

Semantic Web: standards and languages for knowledge representation

N. Aussenac-Gilles

IRIT- CNRS

aussenac@irit.fr

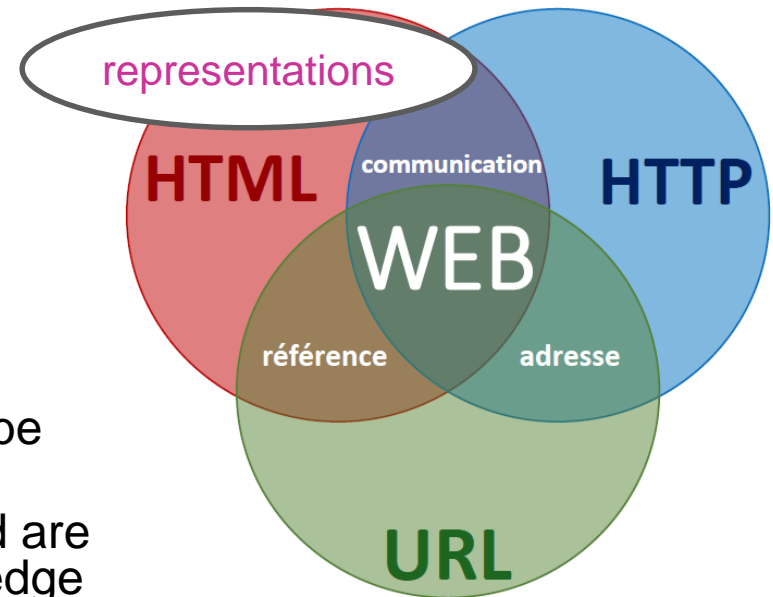
MELODI group

<http://www.irit.fr/-Equipe-MELODI->



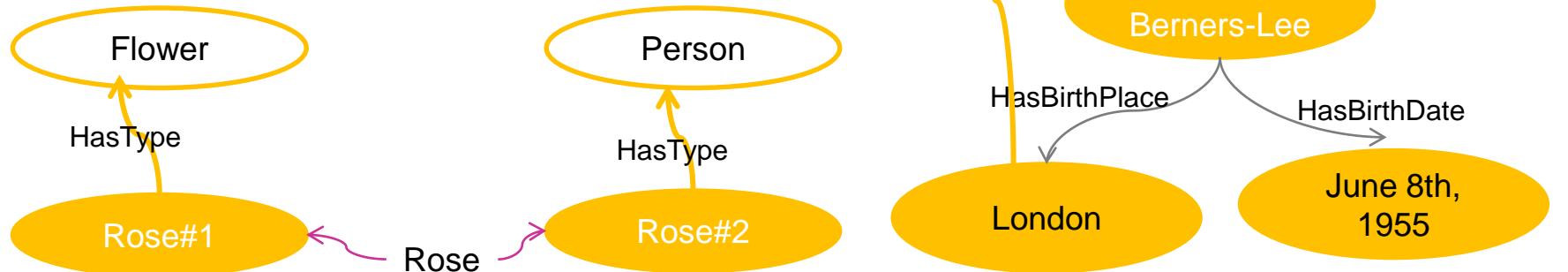
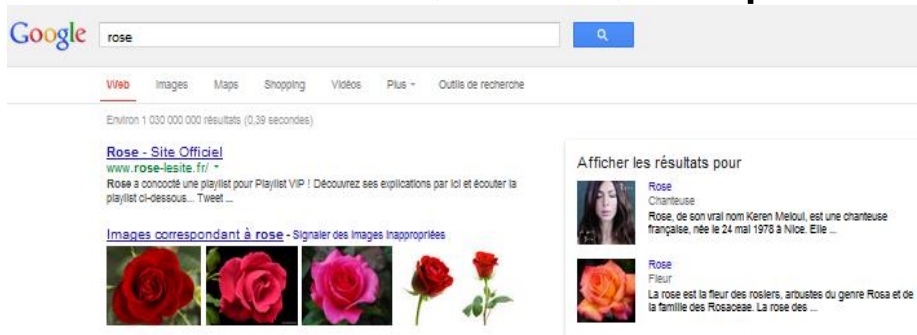
Main features of the last session

- The web, an infrastructure on top of internet
- The semantic web, a project to add a formal representation of the content of web documents and resources
- Important actors of the semantic Web
 - Tim Berners Lee
 - Ted Nelson
- The web relies on standard languages and format defined and promoted by the W3C
- Information retrieval, interoperability, data linking, reuse and sharing are expected to be made easier by semantics
- Various knowledge bases already exist and are used by search engines : DBPedia, Knowlledge vault, ...
- Knowledge can be described at 2 levels
 - Entities / instances
 - Sets of entities / classes or concepts



Example1: Google Knowledge Graph principles

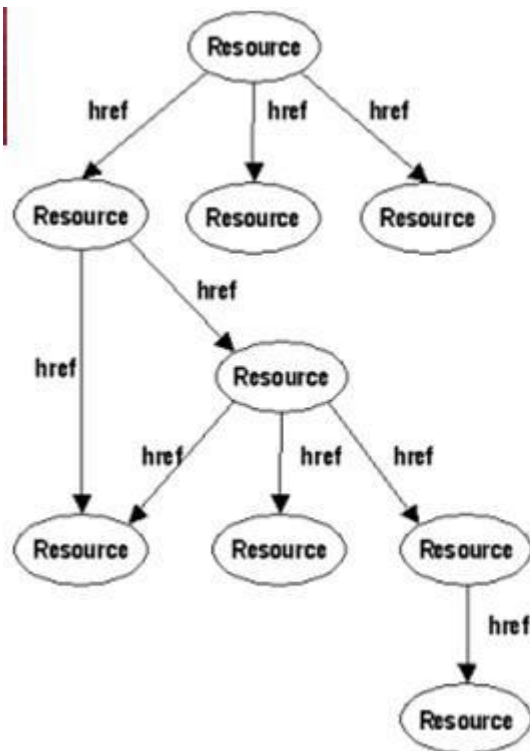
- Connect Google's indexes and structured web data
- Exploit the structure description of entites (persons, monuments, cities, frequent classes ...)



