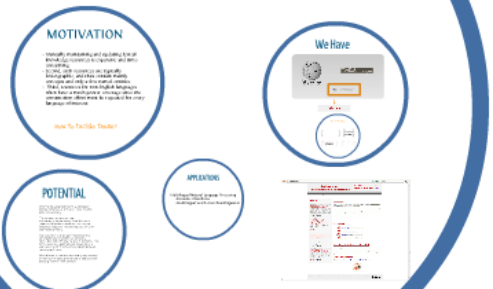
Provide interlingua that allow mapping form scheme A to scheme B

GOLD

Semantic web

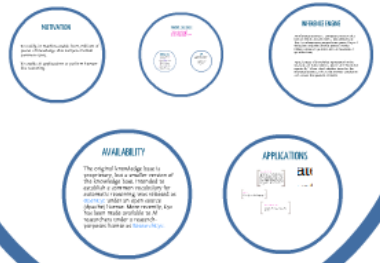
Structures and Ontologies

BABLENET



INDO WORDNET

Cyc (pronounced like syke)



BABLENET

multilingual lexicalized semantic network.
automatically created by linking the largest multilingual Web encyclopedia - i.e., Wikipedia - to the most popular computational lexicon of the English language - i.e., Wordnet

MOTIVATION

- Manually maintaining and updating lexical knowledge resources is expensive and time-consuming.
- Second, such resources are typically lexicographic, and thus contain mainly concepts and only a few named entities.
- Third, resources for non-English languages often have a much poorer coverage since the construction effort must be repeated for every language of interest.

How to tackle these?

We Have

POTENTIAL

Experiments show that this fully-automated approach produces a large-scale lexical resource with high accuracy.

The resource includes millions of semantic relations, mainly from Wikipedia (however, WordNet relations are labeled), and contains almost 3 million concepts (6.7 labels per concept on average).

Such coverage is much wider than that of existing wordnets in non-English languages. While BabelNet currently includes 6 languages, links to freely-available wordnets can immediately be established by utilizing the English WordNet as an interlingua index.

BabelNet can be extended to virtually any language of interest. The translation method allows it to cope with any resource-poor language.

APPLICATIONS

Multilingual Natural Language Processing
-Semantic relatedness
-Multilingual word sense disambiguation

BABLENET

multilingual lexicalized semantic network.
automatically created by linking the largest multilingual Web encyclopedia - i.e., Wikipedia - to the most popular computational lexicon of the English language - i.e., Wordnet

MOTIVATION

DADLINE

multilingual lexicalized semantic network. automatically created by linking the largest multilingual Web encyclopedia - i.e., Wikipedia - to the most popular computational lexicon of the English language - i.e., Wordnet

MOTIVATION

- Manually maintaining and updating lexical knowledge resources is expensive and time-consuming.
- Second, such resources are typically lexicographic, and thus contain mainly concepts and only a few named entities.
- Third, resources for non-English languages often have a much poorer coverage since the construction effort must be repeated for every language of interest.

How to tackle these?

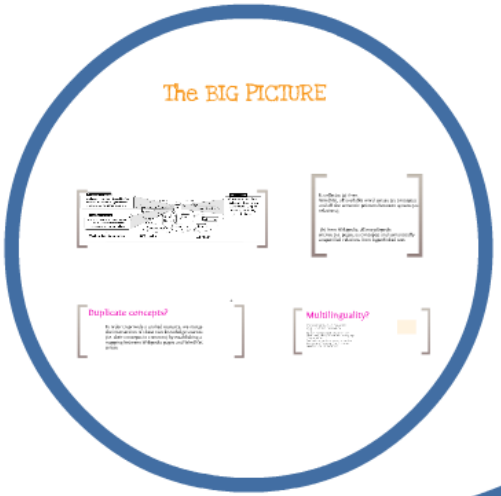
We Have

WIKIPEDIA
The Free Encyclopedia

WordNet
A lexical database for English

Machine Translation

BabelNet
A very large multilingual semantic network



we have

WIKIPEDIA
The Free Encyclopedia

WordNet
A lexical database for English

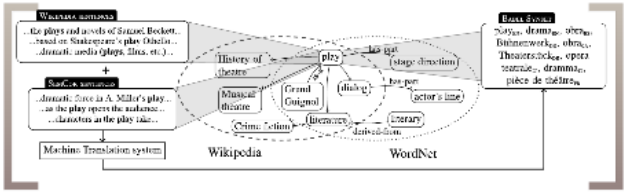
Machine Translation

BabelNet
A very large multilingual semantic network

BabelNet

A very large multilingual semantic network

The BIG PICTURE



It collectss (a) from
WordNet, all available word senses (as concepts) and all the semantic pointers between synsets (as relations);

(b) from Wikipedia, all encyclopedic entries (i.e. pages, as concepts) and semantically unspecified relations from hyperlinked text.

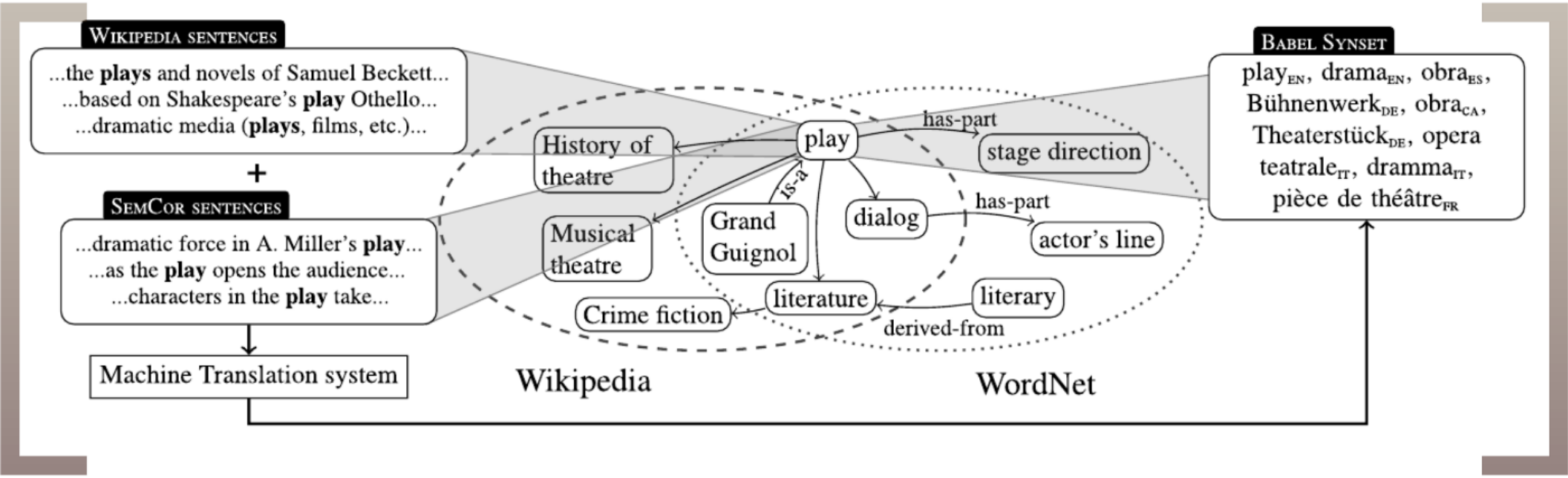
Duplicate concepts?

In order to provide a unified resource, we merge the intersection of these two knowledge sources (i.e. their concepts in common) by establishing a mapping between Wikipedia pages and WordNet senses

+

Multilinguality?

The lexical realizations of the available concepts in different languages by using
(a) the human-generated translations provided in Wikipedia (the so-called inter-language links), as well as
(b) a machine translation system to translate occurrences of the concepts within sense-tagged corpora, namely SemCor


WIKIPEDIA SENTENCES
...the **plays** and novels of Samuel Beckett...
...based on Shakespeare's **play** Othello...
...dramatic media (**plays**, films, etc.)...
+
SEMCOR SENTENCES
...dramatic force in A. Miller's **play**...
...as the **play** opens the audience...
...characters in the **play** take...
Machine Translation system
History of theatre
Musical theatre
Grand Guignol
Crime fiction
play
is-a
has-part
stage direction
dialog
has-part
actor's line
literature
literary
derived-from
Wikipedia
WordNet
BABEL SYNSET
play$_{EN}$, drama$_{EN}$, obra$_{ES}$, Bühnenwerk$_{DE}$, obra$_{CA}$, Theaterstück$_{DE}$, opera teatrale$_{IT}$, dramma$_{IT}$, pièce de théâtre$_{FR}$

It collectss (a) from
WordNet, all available word senses (as concepts) and all the semantic pointers between synsets (as relations);

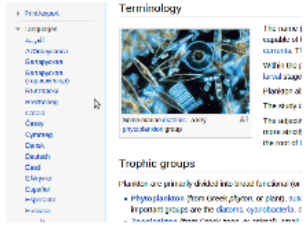
(b) from Wikipedia, all encyclopedic entries (i.e. pages, as concepts) and semantically unspecified relations from hyperlinked text.

