

Linked Open Data

Una “piccola” introduzione

Davide Taibi

Classificazione dei dati aperti

★ Disponibile sul web (in qualunque formato) purché distribuito con una licenza aperta (PDF, immagini, ...)

★★ Disponibile in un formato strutturato che sia elaborabile in modo automatico (machine-readable) (e.g. tabella MS excel invece dell'immagine di una tabella di dati)

★★★ Come il precedente ma distribuito con un format non proprietario (e.g. CSV invece di MS excel)

★★★★ Come sopra con in più l'uso di standards aperti del W3C (e.g. RDF e SPARQL) per identificare le entità, così che chiunque possa indirizzarli

★★★★★ Come sopra, in più: collegare i dati con i dati distribuiti da altri Linked Open Data

Linked (Open) Data

Does all Linked Data need to be Linked Open Data?

In a word, no, and it will likely never be so. The label "Linked Open Data" is widely used, but often to refer to Linked Data in general, rather than to Linked Data that is explicitly published under an open license. **Not all Linked Data will be open, and not all Open Data will be linked.** Therefore care should be taken to use the appropriate term, depending on the licensing terms of the data in question.

<http://linkeddata.org/faq>

Perché collegare i dati?

Condividere dati

Superadditività

$$v(S \cup T) \geq v(S) + v(T)$$

Interoperabilità

Attivare processi di Inferenza

Semantic Reasoning



Inferenza

```
graph TD; Inferenza([Inferenza]) --- Regole([basata su regole]); Inferenza --- Ontologie([basata sulle ontologie]); Regole --- RegoleDesc[Produce valide asserzioni partendo da un sistema basato su regole.]; Ontologie --- OntologieDesc[Deriva informazioni aggiuntive basandosi sulla definizione di classi o istanze di uno specifico dominio.]; RegoleDesc --- SWLR([SWLR]); OntologieDesc --- RDF([RDF]); OntologieDesc --- OWL([OWL]);
```

basata su
regole

Produce valide asserzioni partendo da
un sistema basato su regole.