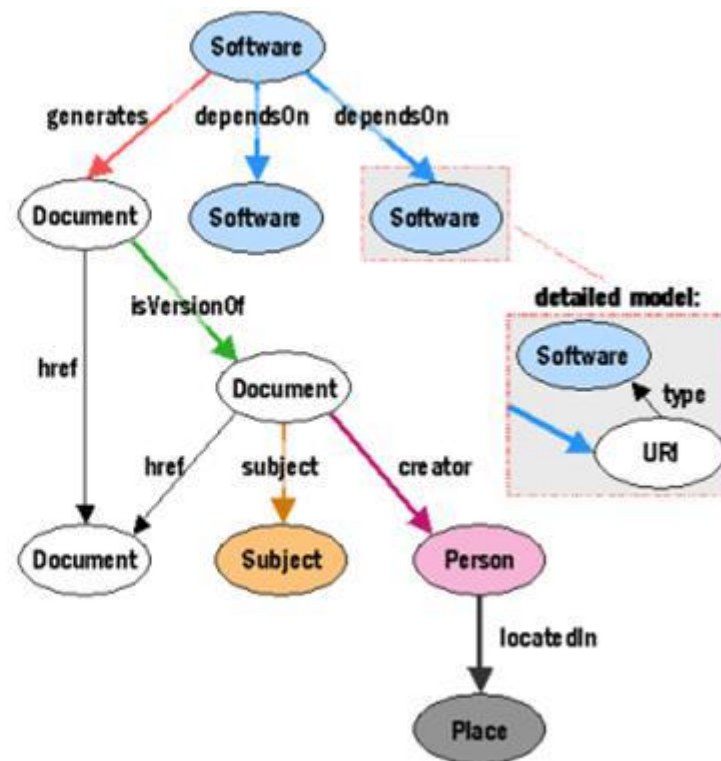
Example 2: better exploit the knowledge content of web pages

- From HTML links to “semantic” links and tags

Classical web



Semantic web



Example 2: html-spip source document

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01 Transitional//EN" "http://www.w3.org/TR/html4/loose.dtd">
<html>
<head> ... <title>IRIT - UMR5505 - Nathalie Aussenac-Gilles</title>
... </head>
<body>
<table> <tr> <td>Permanent - DR au CNRS</td> </tr>
<tr> <td><font class="intertitre_orange" > Equipe </font></td>
<td> <a href= "http://www.irit.fr/MELODI">MEthodes et ing&eacute;nierie des Langues, des Ontologies et du
    Discours</a> </td> </tr>
<tr> <td><font class="intertitre_orange">&nbsp;&nbsp;&nbsp;Contact</font></td>
    <td><a href= "mailto:aussenac@irit.fr">aussenac@irit.fr</a></td> </tr>
</table>
<h2><a name="Encadrement_encours"></a>Encadrement de th&egrave;ses en cours</h2>
<ul>
<p class="spip"><li> <a href = "http://www.irit.fr/annuaire\_detail.php3?code=5633">Julien Corman</a> </li>
<p class="spip"><li> <a href = "http://www.irit.fr/annuaire\_detail.php3?code=8224"> Jean-Philippe
    Fauconnier</a></li>
<p class="spip"><li> <a href = "http://www.irit.fr/-Annuaire-?lang=fr&code=8245">Laurent Sorin</a> </li>
<p class="spip"><li> <a href = "http://www.irit.fr/annuaire\_detail.php3?code=5633">Anis Tissaoui</a> :
    Annotation s&eacute;mantique de corpus dynamiques , depuis nov. 2008, soutenance pr&eacute;vue le 5 d&eacute;c.
    2013. ....</li>
</p> </ul>
```

Example 2: What the browser displays



Aussenac-Gilles Nathalie
Statut Permanent - DR au CNRS
Equipe Méthodes et ingénierie des Langues, des Ontologies et du Discours
Contact aussenac@irit.fr
Localisation IRIT2 / Niveau 2, Pièce: 275
Téléphone + 33 (0)5 61 55 82 93
Publications 

Recherche

Responsabilités au sein de l'IRIT

- Responsable de l'équipe MELODI
- Membre du Conseil scientifique de l'IRIT
- Coanimatrice de l'axe "Masse de données et calcul"

Encadrement

Encadrement de thèses en cours

- Julien Corman
- Jean-Philippe Fauconnier
- Laurent Sorin
- Anis Tissaoui : « Annotation sémantique de corpus dynamiques », depuis nov. 2008, directrice de thèse et co-encadrement avec P. Laublet, Lalic-Paris4 et N. Hernandez, IRIT-IC3. Université Toulouse 3, spécialité Informatique, école doctorale EDIT.

Example 2: What the search engine has access to

(les balises méta)

IRIT - UMR5505 - Nathalie Aussenac-Gilles

Permanent - DR au CNRS

Equipe <a href= "<http://www.irit.fr/MELODI>">MEthodes et ingénierie des Langues, des Ontologies et du Discours

Contact <a href= "<mailto:aussenac@irit.fr>">aussenac@irit.fr

<h2>Encadrement de thèses en cours</h2>

Julien Corman

 Jean-Philippe Fauconnier

<a href = "<http://www.irit.fr/-Annuaire-?lang=fr&code=8245>">Laurent Sorin

Anis Tissaoui :

Annotation sémantique de corpus dynamiques depuis nov. 2008, soutenance prévue le 5 déc. 2013. Directrice de thèse et co-encadrement avec P. Laublet, Lalic-Paris4 et N. Hernandez, IRIT-MELODI et UTM. Université; Toulouse 3, spécialité Informatique, école doctorale EDMITT.

Example 2: behind information is some knowledge

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01 Transitional//EN" "http://www.w3.org/TR/html4/loose.dtd">
```

<html>

My lab

<head>

<title>IRIT - UMR5505 - Nathalie Aussenac-Gilles</title>

... </head>

My name

<body>

Job

Permanent - DR au CNRS

| | **Equipe** |

My group

 [MEthodes et ingérie des Langues, des Ontologies et du Discours](http://www.irit.fr/MELODI) || |
 aussenac@irit.fr |

The groups'web page

</table>

[Encadrement de thèses en cours](#)

<p class="spip"> Julien Corman

<p class="spip"> Jean-Philippe Fauconnier

<p class="spip"> Laurent Sorin

[Anis Tissaoui](http://www.irit.fr/annuaire_detail.php3?code=5633) : Annotation sémantique de corpus dynamiques , depuis nov. 2008, soutenance prévue le 5 déc. 2013.