Duplicate concepts?

In order to provide a unified resource, we merge the intersection of these two knowledge sources (i.e. their concepts in common) by establishing a mapping between Wikipedia pages and WordNet senses

Multilinguality?

The lexical realizations of the available concepts in different languages by using
(a) the human-generated translations provided in Wikipedia (the so-called inter-language links), as well as
(b) a machine translation system to translate occurrences of the concepts within sense-tagged corpora, namely SemCor



Multilingual

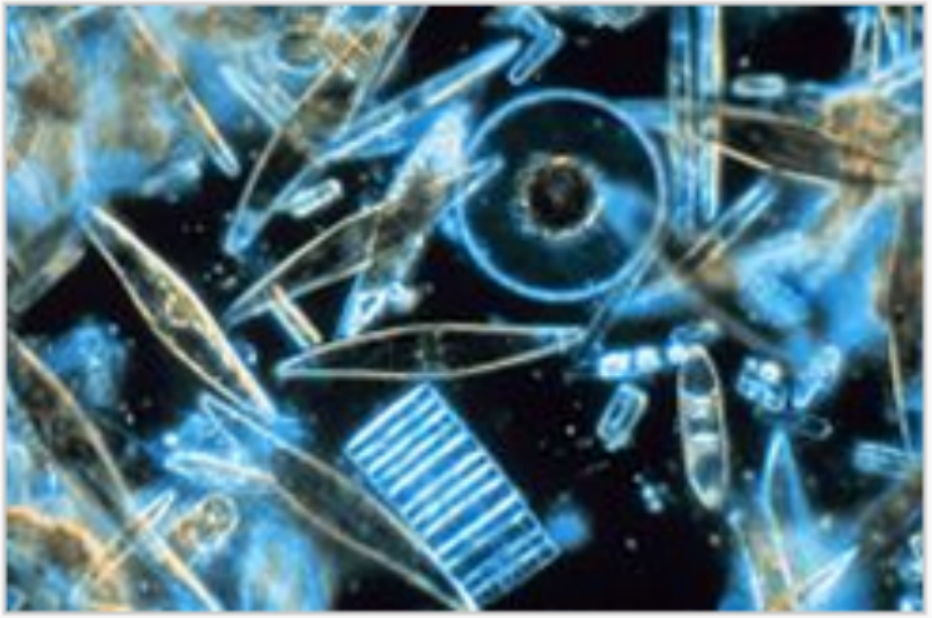
The lexical realizations of the available concepts in different languages by using
(a) the human-generated translations provided in Wikipedia (the so-called inter-language links), as well as
(b) a machine translation system to translate occurrences of the concepts within sense-tagged corpora, namely SemCor

Print/export

Languages

العربية
Azərbaycanca
Беларуская
Беларуская (тарашкевіца)
Български
Brezhoneg
Català
Česky
Cymraeg
Dansk
Deutsch
Eesti
Ελληνικά
Español
Esperanto
Euskara

Terminology



Some marine diatoms - a key phytoplankton group

The name p
capable of i
currents. Th

Within the p
larval stage

Plankton ab

The study o

The adjecti
more strictl
the root of t

Trophic groups

Plankton are primarily divided into broad functional (or

- **Phytoplankton** (from Greek *phyton*, or plant), auto
important groups are the diatoms, cyanobacteria, d
- **Zooplankton** (from Greek *zoon*, or animal), small

Multilingual

The lexical realizations of the available concepts in different languages by using
(a) the human-generated translations provided in Wikipedia (the so-called inter-language links), as well as
(b) a machine translation system to translate occurrences of the concepts within sense-tagged corpora, namely SemCor

<s snum=3>
<wf cmd=ignore pos=DT>The</wf>
<wf cmd=done pos=NN lemma=september wnsn=1
lexsn=1:28:00::>September</wf>
<wf cmd=done pos=NN lemma=october wnsn=1
lexsn=1:28:00::>October</wf>
<wf cmd=done pos=NN lemma=term wnsn=2
lexsn=1:28:00::>term</wf>
<wf cmd=done pos=NN lemma=jury wnsn=1
lexsn=1:14:00::>jury</wf>
<wf cmd=done pos=VBD ot=notag>had</wf>
<wf cmd=done pos=VBN ot=notag>been</wf>
<wf cmd=done pos=VB lemma=charge wnsn=5
lexsn=2:41:00::>charged</wf>
<wf cmd=ignore pos=IN>by</wf>
<wf cmd=done rdf=location pos=NNP
lemma=location wnsn=1 lexsn=1:03:00::
pn=location>Fulton</wf>
<wf cmd=done rdf=person pos=NNP lemma=person
wnsn=1 lexsn=1:03:00::
pn=person>Superior_Court_Judge_Durwood_Pye</w
f>
<wf cmd=ignore pos=TO>to</wf>
<wf cmd=done pos=VB lemma=investigate wnsn=2
lexsn=2:32:01::>investigate</wf>
<wf cmd=done pos=NN lemma=report wnsn=2
lexsn=1:10:00::>reports</wf>
<wf cmd=ignore pos=IN>of</wf>
<wf cmd=done pos=JJ lemma=possible wnsn=2
lexsn=3:00:04::>possible</wf>
<punc>``</punc>
<wf cmd=done pos=NN lemma=irregularity wnsn=1
lexsn=1:04:00::>irregularities</wf>
<punc>''</punc>
<wf cmd=ignore pos=IN>in</wf>
<wf cmd=ignore pos=DT>the</wf>
<wf cmd=done pos=JJ lemma=hard-fought wnsn=1
lexsn=5:00:00:difficult:00>hard-fought</wf>
<wf cmd=done pos=NN lemma=primary wnsn=1
lexsn=1:04:00::>primary</wf>
<wf cmd=ignore pos=WDT>which</wf>
<wf cmd=done pos=VBD ot=notag>was</wf>
<wf cmd=done pos=VB lemma=win wnsn=1
lexsn=2:33:00::>won</wf>
<wf cmd=ignore pos=IN>by</wf>
<wf cmd=done rdf=person pos=NNP lemma=person
wnsn=1 lexsn=1:03:00:: pn=person>Mayor-
nominate_Ivan_Allen_Jr.</wf>
<punc>.</punc>
</s>

POTENTIAL

Experiments show that this fully-automated approach produces a large-scale lexical resource with high accuracy.

The resource includes millions of semantic relations, mainly from Wikipedia (however, WordNet relations are labeled), and contains almost 3 million concepts (6.7 labels per concept on average).

Such coverage is much wider than that of existing wordnets in non-English languages. While BabelNet currently includes 6 languages, links to freely-available wordnets can immediately be established by utilizing the English WordNet as an interlanguage index.

BabelNet can be extended to virtually any language of interest. The translation method allows it to cope with any resource-poor language.

APPLICATIONS

Multilingual Natural Language Processing

-Semantic relatedness

-Multilingual word sense disambiguation.

BabelNet

A very large multilingual semantic network

BabelNetXplorer, an online interface, is available here.

Papers

Roberto Navigli & Simone Paolo Ponzetto
BabelNet: Building a very large multilingual semantic network
In: Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics, Uppsala, Sweden, 11-16 July 2010, pp. 216-225.

PDF BibTeX

Roberto Navigli & Simone Paolo Ponzetto
BabelNet: The Automatic Construction, Evaluation and Application of a Wide-Coverage Multilingual Semantic Network
Artificial Intelligence, 193, Elsevier, 2012, pp. 217-250.

PDF

Roberto Navigli & Simone Paolo Ponzetto
BabelNetXplorer: A Platform for Multilingual Lexical Knowledge Base Access and Exploration
In: Companion Volume to the Proceedings of the 21st World Wide Web Conference, Lyon, France, 16-20 April 2012, pp. 393-396.

PDF BibTeX

Roberto Navigli & Simone Paolo Ponzetto
Multilingual WSD with Just a Few Lines of Code: The BabelNet API
In: Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics, Jeju Island, South Korea, 8-14 July 2012.

PDF BibTeX

Overview

From this page you can download: BabelNet, a **very large multilingual semantic network** with millions of concepts obtained from:

- an integration of WordNet and Wikipedia based on an automatic mapping algorithm
- translations of the concepts (i.e. English Wikipedia pages and WordNet synsets) based on Wikipedia cross-language links and the output of a machine translation system

The resource was **automatically generated using the methods presented in our ACL-10 and AIJ papers** (see "Papers" in the left box).