SWLR

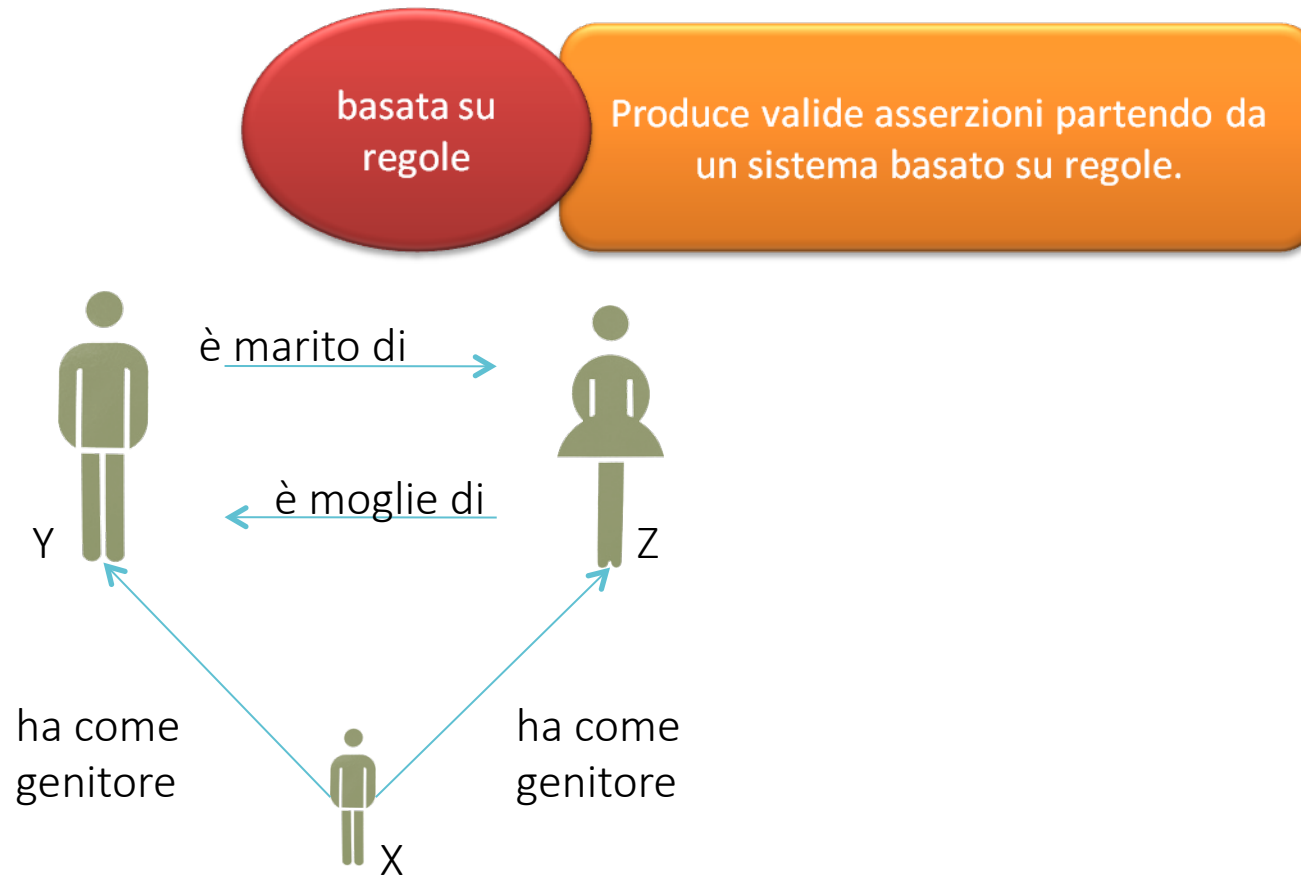
basata sulle
ontologie

Deriva informazioni aggiuntive
basandosi sulla definizione di classi o
istanze di uno specifico dominio.

RDF