</p>

His web page

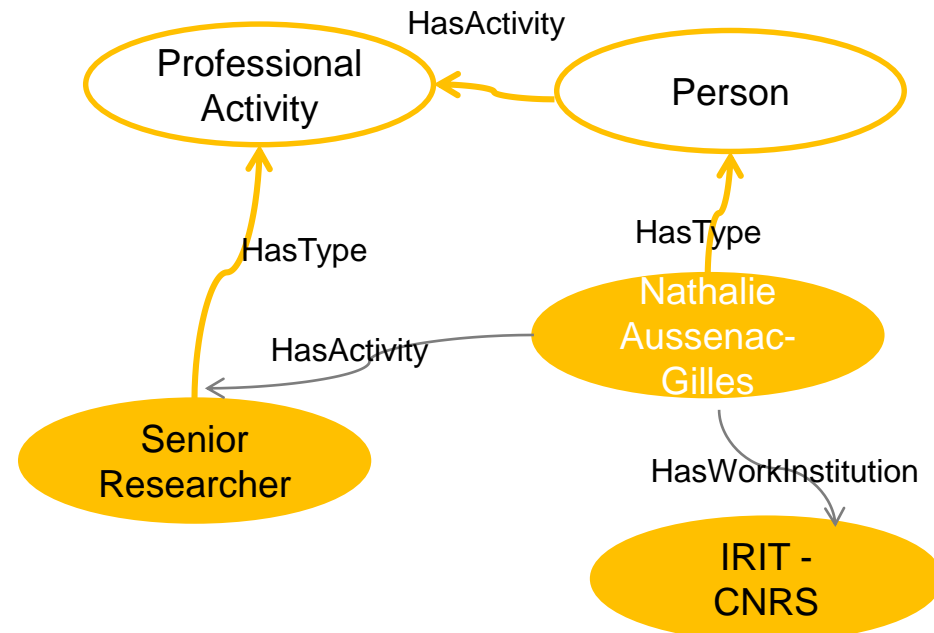
A PhD student under my supervision

Example 2:

- Natural language on web documents can be a knowledge source to build structured data
 - Using Natural Language Processing (NLP) to build structured (RDF) data

□ Either classes

□ Or instances



Example 2: From strings to things

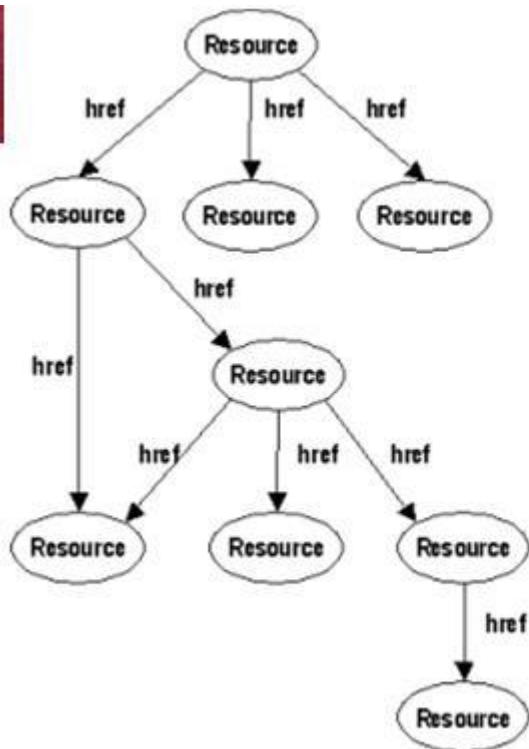
- Documents can be enriched with structured data
 - Semantic annotation
 - Inside the document itself : RDFa
 - Outside the document: locally or publically
- Structured data can be reused
 - Purposes: knowledge bases > K. discovery, web service combination, ...
 - Where
 - classes and their relations : ontologies
 - Instances: triple stores, ex DBPEDIA



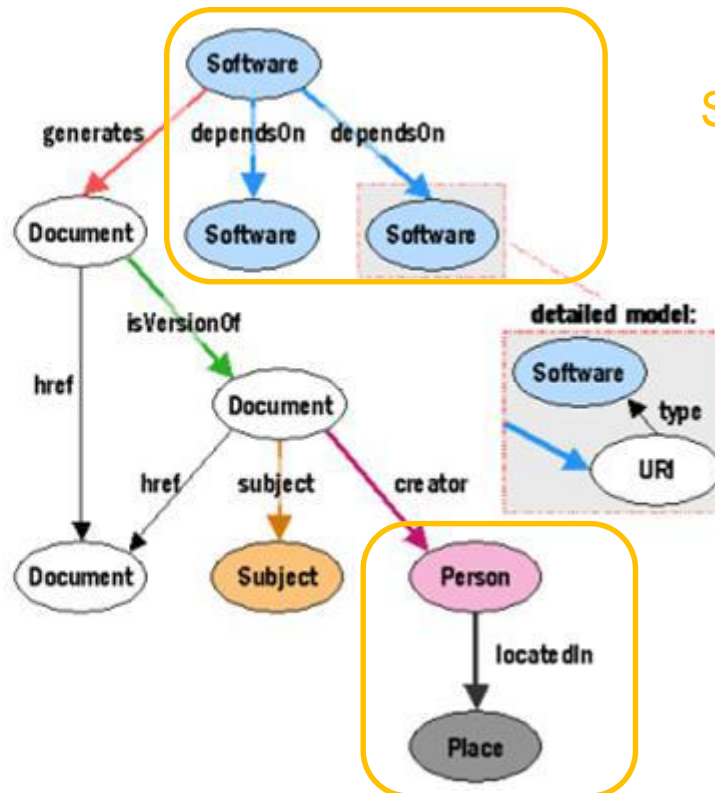
Example 2: Access to knowledge in web pages

- Ontologies as shared resources to assign types to data

Classical web



Semantic web



Ontologies

Software ontology

Foaf ontology



definitions

- Knowledge Graph
 - Knowledge base used by Google to improve information retrieval
 - Includes other knowledge bases like FreeBase and DbPedia
 - Any graph-based representation where nodes and edges are symbolic data
 - <https://arxiv.org/abs/2003.02320> « knowledge graphs » (survey)
- Knowledge vault
 - Knowledge base approach for information retrieval
 - Data integration from the knowledge graph, user's queries, information extracted from news
- Semantic annotation
 - Add « semantic » representations to documents as meta-data or tags
 - Requires a formal vocabulary or an ontology

Example 3 : Semantic annotation with a knowledge base

<http://labs.sparna.fr/> Thomas Francart

What is this page ? see [below](#), that [blog post](#), and more examples on [labs.sparna.fr](#).

Français ? rechargez la page [en français](#).

Horsemeat scandal: Scotland bans frozen beefburgers in schools

School kitchen in North Lanarkshire finds burger testing positive for horse DNA as Birds Eye withdraws some products

Schools in Scotland have been told not to serve frozen beefburgers, while in Wales supplies to schools in seven local authorities have been withdrawn.

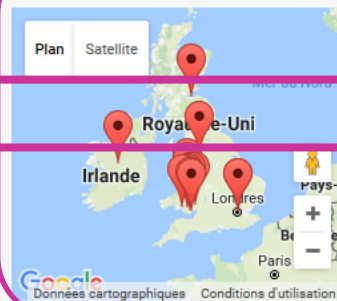
The move in Scotland came after a frozen beefburger from a school kitchen in North Lanarkshire tested positive for horse DNA while in Wales it was discovered that some burgers delivered to schools had been made at the Burger Manufacturing Company (BMC), in Builth Wells, Powys, one of the most recent producers to be caught up in the growing horsemeat scandal.