Download

- NEW BabelNet API version 1.0.1 (August 2012)
- NEW BabelNet precompiled index version 1.0 (July 2012)
- NEW BabelNet dump version 1.0 (July 2012)
- NEW BabelNet glosses version 1.0 (July 2012)

How to get started with BabelNet

The best way to start is to **download the BabelNet API** and follow the instructions found in the README file. You can also download the `raw' data (i.e., the **BabelNet dump** and **glosses** files) and use them with other software (e.g., your favourite API). Details on the APIs and the data format can be found in our demo papers (again, please refer to the papers found in the left box).

Acknowledgments

BabelNet is an output of the ERC Starting Grant MultiJEDI No. 259234.

erc

INDO WORDNET

linked lexical knowledge base of wordnets of 18 scheduled languages of India
Assamese, Bangla, Bodo, Gujarati, Hindi, Kannada, Kashmiri, Konkani, Malayalam, Manipuri, Marathi, Nepali, Oriya, Punjabi, Sanskrit, Tamil, Telugu and Urdu.

MOTIVATION

Indian languages form a very significant component of the languages landscape of the world.

Many languages rank within top 10 in the world in terms of the population speaking them, e.g., Hindi-Urdu 5th, Bangla 7th, Marathi 12th and so on as per the List of languages by number of native speakers.

Indowordnet project started with creations of Hindi WordNet by the Natural Language Processing group at the Center for Indian Language Technology (CFILT) in the Computer Science and Engineering Department at IIT Bombay.

AVAILABILITY

Publicaly browsable at
http://www.cfilt.iitb.ac.in/indowordnet
Available under GNU licence

INDO WORDNET

linked lexical knowledge base of wordnets of 18 scheduled languages of India

Assamese, Bangla, Bodo, Gujarati, Hindi, Kannada, Kashmiri, Konkani, Malayalam, Manipuri, Marathi, Nepali, Oriya, Punjabi, Sanskrit, Tamil, Telugu and Urdu.

WORD

linked lexical knowledge base of wordnets of 18 scheduled languages of India

Assamese, Bangla, Bodo, Gujarati, Hindi, Kannada, Kashmiri, Konkani, Malayalam, Manipuri, Marathi, Nepali, Oriya, Punjabi, Sanskrit, Tamil, Telugu and Urdu.

MOTIVATION

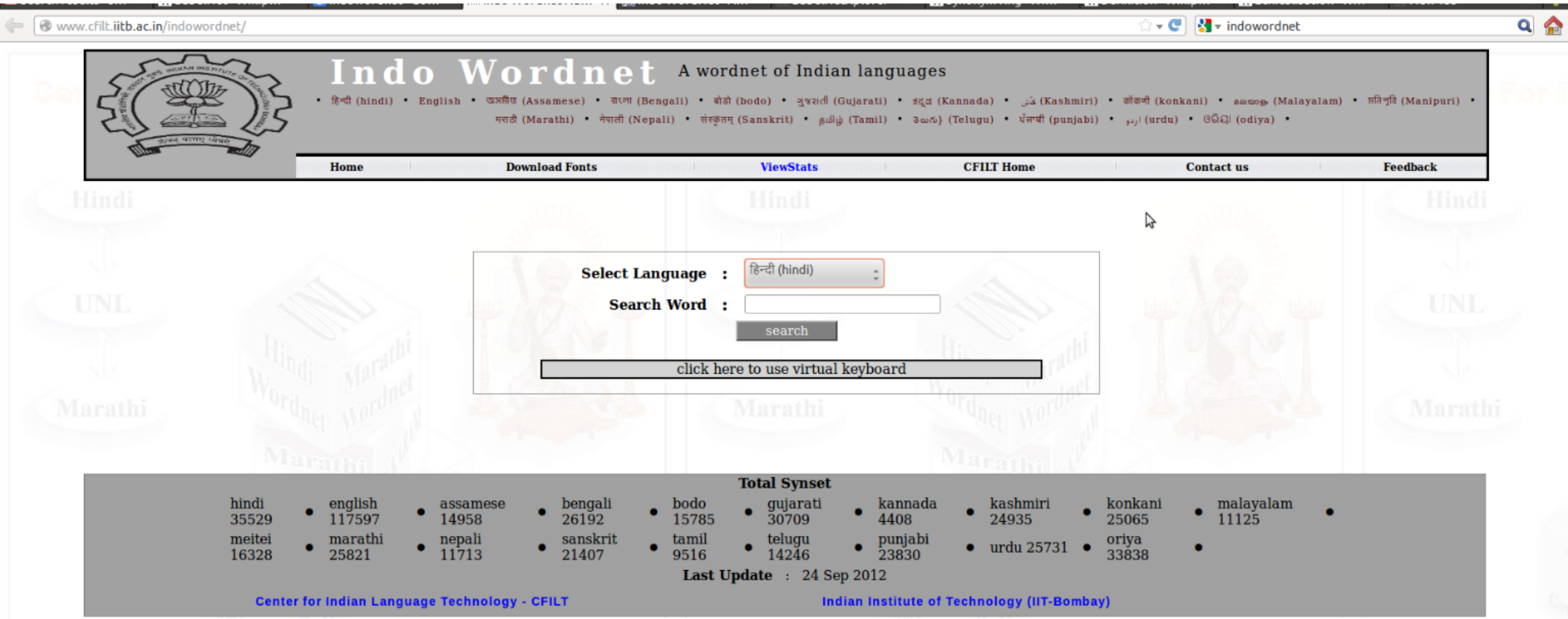
Indian languages form a very significant component of the languages landscape of the world.

Many languages rank within top 10 in the world in terms of the population speaking them, e.g., Hindi-Urdu 5th, Bangla 7th, Marathi 12th and so on as per the List of languages by number of native speakers.

Indowordnet project started with creations of Hindi WordNet by the Natural Language Processing group at the Center for Indian Language Technology (CFILT) in the Computer Science and Engineering Department at IIT Bombay.

AVAILABILITY

Publicaly browsable at
http://www.cfilt.iitb.ac.in/indowordnet
Available under GNU licence


www.cfilt.iitb.ac.in/indowordnet/
indowordnet
Indo Wordnet
A wordnet of Indian languages
• हिन्दी (hindi) • English • অসমীয়া (Assamese) • বাংলা (Bengali) • बोडो (bodo) • ગુજરાતી (Gujarati) • ಕನ್ನಡ (Kannada) • کشمیری (Kashmiri) • कोंकणी (konkani) • മലയാളം (Malayalam) • মণিপুরি (Manipuri) •
• मराठी (Marathi) • नेपाली (Nepali) • संस्कृतम् (Sanskrit) • தமிழ் (Tamil) • తెలుగు (Telugu) • ਪੰਜਾਬੀ (punjabi) • اردو (urdu) • ଓଡିଆ (odiya) •
Home
Download Fonts
ViewStats
CFILT Home
Contact us
Feedback
Select Language :
हिन्दी (hindi)
Search Word :
search
click here to use virtual keyboard
Total Synset
hindi 35529 • english 117597 • assamese 14958 • bengali 26192 • bodo 15785 • gujarati 30709 • kannada 4408 • kashmiri 24935 • konkani 25065 • malayalam 11125 •
meitei 16328 • marathi 25821 • nepali 11713 • sanskrit 21407 • tamil 9516 • telugu 14246 • punjabi 23830 • urdu 25731 • oriya 33838 •
Last Update : 24 Sep 2012
Center for Indian Language Technology - CFILT
Indian Institute of Technology (IIT-Bombay)

www.cfilt.iitb.ac.in/indowordnet/first?langno=0&queryword=वृक्ष

indowordnet

Indo Wordnet

A wordnet of Indian languages

• हिन्दी (hindi) • English • অসমীয়া (Assamese) • বাংলা (Bengali) • बोडो (bodo) • ગુજરાતી (Gujarati) • ಕನ್ನಡ (Kannada) • کشمیر (Kashmiri) • कोंकणी (konkani) • മലയാളം (Malayalam) • মনিপুরি (Manipuri) • मराठी (Marathi) • नेपाली (Nepali) • संस्कृतम् (Sanskrit) • தமிழ் (Tamil) • తెలుగు (Telugu) • ਪੰਜਾਬੀ (punjabi) • اردو (urdu) • ଓଡ଼ିଆ (odiya) •