OWL

Inferenza basata su regole



Regole:

Se: X ha come genitore Y

Se: X ha come genitore Z

Se: Y è uomo

Se: Z è donna

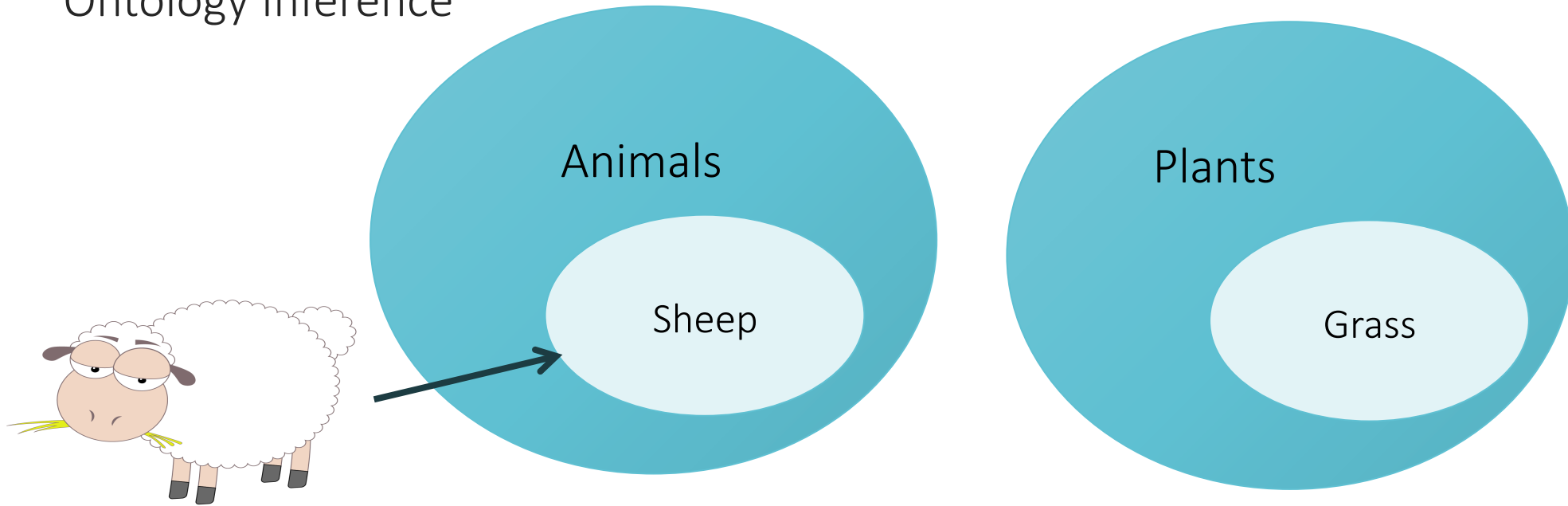
Asserzioni:

Allora: Z è moglie di Y

Allora: Y è marito di Z

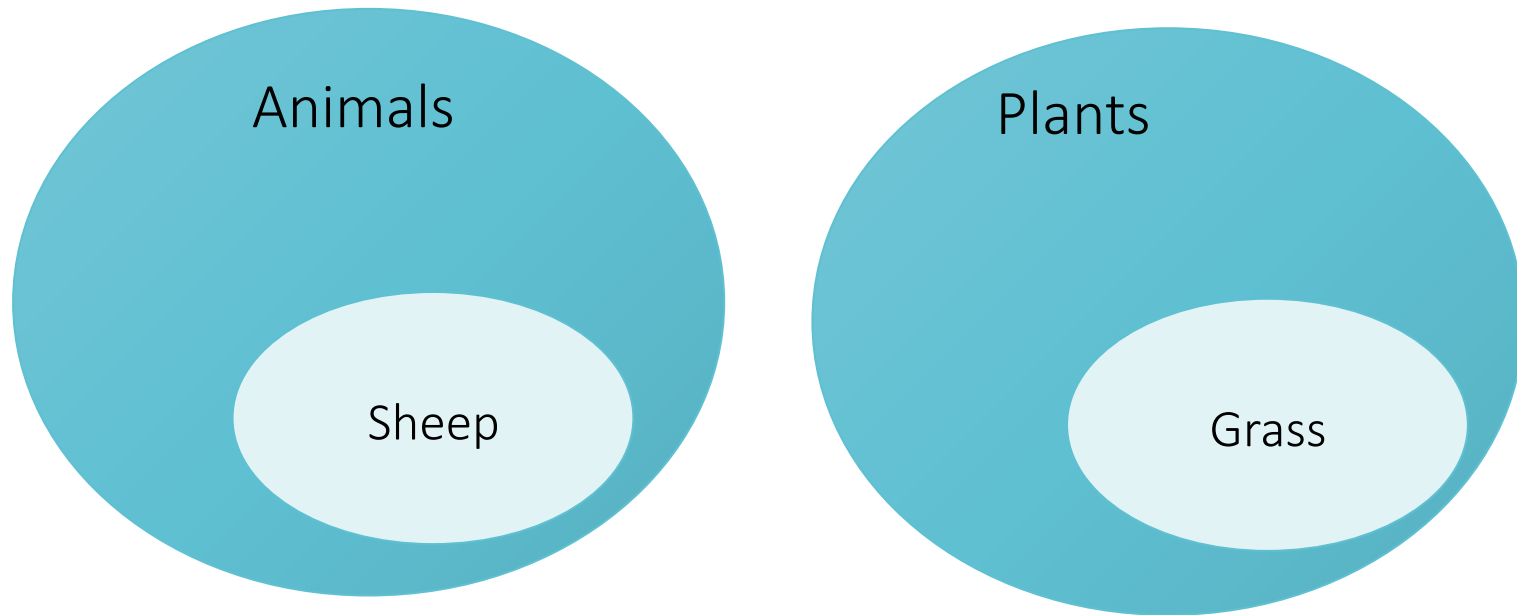
Semantic Reasoning

Ontology Inference



Semantic Reasoning

Combine reasoning methods



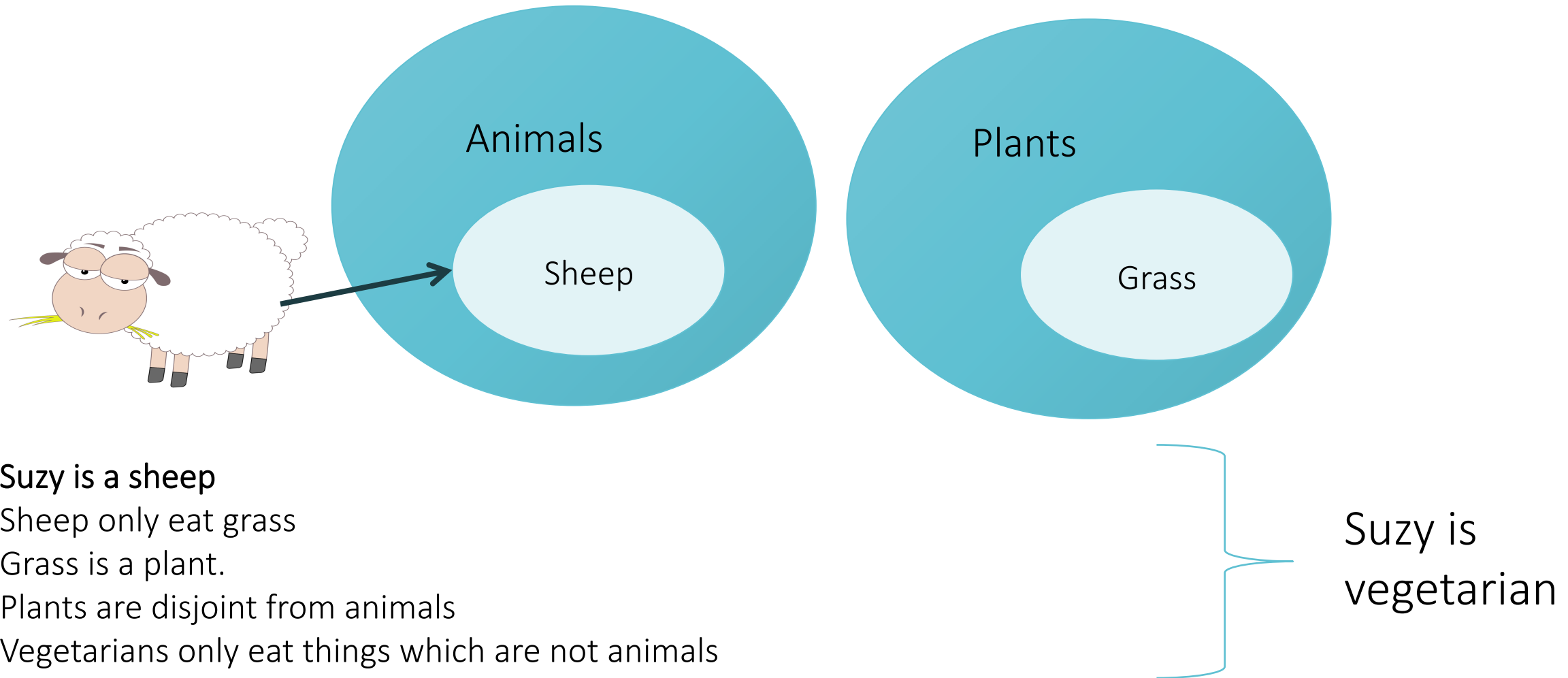
Rules:

Sheep only eat grass

Vegetarians only eat things which are not animals

Semantic Reasoning

The sheep Suzy is vegetarian

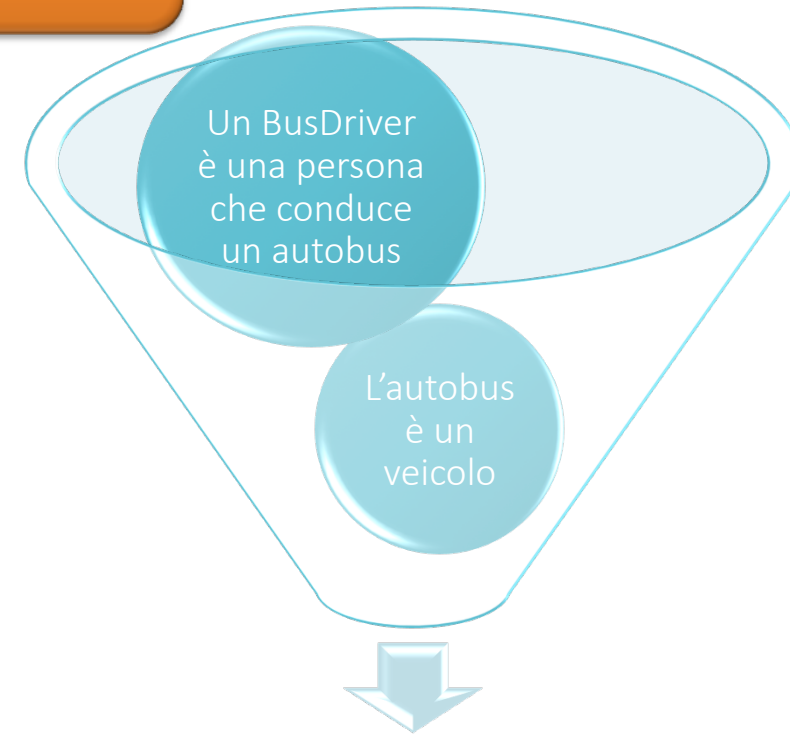


Inferenza basata su ontologie

basata sulle
ontologie

Deriva informazioni aggiuntive
basandosi sulla definizione di classi o
istanze di uno specifico dominio.

- ▶ *Class(a:bus_driver complete
intersectionOf(a:person restriction(a:drives
someValuesFrom (a:bus))))*
- ▶ *Class(a:driver complete
intersectionOf(a:person restriction(a:drives
someValuesFrom (a:vehicle))))*
- ▶ *Class(a:bus partial a:vehicle)*
- ▶ *A bus_driver is a person that drives a bus.*
- ▶ *A bus is a vehicle.*
- ▶ *A bus_driver drives a vehicle.*



L'essenza del Linked Data

«Proprio come i collegamenti ipertestuali nel Web classico che collegano documenti in un unico spazio di informazione globale, i Linked Data consentono collegamenti tra entità di diverse fonti di dati e quindi consente di collegare queste fonti di dati in un unico spazio di dati globale Questa è l'essenza del Linked Data.»

“Linked Data: Evolving the Web into a Global Data Space”

Tom Heath and Christian Bizer (2011)