The developments came as Birds Eye withdrew three beef ready-meals from sale in the UK and Ireland as a precaution after 2% horse DNA was found in its chilli con carne made for the Belgian market.

Local authorities in Scotland were advised to "place a hold" on the use of the products following the discovery in a burger at a North Lanarkshire school kitchen.

The measure also applies to council leisure facilities and some social care establishments.

Where



Who

- Birds Eye (5)
- Richard Lochhead (1)
- Food Standards Agency (2)

What

- Scotland (4)
- North Lanarkshire (4)
- Wales (4)
- United Kingdom (1)
- Ireland (1)
- Monmouthshire (1)
- Blaenau Gwent (2)
- Merthyr Tydfil (1)
- Rhondda Cynon Taf (1)
- Caerphilly (2)
- Powys (2)
- Neath Port Talbot (1)
- England (2)
- Lancashire (1)
- Iglo (1)



```
<span rel="dc:subject"
resource="http://dbpedia.org/resource/Hamburger">
beefburgers</span>
```

What is this page ? see [below](#), that [blog post](#), and more examples on [labs.sparna.fr](#).

Français ? rechargez la page [en français](#).

Horsemeat scandal: Scotland bans frozen beefburgers in schools

School kitchen in North Lanarkshire finds burger testing positive for horse DNA as Birds Eye withdraws some products

Schools in Scotland have been told not to serve frozen beefburgers, while in Wales supplies to schools in seven local authorities have been withdrawn.

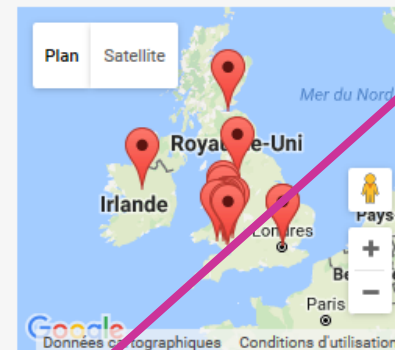
The move in Scotland came after a frozen beefburger from a school kitchen in North Lanarkshire tested positive for horse DNA while in Wales it was discovered that some burgers delivered to schools had been made at the Burger Manufacturing Company (BMC), in Builth Wells, Powys, one of the most recent producers to be caught up in the growing horsemeat scandal.

The developments came as Birds Eye withdrew three beef ready-meals from sale in the UK and Ireland as a precaution after 2% horse DNA was found in its chilli con carne made for the Belgian market.

Local authorities in Scotland were advised to "place a hold" on the use of the products following the discovery in a burger at a North Lanarkshire school kitchen.

The measure also applies to council leisure facilities and some social care establishments.

Where



Who

- [Birds Eye](#) (5)
- [Richard Lochhead](#) (1)
- [Food Standards Agency](#) (2)

What

- [Scotland](#) (4)
- [North Lanarkshire](#) (4)
- [Wales](#) (4)
- [United Kingdom](#) (1)
- [Ireland](#) (1)
- [Monmouthshire](#) (1)
- [Blaenau Gwent](#) (2)
- [Merthyr Tydfil](#) (1)
- [Rhondda Cynon Taf](#) (1)
- [Caerphilly](#) (2)
- [Powys](#) (2)
- [Neath Port Talbot](#) (1)
- [England](#) (2)
- [Lancashire](#) (1)
- [Iglo](#) (1)

Hamburger



A hamburger (also called a beef burger, sandwich, burger, hamburger or cheeseburger when served with a slice of cheese) is a sandwich consisting of one or

more cooked patties of ground meat, usually beef, placed inside a sliced bun. Hamburgers may be cooked in a variety of ways, including pan-frying, barbecuing, and flame-broiling. Hamburgers are often served with lettuce, bacon, tomato, onion, pickles, cheese and condiments such as mustard, mayonnaise, ketchup, relish, and chiles. The term "burger" can also be applied to the meat patty on its own, especially in the UK where the term "patty" is rarely used. The term may be prefixed with the type of meat or meat substitute used, as in "turkey burger", "bison burger", or "veggie burger". Hamburgers are sold at fast-food restaurants, diners, specialty restaurants (where burgers may sell for several times the cost of a burger) and high-end restaurants. There are many international and regional variations of the hamburger.

See [Wikipedia page](#)
See [DBpedia page](#)

<https://dbpedia.org/page/Hamburger>

About: Hamburger

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

Un hamburger, parfois hambourgeois (au Canada francophone) ou par aphérèse burger, est un sandwich d'origine allemande, composé de deux pains de forme ronde (bun) garnis de viande hachée (souvent du bœuf) et généralement de crudités — salade, tomate, oignon, cornichon (pickles) —, de fromage et de sauce. C'est un plat typique de la restauration rapide, emblématique de la cuisine américaine.

Property

Value

[dbo:abstract](#)

- Un hamburger, parfois hambourgeois (au Canada francophone) ou par aphérèse burger, est un sandwich d'origine allemande, composé de deux pains de forme ronde (bun) garnis de viande hachée (souvent du bœuf) et généralement de crudités — salade, tomate, oignon, cornichon (pickles) —, de fromage et de sauce. C'est un plat typique de la restauration rapide, emblématique de la cuisine américaine. (fr)
- A hamburger (also called a beef burger, sandwich, burger, hamburg or cheeseburger when served with a slice of cheese) is a sandwich consisting of one or more cooked patties of ground meat, usually beef, placed inside a sliced bun. Hamburgers may be cooked in a variety of ways, including pan-frying, barbecuing, and flame-broiling. Hamburgers are often served with lettuce, bacon, tomato, onion, pickles, cheese and condiments such as mustard, mayonnaise, ketchup, relish, and chiles. The term "burger" can also be applied to the meat patty on its own, especially in the UK where the term "patty" is rarely used. The term may be prefixed with the type of meat or meat substitute used, as in "turkey burger", "bison burger", or "veggie burger". Hamburgers are sold at fast-food restaurants, diners, specialty restaurants (where burgers may sell for several times the cost of a burger) and high-end restaurants. There are many international and regional variations of the hamburger. (en)

[dbo:country](#)

- [dbr:Germany](#)
- [dbr:United_States](#)

[dbo:ingredient](#)

- [dbr:Meat](#)
- [dbr:Bread](#)

Knowledge bases

- DBpedia <http://dbpedia.org/>

- Wikipedia categories as resources
- Wikipedia infoboxes to add properties to these resources
- 1 version for each language



- Yago

- Includes WordNet, DBpedia ...
- Knowledge extracted from definitions in the Wictionary etc.