Home | Download Fonts | ViewStats | CFILT Home | Contact us | Feedback

Number of Synset for " वृक्ष" : 1 showing 1 / 1

Synset ID : 2349 **POS :** noun

Synonyms : पेड़, वृक्ष, पादप, द्रुम, तरु, तरुवर, विटप, रुक्ष, रूख, विटपी, रूँख, अघ्रिप, अग, अनोकह, साखी, साखि, अमंद, अमन्द, शिखरी, शिखी, अर्क, स्कंधी, स्कन्धी, बीरो, जर्ण, पुलाकी, भूमिजात, आसना, प्रतिबंधक, प्रतिबन्धक, पल्लवी,

Gloss : जड़,तने,शाखा,तथा पत्तियों से युक्त बहुवर्षीय वनस्पति

Example statement : "पेड़ मनुष्य के लिए बहुत ही उपयोगी हैं"

Gloss in Hindi : जड़,तने,शाखा,तथा पत्तियों से युक्त बहुवर्षीय वनस्पति:"पेड़ मनुष्य के लिए बहुत ही उपयोगी हैं"

Gloss in English : a tall perennial woody plant having a main trunk and branches forming a distinct elevated crown; includes both gymnosperms and angiosperms

Select Language : മലയാളം (Malayalam)

Search Word :

search

click here to use virtual keyboard

Current language : हिन्दी (hindi)

Change language : हिन्दी (hindi)

Relations

- hypernymy
- hyponymy
- holonymy
- meronymy
- antonymy
- Onto tree
- noun relation
- verb relation
- derived from
- modifies

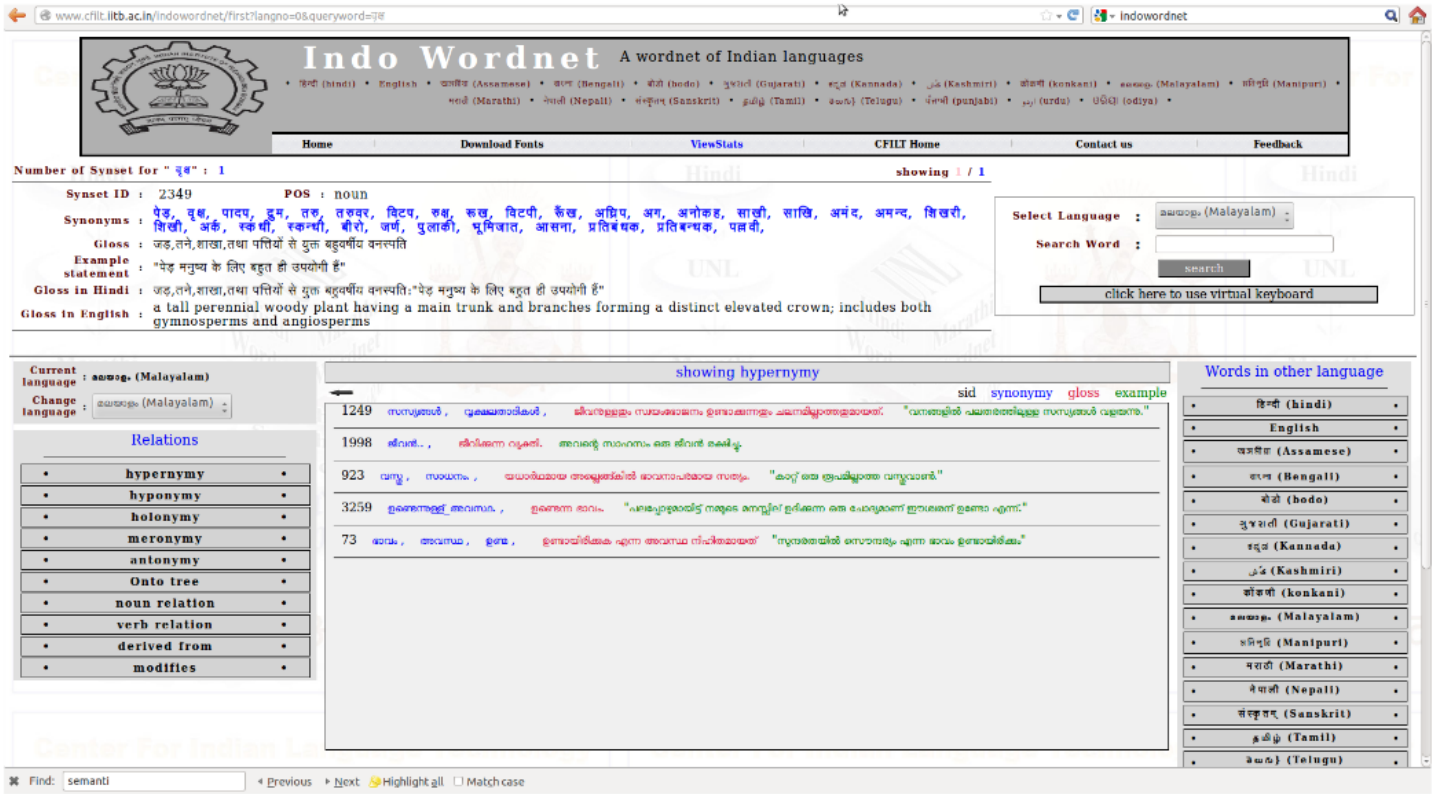
showing hypernymy

sid	synonymy	gloss	example
1249	वनस्पति , पेड़-पौधा ,	वह सजीव जिसमें गति नहीं होती है और अधिकांशतः वह अपना भोजन स्वयं बनाता है	"जंगलों में तरह-तरह की वनस्पतियाँ पायी जाती हैं"
1998	जीव , प्राणी , जीवधारी , जीवात्मा , अनीश , सजीव , प्राणधारी , तनुधारी , जीवक , प्राणक , आसना , मंदसानु , मन्दसानु ,	सजीव प्राणी या वह जिसमें प्राण हो	"पृथ्वी पर विभिन्न प्रकार के जीव पाये जाते हैं"
923	वस्तु , चीज़ , चीज ,	वास्तविक या कल्पित सत्ता	"हवा एक अमूर्त वस्तु है"
3259	अस्तित्व , मौजूदगी , मौजूदगी , वजूद , वजूद , संभूति , विद्यमानता , सत्ता , हस्ती , भव , अस्ति ,	सत्ता का भाव	"कभी-कभी हमारे मन में यह प्रश्न उठता है कि क्या ईश्वर का अस्तित्व है"
73	भाव ,	वह जिसमें होने की क्रिया निहित हो	"सुंदरता में सुंदर होने का भाव है"

Words in other language

- हिन्दी (hindi)
- English
- অসমীয়া (Assamese)
- বাংলা (Bengali)
- बोडो (bodo)
- ગુજરાતી (Gujarati)
- ಕನ್ನಡ (Kannada)
- کشمیر (Kashmiri)
- कोंकणी (konkani)
- മലയാളം (Malayalam)
- মনিপুরি (Manipuri)
- मराठी (Marathi)
- नेपाली (Nepali)
- संस्कृतम् (Sanskrit)
- தமிழ் (Tamil)
- తెలుగు (Telugu)