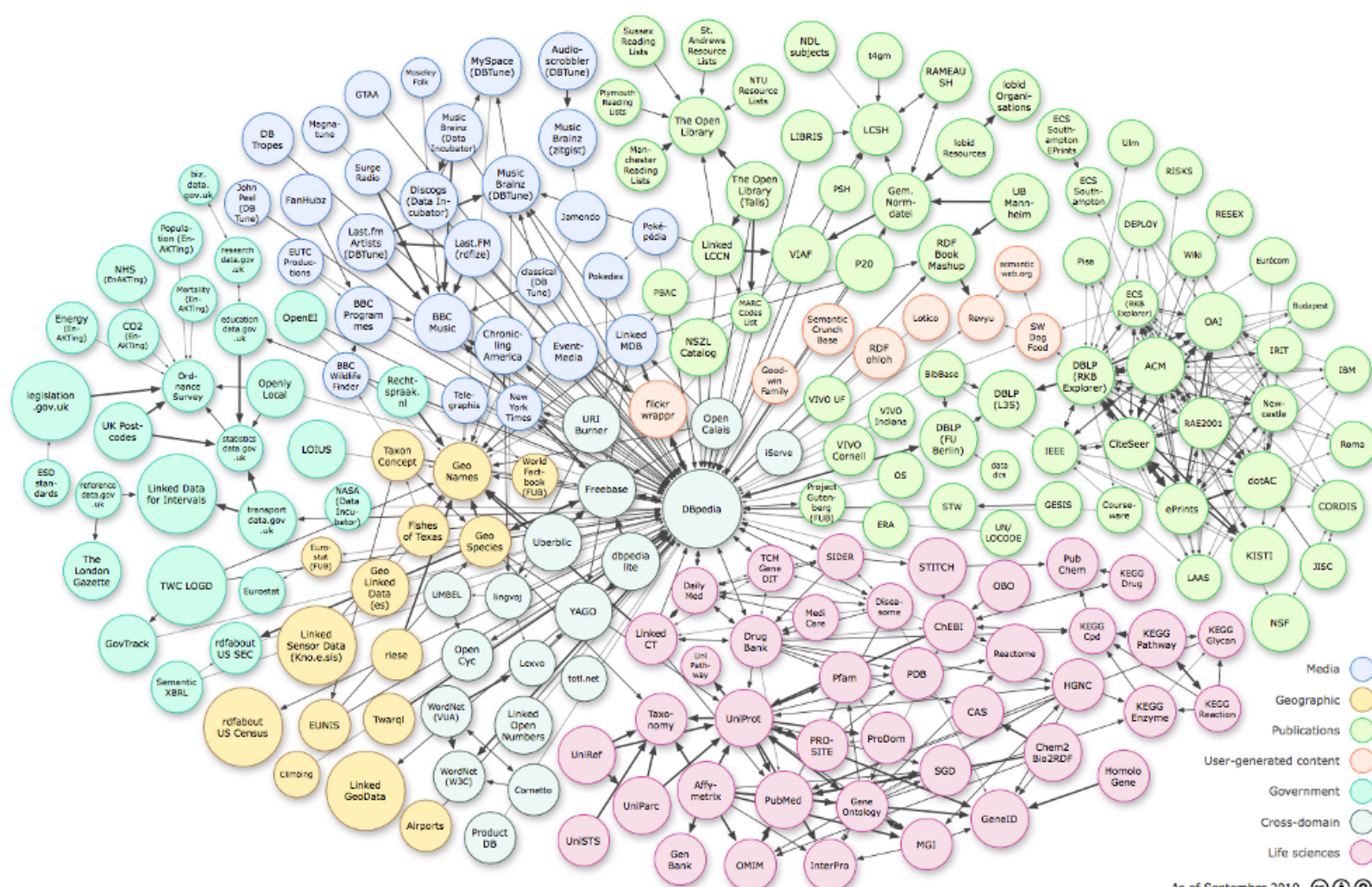
I 4 principi del Linked Data

1. Utilizzare URI (Uniform Resource Identifiers) per identificare entità
2. Usare HTTP URI in modo che le entità possano essere individuati da persone e da user agent (browser, programmi)
3. Fornire informazioni utili sull'oggetto quando qualcuno accede ad un URI usando formati standard come RDF, SPARQL
4. Includere collegamenti ad altri URI cosicché si possano scoprire maggiori informazioni

<http://www.w3.org/DesignIssues/LinkedData.html>

```

graph LR
    MusicBrainz --- Discogs
    MusicBrainz --- Geonames
    MusicBrainz --- PGP
    Discogs --- Geonames
    Discogs --- PGP
    Geonames --- US Census Data
    Geonames --- PGP
    Geonames --- DBLP
    US Census Data --- PGP
    Project Gutenberg --- PGP
    Project Gutenberg --- DBLP
    DBLP --- RDF Book Mashup
    PGP --- Bayou
    PGP --- RDF Book Mashup
    Bayou --- RDF Book Mashup
  
```

[illegible]

Linked Data Cloud
<http://lod-cloud.net/>

Web semantico e semantica del Web

Per estrarre i frammenti di dati strutturati da pagine HTML sono stati inventati una varietà di microformats

I Microformats possono essere utilizzati per pubblicare dati strutturati che descrivono tipi specifici di entità, come persone e organizzazioni, eventi, recensioni e voti, attraverso l'inclusione dei dati all'interno delle pagine HTML.

Riferimenti

T. Heath and C. Bizer. Linked Data: Evolving the Web into a Global Data Space.
Morgan & Claypool, 2011.
<http://linkeddatabook.com/book>

K. Janowicz, P.Hitzler, B. Adams, D. Kolas, C. Vardeman II, Five Stars of Linked Data
Vocabulary Use, Semantic Web, 5 (3), 2014, 173-176