HTML pages with structured data

- Approche générale google
 - <https://developers.google.com/search/docs/guides/intro-structured-data>
- Démonstrateur de
<http://labs.sparna.fr/schema.org/apres/index.html>
- Un outil pour extraire les données structurées à partir de pages web ;
<https://search.google.com/structured-data/testing-tool>



hamburger, burger, hambourgeois

Un hamburger, parfois hambourgeois ou par aphérèse burger, est un sandwich composé de deux pains de forme ronde garnis de viande hachée et généralement de crudités — salade, tomate, oignon, cornichon —, de

ID: 00009566n | Concept



Steak Salisbury, hamburger, hambourgeois

Le steak Salisbury est un plat composé de bœuf haché et d'autres ingrédients, modelés en forme de steak, généralement assorti d'une sauce brune.

ID: 00068946n | Concept



Steak haché, boulette, hamburger

Un steak haché est une viande hachée, souvent conditionnée en forme de galette prête à la cuisson.

ID: 00014608n | Concept



hamburger, bœuf haché

EN Beef that has been ground

Semantic web - 2

Example 4: BBC – a “real application”

NATURE WILDLIFE


Home | News | Features | Video collections | **Wildlife** | Prehistoric life | Places | Contact

Life | Animals | Mammals | Kangaroos and wallabies

Kangaroos and wallabies


This family of marsupials includes the kangaroos, tree kangaroos, wallabies and quokkas. When they need to move quickly, they adopt a bipedal hopping gait and use their tail for balance.

Scientific name: Macropodidae
Rank: Family
Common names: Macropods



10 The Abbott/Getty Images

Distribution



The shading illustrates the diversity of this group – the darker the colour the greater the number of species. Data provided by WWF's Wildfinder.

Classification

- Life
- Animals
- Vertebrates
- Mammals
- Kangaroos, possums and wallabies
- Kangaroos and wallabies

Share f t

Find wildlife

Search

BBC News about Kangaroos and wallabies

How the kangaroo got its bounce


The newly sequenced DNA of the kangaroo reveals some of the genetic secrets behind its remarkable biology.

- Kangaroo bounce mechanics snapped
- Australia's marsupials 'have American roots'
- Kidnap a joey to save a species

Elsewhere on the web

- Animal Diversity Web (animaldiversity.umnz.umnich.edu)
- Macropod fact sheet (bird.net.au)
- Fascinating facts about macropods (rootourism.com)
- Introduction to macropods (zipcodezoo.com)
- Macropod (wikipedia.org)

Explore this group



Tree-Kangaroos Kangaroos, wallabies and wallaroos Rock wallabies

Situation in a taxonomy

Close and similar animals

Additional informations

Example 4: BBC – “real ontologies”

- Ontologies and vocabularies for
 - Tagging videos and papers
 - Collecting additional information on the web
 - Providing synthetic information about wildlife, sports, news, etc.

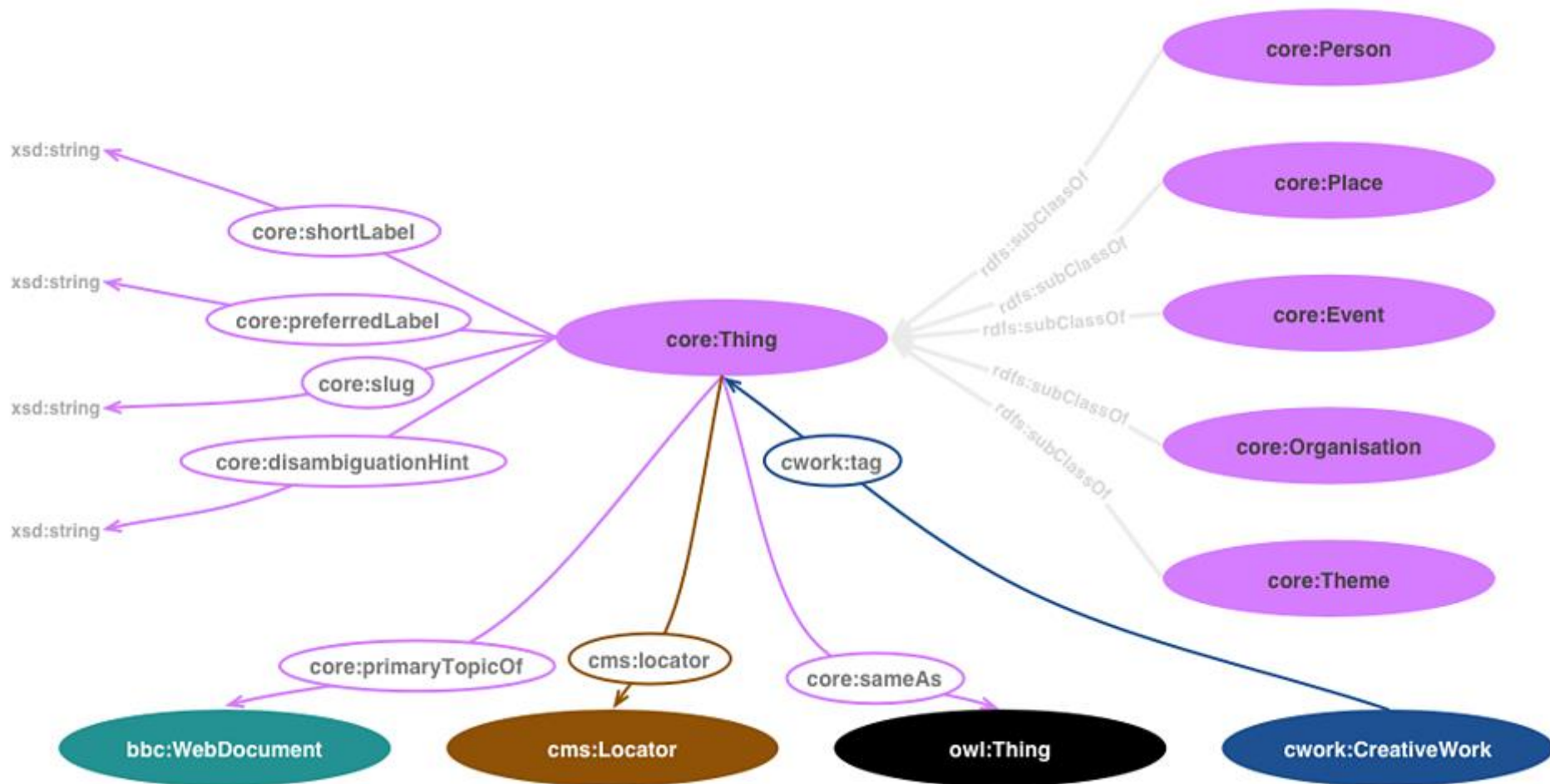
<http://www.bbc.co.uk/ontologies/bbc> sur le LOV http://liveschema.eu/dataset/lov_bbc