Find: semanti ◂ Previous ▸ Next Highlight all Match case


www.cfilt.iitb.ac.in/indowordnet/first?langno=0&queryword=वृक्ष
Indowordnet
Indo Wordnet
A wordnet of Indian languages
Home
Download Fonts
ViewStats
CFILT Home
Contact us
Feedback
Number of Synset for " वृक्ष" : 1
showing 1 / 1
Synset ID : 2349
POS : noun
Synonyms : पेड़, वृक्ष, पादप, द्रुम, तरु, तरुवर, विटप, रुख, रूख, विटपी, कैंख, अघ्रिप, अग, अनोकह, साखी, साखि, अमंद, अमन्द, शिखरी, शिखी, अंकै, स्कंधी, स्कन्धी, बीरी, जर्ण, पुलाकी, भूमिजात, आसना, प्रतिबंधक, प्रतिबन्धक, पल्वी,
Gloss : जड़,तने,शाखा,तथा पत्तियों से युक्त बहुवर्षीय वनस्पति
Example statement : "पेड़ मनुष्य के लिए बहुत ही उपयोगी हैं"
Gloss in Hindi : जड़,तने,शाखा,तथा पत्तियों से युक्त बहुवर्षीय वनस्पति:"पेड़ मनुष्य के लिए बहुत ही उपयोगी हैं"
Gloss in English : a tall perennial woody plant having a main trunk and branches forming a distinct elevated crown; includes both gymnosperms and angiosperms
Select Language : മലയാളം (Malayalam)
Search Word :
search
click here to use virtual keyboard
Current language : മലയാളം (Malayalam)
Change language : മലയാളം (Malayalam)
Relations
hypernymy
hyponymy
holonymy
meronymy
antonymy
Onto tree
noun relation
verb relation
derived from
modifies
showing hypernymy
sid synonymy gloss example
1249 സസ്യങ്ങൾ , വൃക്ഷജാതികൾ , ജീവനുള്ളതും സ്വയംപ്രാപ്തവുമായ ഉണ്ടാക്കുന്നതും ചലനമില്ലാത്തതുമായത്. "മനുഷ്യരിൽ പലതരത്തിലുള്ള സസ്യങ്ങൾ വളരുന്നു."
1998 ജീവൻ.. , ജീവിക്കുന്ന വ്യക്തി. അവന്റെ സാഹസം കൊണ്ട് ഒരു ജീവൻ രക്ഷിച്ചു
923 വസ്തു , സാധനം , യഥാർത്ഥമായ അല്ലെങ്കിൽ ഭാവനാപരമായ സത്ത. "കാറ്റ് ഒരു ശൂന്യിലാത്ത വസ്തുവാണ്."
3259 ഉണ്ടെന്നതിന്റെ അവസ്ഥ , ഉണ്ടെന്ന ഭാവം. "പലപ്പോഴായിട്ട് നമ്മുടെ മനസ്സിലെ ഉറപ്പിക്കുന്ന ഒരു പോരായ്മകൾ ഈശ്വരന് ഉണ്ടോ എന്ന്."
73 ഭാവം , അവസ്ഥ , ഉണ്ടു , ഉണ്ടായിരിക്കുക എന്ന അവസ്ഥ നിഷ്കർഷയായത് "സുരക്ഷയിൽ സൌന്ദര്യം എന്ന ഭാവം ഉണ്ടായിരിക്കും."
Words in other language
हिन्दी (hindi)
English
অসমীয়া (Assamese)
বাংলা (Bengali)
बोडो (bodo)
ગુજરાતી (Gujarati)
ಕನ್ನಡ (Kannada)
کٲشُر (Kashmiri)
कोंकणी (konkani)
മലയാളം (Malayalam)
মনিপুরি (Manipuri)
मराठी (Marathi)
नेपाली (Nepali)
संस्कृतम् (Sanskrit)
தமிழ் (Tamil)
తెలుగు (Telugu)
Find: semanti
Previous
Next
Highlight all
Match case

Cyc (pronounced like syke)

Cyc is an artificial intelligence project that attempts to assemble a comprehensive ontology and knowledge base of everyday common sense knowledge

MOTIVATION

To codify, in machine-usable form, millions of pieces of knowledge that compose human common sense.

To enable AI applications to perform human-like reasoning.

KNOWLEDGE BASE

INFERENCE ENGINE

An inference engine is a computer program that tries to derive answers from a knowledge base. The Cyc inference engine performs general logical deduction (including modus ponens, modus tollens, universal quantification and existential quantification).

Typical pieces of knowledge represented in the database are "Every tree is a plant" and "Plants die eventually". When asked whether trees die, the inference engine can draw the obvious conclusion and answer the question correctly.

AVAILABILITY

The original knowledge base is proprietary, but a smaller version of the knowledge base, intended to establish a common vocabulary for automatic reasoning, was released as OpenCyc under an open source (Apache) license. More recently, Cyc has been made available to AI researchers under a research-purposes license as ResearchCyc.

APPLICATIONS

pronounced like

Cyc is an artificial intelligence project that attempts to assemble a comprehensive ontology and knowledge base of everyday common sense knowledge

MOTIVATION

To codify, in machine-usable form, millions of pieces of knowledge that compose human common sense.

To enable AI applications to perform human-like reasoning.

KNOWLEDGE BASE

The Knowledge Base (KB) contains over one million human-defined assertions, rules or common sense ideas.

REPRESENTATION

The concept names in Cyc are known as constants. Constants start with an optional "#$" and are case-sensitive. There are constants for:

Individual items known as Individuals, such as #$BillClinton or #$France.

Collections, such as #$Tree-ThePlant (containing all trees) or #$EquivalenceRelation (containing all equivalence relations). A member of a collection is called an instance of that collection.

Truth Functions which can be applied to one or more other concepts and return either true or false. For example #$siblings is the sibling relationship, true if the two arguments are siblings.

Functions, which produce new terms from given ones. For example, #$FruitFn, when provided with an argument describing a type (or collection) of plants, will return the collection of its fruits. By convention, function constants start with an upper-case letter and end with the string "Fn".

CycL

These are formulated in the language CycL, which is based on predicate calculus and has a syntax similar to that of the Lisp programming language.

```
;;
(#$isa #$BillClinton #$UnitedStatesPresident)
"Bill Clinton belongs to the collection of U.S. presidents"
(#$genls #$Tree-ThePlant #$Plant)
"All trees are plants".
```

KNOWLEDGE BASE

The Knowledge Base (KB) contains over one million human-defined assertions, rules or common sense ideas.

REPRESENTATION

The concept names in Cyc are known as constants. Constants start with an optional "#$" and are case-sensitive. There are constants for:

Individual items known as individuals, such as #$BillClinton or #$France.

Collections, such as #$Tree-ThePlant (containing all trees) or #$EquivalenceRelation (containing all equivalence relations). A member of a collection is called an instance of that collection.

Truth Functions which can be applied to one or more other concepts and return either true or false. For example #$siblings is the sibling relationship, true if the two arguments are siblings.

Functions, which produce new terms from given ones. For example, #$FruitFn, when provided with an argument describing a type (or collection) of plants, will return the collection of its fruits. By convention, function constants start with an upper-case letter and end with the string "Fn".

CycL

These are formulated in the language CycL, which is based on predicate calculus and has a syntax similar to that of the Lisp programming language.

eg:

(#$isa #$BillClinton #$UnitedStatesPresident)

"Bill Clinton belongs to the collection of U.S. presidents"

(#$genls #$Tree-ThePlant #$Plant)

"All trees are plants".

INFERENCE ENGINE

An inference engine is a computer program that tries to derive answers from a knowledge base. The Cyc inference engine performs general logical deduction (including modus ponens, modus tollens, universal quantification and existential quantification).

Typical pieces of knowledge represented in the database are "Every tree is a plant" and "Plants die eventually". When asked whether trees die, the inference engine can draw the obvious conclusion and answer the question correctly.

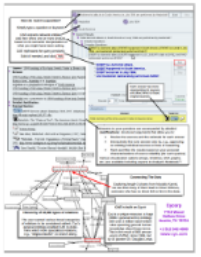
AVAILABILITY

The original knowledge base is proprietary, but a smaller version of the knowledge base, intended to establish a common vocabulary for automatic reasoning, was released as OpenCyc under an open source (Apache) license. More recently, Cyc has been made available to AI researchers under a research-purposes license as ResearchCyc.

APPLICATIONS

Terrorism Knowledge Base

The comprehensive Terrorism Knowledge Base is an application of Cyc in development that will try to ultimately contain all relevant knowledge about "terrorist" groups, their members, leaders, ideology, founders, sponsors, affiliations, facilities, locations, finances, capabilities, intentions, behaviors, tactics, and full descriptions of specific terrorist events. The knowledge is stored as statements in mathematical logic, suitable for computer understanding and reasoning.

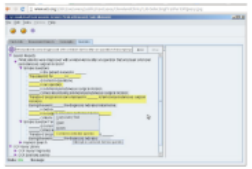


Cyclopedia

Cyclopedia is being developed; it superimposes Cyc keywords on pages taken from Wikipedia pages

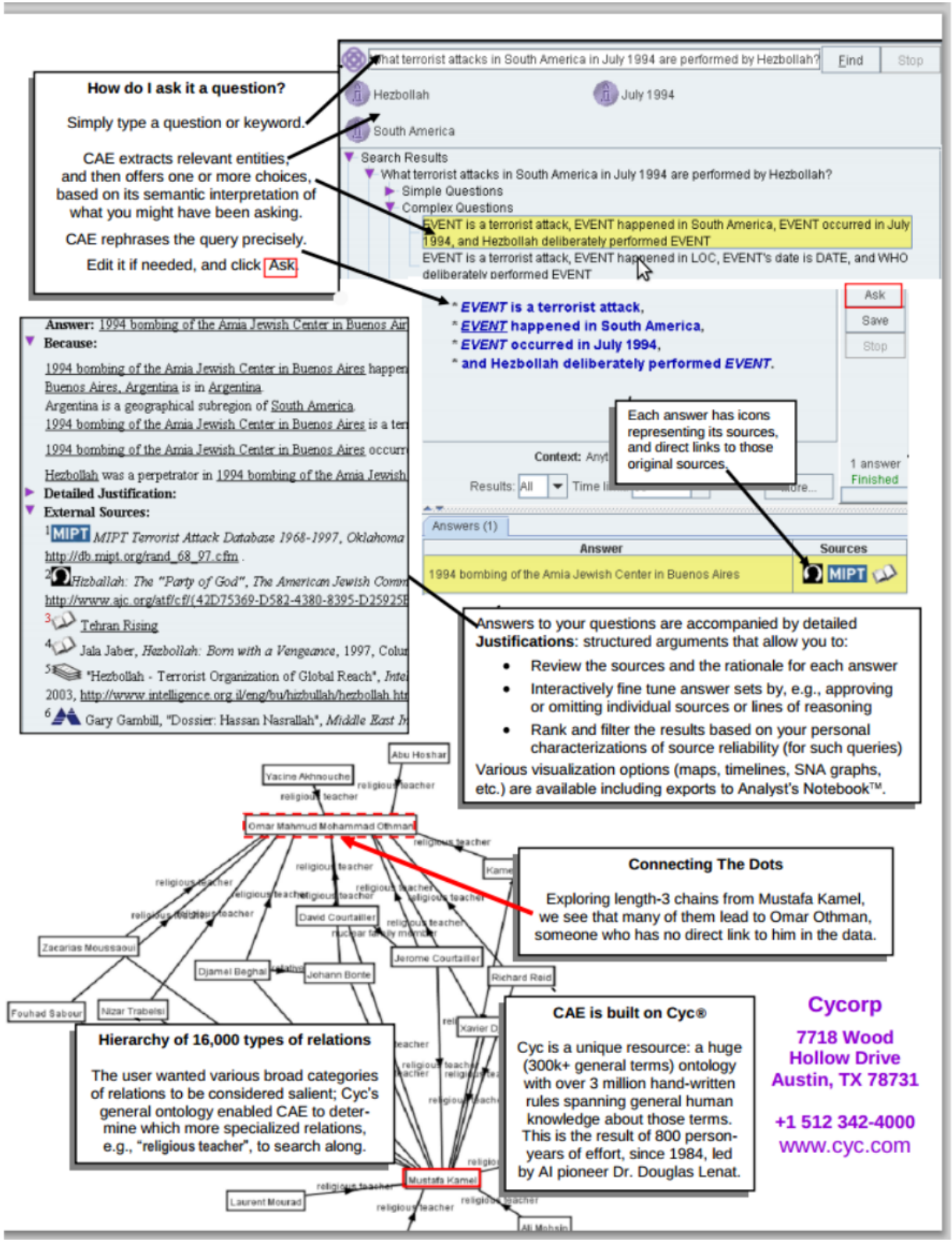
Cleveland Clinic Foundation

The Cleveland Clinic has used Cyc to develop a natural language query interface of biomedical information.



Terrorism Knowledge Base

The comprehensive Terrorism Knowledge Base is an application of Cyc in development that will try to ultimately contain all relevant knowledge about "terrorist" groups, their members, leaders, ideology, founders, sponsors, affiliations, facilities, locations, finances, capabilities, intentions, behaviors, tactics, and full descriptions of specific terrorist events. The knowledge is stored as statements in mathematical logic, suitable for computer understanding and reasoning.


How do I ask it a question?
Simply type a question or keyword.
CAE extracts relevant entities, and then offers one or more choices, based on its semantic interpretation of what you might have been asking.
CAE rephrases the query precisely.
Edit it if needed, and click Ask.
What terrorist attacks in South America in July 1994 are performed by Hezbollah?
Find
Stop
Hezbollah
July 1994
South America
Search Results
What terrorist attacks in South America in July 1994 are performed by Hezbollah?
Simple Questions
Complex Questions
EVENT is a terrorist attack, EVENT happened in South America, EVENT occurred in July 1994, and Hezbollah deliberately performed EVENT
EVENT is a terrorist attack, EVENT happened in LOC, EVENT's date is DATE, and WHO deliberately performed EVENT
* EVENT is a terrorist attack,
* EVENT happened in South America,
* EVENT occurred in July 1994,
* and Hezbollah deliberately performed EVENT.
Ask
Save
Stop
Answer: 1994 bombing of the Amia Jewish Center in Buenos Air
Because:
1994 bombing of the Amia Jewish Center in Buenos Aires happen
Buenos Aires, Argentina is in Argentina.
Argentina is a geographical subregion of South America.
1994 bombing of the Amia Jewish Center in Buenos Aires is a ter
1994 bombing of the Amia Jewish Center in Buenos Aires occurr
Hezbollah was a perpetrator in 1994 bombing of the Amia Jewish
Detailed Justification:
External Sources:
1 MIPT MIPT Terrorist Attack Database 1968-1997, Oklahoma
http://db.mipt.org/rand_68_97.cfm .
2 Hizballah: The "Party of God", The American Jewish Comm
http://www.ajc.org/atf/cf/{42D75369-D582-4380-8395-D25925E
3 Tehran Rising
4 Jala Jaber, Hezbollah: Born with a Vengeance, 1997, Colu
5 "Hezbollah - Terrorist Organization of Global Reach", Inte
2003, http://www.intelligence.org.il/eng/bu/hizbullah/hezbollah.htr
6 Gary Gambill, "Dossier: Hassan Nasrallah", Middle East I
Each answer has icons representing its sources, and direct links to those original sources.
Context: Anyt
Results: All
Time li
more...
1 answer
Finished
Answers (1)
Answer
Sources
1994 bombing of the Amia Jewish Center in Buenos Aires
MIPT
Answers to your questions are accompanied by detailed Justifications: structured arguments that allow you to:
Review the sources and the rationale for each answer
Interactively fine tune answer sets by, e.g., approving or omitting individual sources or lines of reasoning
Rank and filter the results based on your personal characterizations of source reliability (for such queries)
Various visualization options (maps, timelines, SNA graphs, etc.) are available including exports to Analyst's Notebook™.
Abu Hoshar
Yacine Akhnouche
religious teacher
Omar Mahmud Mohammad Othman
David Courtailler
Zacarias Moussaoui
Djamel Beghal
Johann Bonte
Jerome Courtailler
Richard Reid
Fouhad Sabour
Nizar Trabelsi
Mustafa Kamel
Laurent Mourad
Connecting The Dots
Exploring length-3 chains from Mustafa Kamel, we see that many of them lead to Omar Othman, someone who has no direct link to him in the data.
Hierarchy of 16,000 types of relations
The user wanted various broad categories of relations to be considered salient; Cyc's general ontology enabled CAE to determine which more specialized relations, e.g., "religious teacher", to search along.
CAE is built on Cyc®
Cyc is a unique resource: a huge (300k+ general terms) ontology with over 3 million hand-written rules spanning general human knowledge about those terms. This is the result of 800 person-years of effort, since 1984, led by AI pioneer Dr. Douglas Lenat.
Cycorp
7718 Wood Hollow Drive
Austin, TX 78731
+1 512 342-4000
www.cyc.com

reasoning.

Cyclopedia

Cyclopedia is being developed; it superimposes Cyc keywords on pages taken from Wikipedia pages

www.cycfoundation.org/cyclopedia.jpg

CYCLOPEDIA

concept navigator

Poodle

more specific

- Poodle
- German Shepherd
- Collie

articles

- Canine
- Dog
- Dog Training

search

Go Search

Sign in / create account

article | discussion | view source | history

Your continued donations keep Wikipedia running!

Dog

From Wikipedia, the free encyclopedia

This article is about the domestic dog. For other uses, see Dog (disambiguation).

The **dog** (*Canis lupus familiaris*) is a domestic subspecies of the wolf, a mammal of the Canidae family the order Carnivora. The term encompasses both feral and pet variants. It i… species.

Over time, t… variation. For example, heights at th… the Chihuahua) to roughly three feet … white through grays (usually called *b*… ("red" or "chocolate") in a tremendou… from very short to several centimete… wool, straight or curly, or smooth.