- Ontologies
 - 12 main ontologies
 - wildlife ontology
 - 6 reused ontologies

<https://www.bbc.co.uk/blogs/internet/entries/78d4a720-8796-30bd-830d-648de6fc9508>

<http://www.bbc.co.uk/ontologies/wildlife/2010-11-04.shtml>

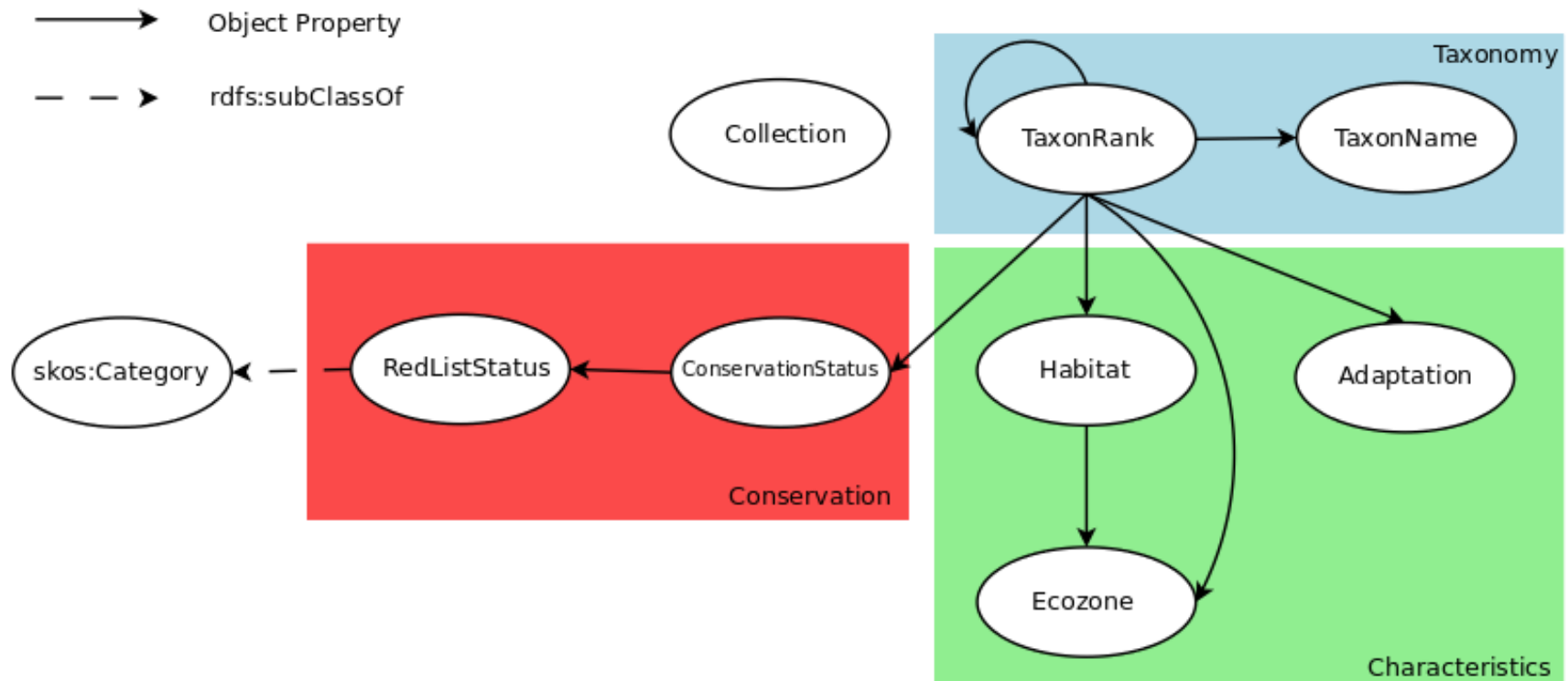
Ex 4 : BBC Core Concepts Ontology



Ex 4: BBC wild life ontology

Vocabulary Diagram

The following diagram illustrates the relationships between the key classes in the ontology. A number of classes of TaxonRank, Habitat and Adaptation have been omitted for clarity.



New needs

- Data are not just web pages ...
- ... but also
 - Data collected from data bases, repositories
 - Meta-data, multi-media documents
 - Lexical ressources, terminologies, thesaurus ...
- ... and also
 - Cookies, data collected from the social web
 - Open data from governments, scientific groups...
 - Data flows produce in real time