Biological Family

Family-Human

Family-Nuclear

The collection of biological taxonomic subdivisions below biological order (or biological suborder) and above biological genus. Especially important in Botanical classification.

related articles:

- Family (biology)

*Yes, this is the correct meaning of **family**.*

Contents [hide]

1 Overview
- 1.1 Origins
- 1.2 Human relationships

2 Terminology

3 Intelligence
- 3.1 Evaluation of a dog's intelligence

4 Physical characteristics
- 4.1 Sight
- 4.2 Hearing
- 4.3 Smell

5 Working, utility and assistance dogs

6 Show and sport (competition) dogs

7 Diet
- 7.1 Dangerous substances

8 Reproduction
- 8.1 General
- 8.2 Males
- 8.3 Females
- 8.4 Copulation
- 8.5 Gestation and litters

Domestic dog

Fossil range: Late Pleistocene - Recent

other images of dogs

Conservation status

Domesticated

Scientific classification

Domain:	Eukaryota
Kingdom:	Animalia
Phylum:	Chordata
Class:	Mammalia
Order:	Carnivora
Family:	Canidae
Genus:	*Canis*
Species:	*C. lupus*
Subspecies:	***C. l. familiaris***

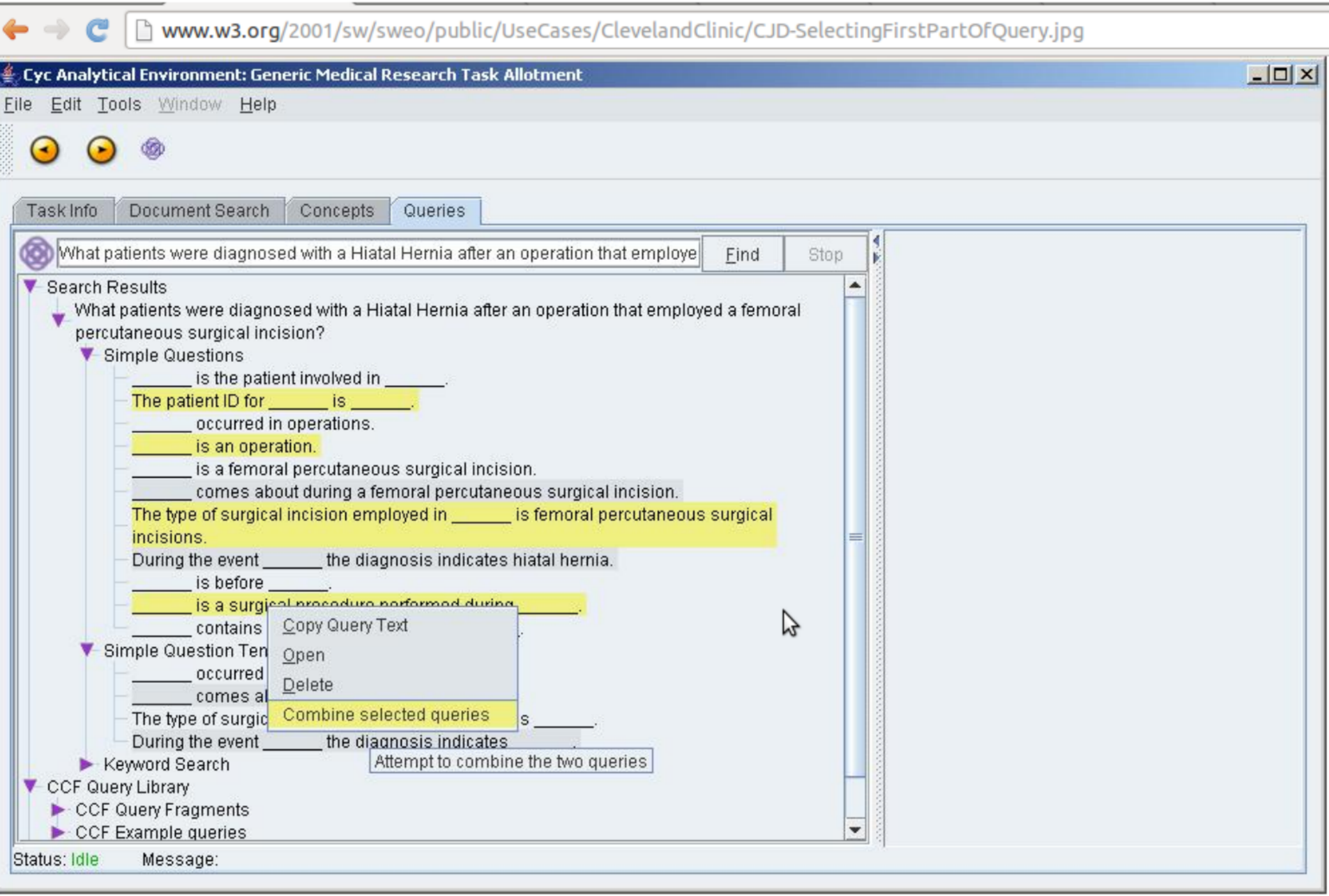
Trinomial name

Canis lupus familiaris

Dogs Portal

Cleveland Clinic Foundation

The Cleveland Clinic has used Cyc to develop a natural language query interface of biomedical information.


www.w3.org/2001/sw/sweo/public/UseCases/ClevelandClinic/CJD-SelectingFirstPartOfQuery.jpg
Cyc Analytical Environment: Generic Medical Research Task Allotment
File Edit Tools Window Help
Task Info
Document Search
Concepts
Queries
What patients were diagnosed with a Hiatal Hernia after an operation that employe
Find
Stop
Search Results
What patients were diagnosed with a Hiatal Hernia after an operation that employed a femoral percutaneous surgical incision?
Simple Questions
_____ is the patient involved in _____.
The patient ID for _____ is _____.
_____ occurred in operations.
_____ is an operation.
_____ is a femoral percutaneous surgical incision.
_____ comes about during a femoral percutaneous surgical incision.
The type of surgical incision employed in _____ is femoral percutaneous surgical incisions.
During the event _____ the diagnosis indicates hiatal hernia.
_____ is before _____.
_____ is a surgical procedure performed during _____.
_____ contains
Simple Question Ten
_____ occurred
_____ comes al
The type of surgic
s _____.
During the event _____ the diagnosis indicates
Copy Query Text
Open
Delete
Combine selected queries
Attempt to combine the two queries
Keyword Search
CCF Query Library
CCF Query Fragments
CCF Example queries
Status: Idle
Message:

NLP & Semantic web

- Relationship between NLP & Semantic Web
- Mutual Benefit

Why NLP needed for Semantic Web ?

Ontology Query

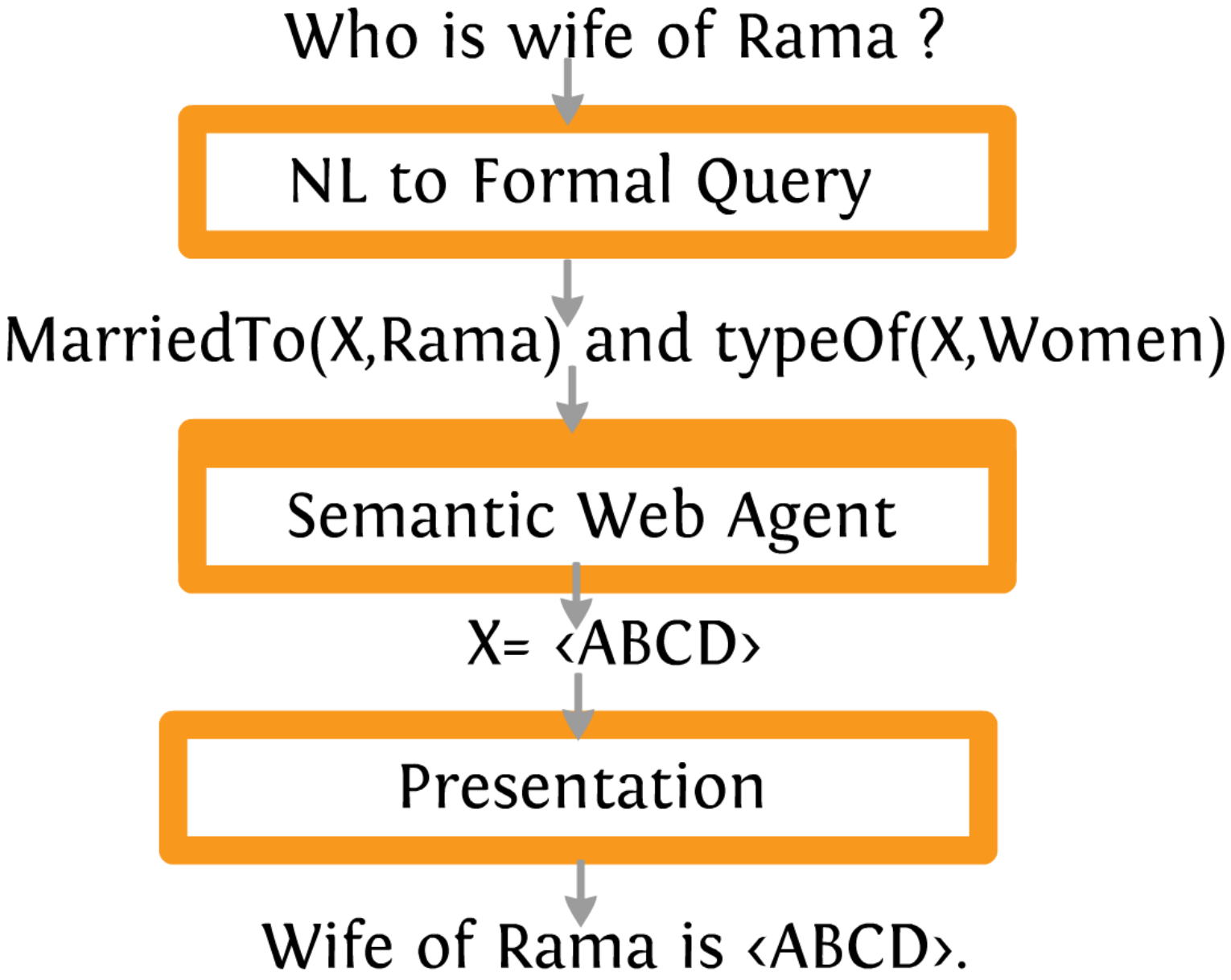
-Semantic Web has the quite complex structure