Internet of things

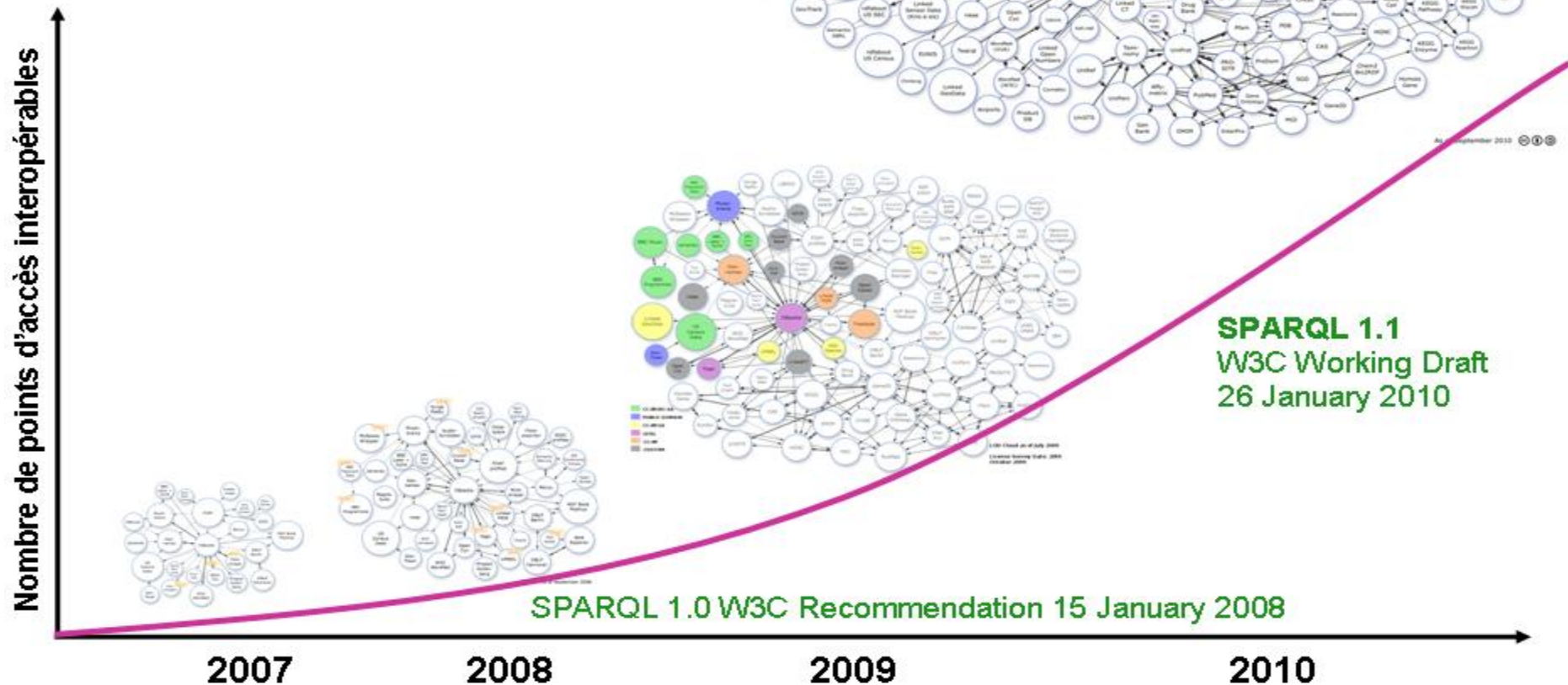


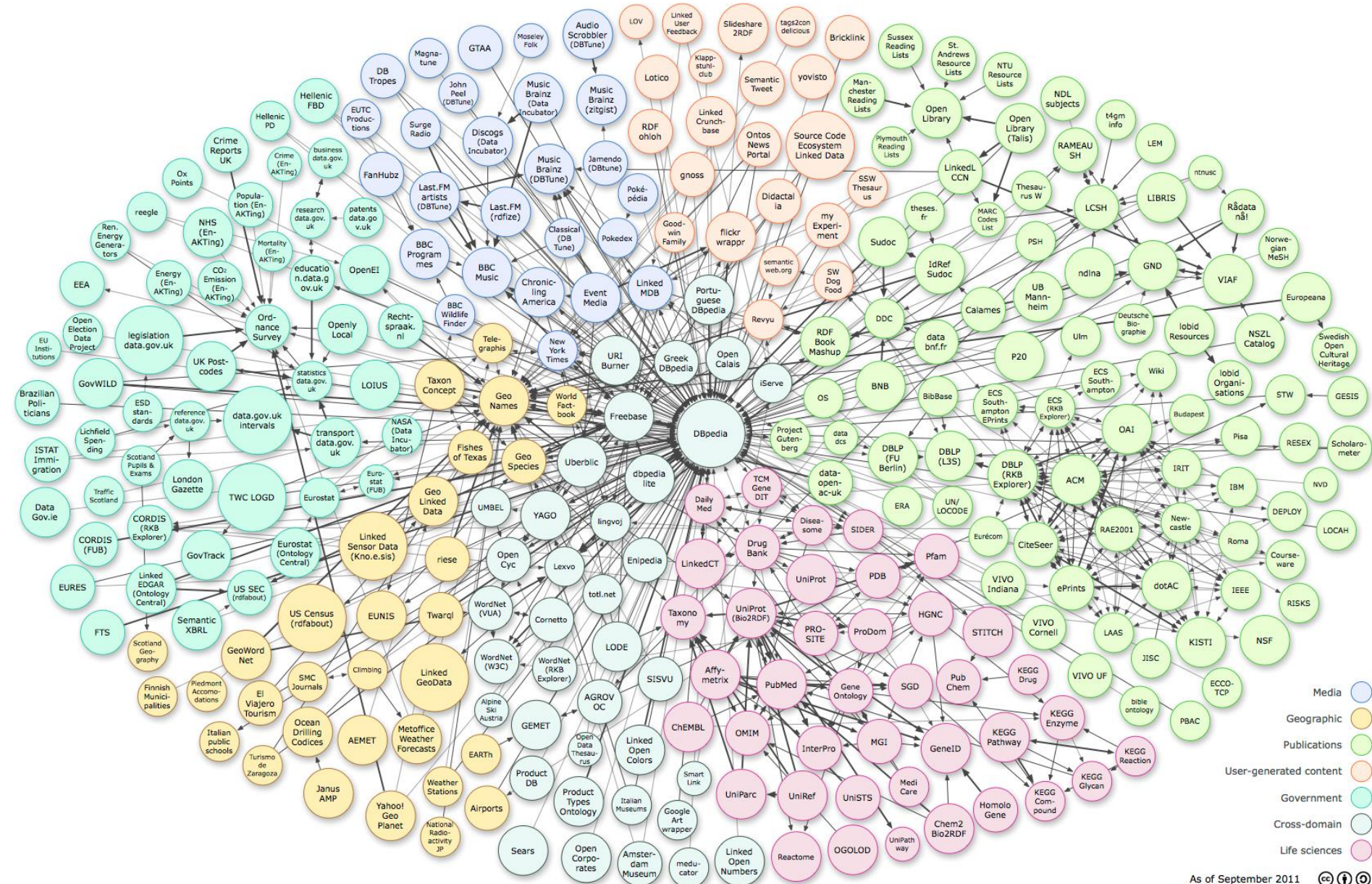
- Connecting sensors and any object with and RFID cheap
- *objets ayant des identités et des personnalités virtuelles, opérant dans des espaces intelligents et utilisant des interfaces intelligentes pour se connecter et communiquer au sein de contextes d'usages variés*
- Objects are identified with URIs
- Standardization of information exchange format
- Store and exploit context

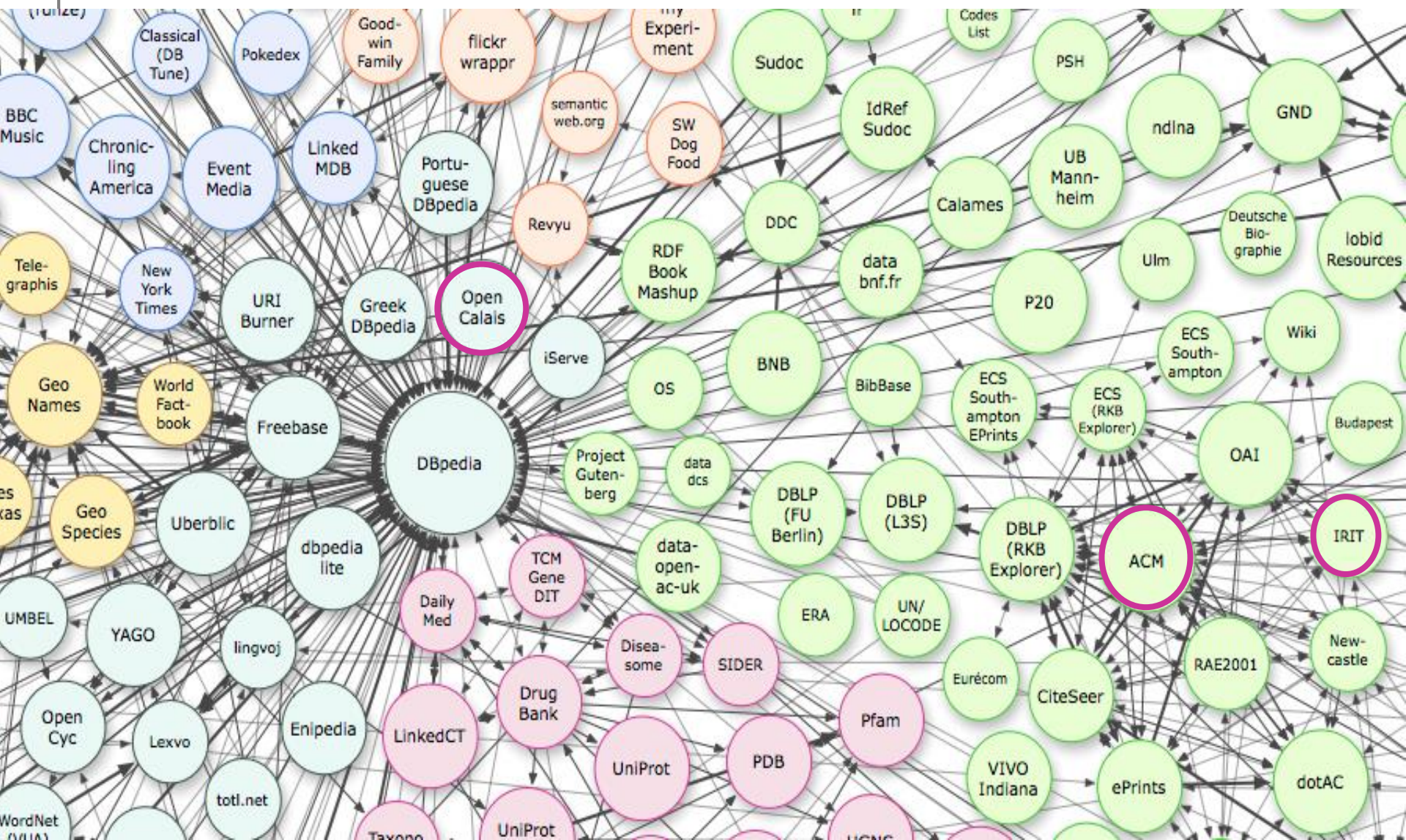
New needs

■ More and more data

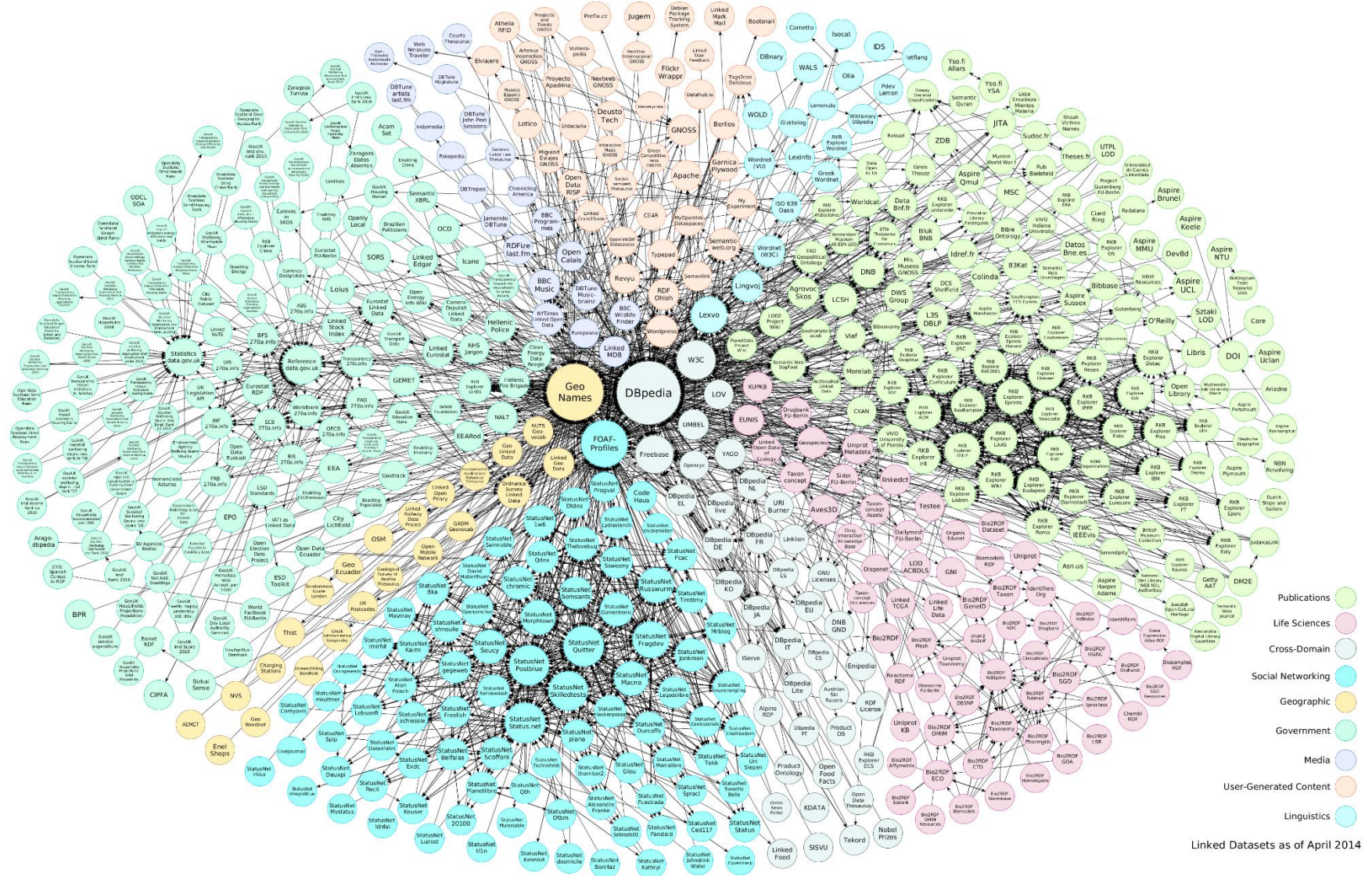
http://fr.wikipedia.org/wiki/Tim_Berners_Lee







<http://data.dws.informatik.uni-mannheim.de/lodcloud/2014/LinkedOpenDataCloud2014.png>



LOD 2021

- 2020 <https://www.lod-cloud.net/>
- Linked open research cloud
<https://linkedresearch.org/cloud>