1.Ontology Query
2.Ontology Learning
3.Multilingual
Ontology Mapping

Ontology Query

-Semantic Web has the quite complex structure and strict logical form
- It is not expected that Users will make a query logic form

NL Query : Who is wife of Rama ?
Logical Query :
MarriedTo(?,Rama) and typeOf(?,Men)


Who is wife of Rama ?
NL to Formal Query
MarriedTo(X,Rama) and typeOf(X,Women)
Semantic Web Agent
X= ‹ABCD›
Presentation
Wife of Rama is ‹ABCD›.

Ontology Learning

Layer - Learning

Terms

Unstructured sources

Involves NLP techniques, morphological and syntactic analysis, etc.

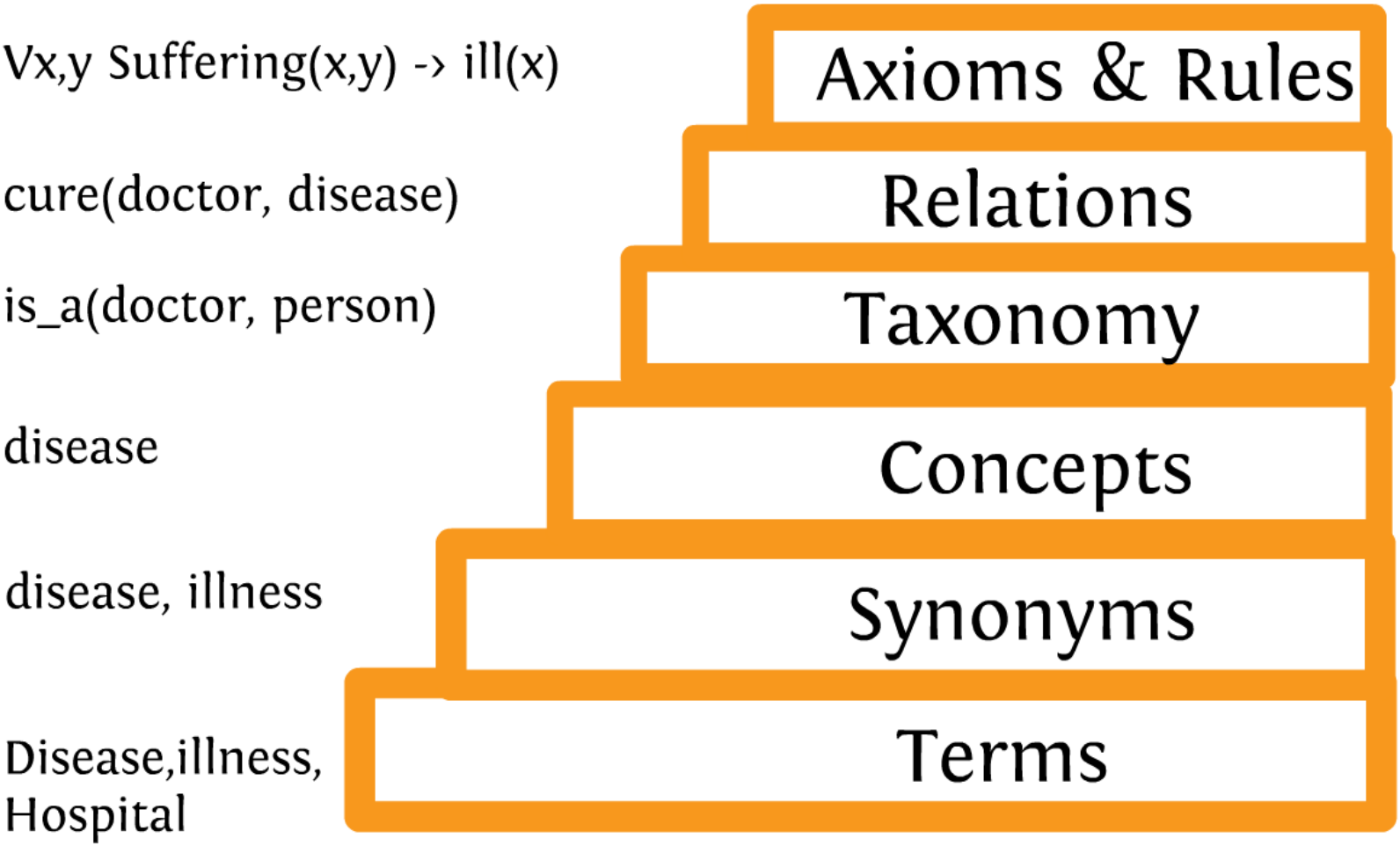
Semi-structured source

elicit an ontology from sources that have some predefined structure, such as XML Schema

Structured data

Extracting concepts and relations from knowledge contained in structured data, such as databases

Layer - Learning



Terms

- Linguistic realizations of domain-specific concepts
- Are the basis of the ontology learning process

Term Extraction

-Run a Part-Of-Speech (POS) tagger over the domain corpus

-Identify possible terms by constructing patterns, such as: Adj-Noun, Noun-noun, Adj-Noun-Noun,...

-Ignore Names

-Identify only the relevant to the text terms by applying statistical metrics

Taxonomy Extraction

1.With the use of WordNet
2.Co-occurrence Analysis
3.
Lexico-syntactic patterns

Taxonomy Extraction with wordnet

-Given two terms t1 and t2, check if they stand in a hypernym relation with regard to WordNet

-Normalize the number of hypernym paths by dividing by the number of senses of t1

Example: 4 different hypernym paths between synsets 'country' and 'region'

- 'country' has 5 senses

value of isa (country, region) = 0.8

Lexico-syntactic patterns

- NPo such as {NP1, NP2,..., (and | or)} NPn
Vehicles such as cars, trucks and bikes....

- such NP as {NP,} * { (or | and) } NP
Such fruits as oranges, nectarines or apples...

S

- NP {, NP} * { , } { or | and } other NP
Swimming, running, or/and other activities...

- NP { , } including {NP, } * { or | and } NP
Injuries, including broken bones, wounds and bruises...

Relations

-General Relations

Exploiting linguistic structure

Ex: Author wrote a book.

Relation: write(Author, Book)

Why Semantic Web is needed for NLP ?

Why Semantic Web is needed for NLP ?

Word Sense Disambiguation

- Locate the Least Weighted Path from One Ontology Concept to Other Concept

Example

1. Sachin's play is awesome.
2. Shakesphere's play is awesome.

Example

1. Sachin's play is awesome.

2. Shakesphere's play is awesome.

Sachin -› Cricketer-› Play_1
Shakesphere -› Playwriter-› Play_2

Conclusion

Conclusion

-Without any doubt, Semantic Web needs natural language technology

-To acquire knowledge from massive unstructured/ semi-structured Web 2.0

-To understand user's Natural Language Query

-Finally, To present answers to users

-Semantic Web should improve the NLP's performance in IE and WSD

-Almost all problems are in open to research , a long journey to go.

References

1. Philipp Cimiano and Bernardo Magnini, Ontology Learning from Text: An Overview, IOS Press, 2003
2. Ruiqiang Guo, Fuji Ren, Towards the Relationship Between Semantic Web and NLP, IEEE, 2009
3. C'ssia Trojahn , Paulo Quaresma , Renata Vieira, A Framework for Multilingual Ontology Mapping, The International Conference on Language Resources and Evaluation, 2008
4. Christian Chiarcos, Interoperability of Corpora and Annotations, DFG, 2011
5. John McCrae, Dennis Spohr, and Philipp Cimiano Linking Lexical Resources and Ontologies on the Semantic Web with lemon, AG Semantic Computing, CITEC, University of Bielefeld, 2012
6. Roberto Navigli, Simone Paolo Ponzetto, BabelNet: Building a Very Large Multilingual Semantic Network, ACL, 2012