

Jan Zemánek

jan.zemanek@gmail.com

# LINKED DATA

# Slidy na Webu

- ⦿ Dočasně na

- <http://keg.vse.cz/4iz440/linkedata.ppt>

# „Web of Documents“ (1)

## ⊙ Dnešní Web

- Webové stránky = dokumenty
- XHTML = XML = struktura dokumentů
- Dokumenty vzájemně „prolinkovány“ odkazy
- URL
- Odkazy
  - „the most powerful – feature of the Web“
  - „only one direction and untyped“
  - Evolving the Link - Danny Ayers
    - <http://tinyurl.com/5l4u4g>

# „Web of Documents“ (2)

- ⊙ Analogy
  - a global **filesystem**
- ⊙ Designed for
  - **human consumption**
- ⊙ Primary objects
  - **documents**
- ⊙ Links between
  - **documents (or sub-parts of)**
- ⊙ Degree of structure in objects
  - **fairly low**
- ⊙ Semantics of content and links
  - **Implicit**

# „The Web of Documents and the Web of Linked Data“

- ◎ „From global filesystem to global database“
- ◎ Databáze na Webu dnes
  - Přímě nepřístupná a vzájemně nepropojená „datová sila“



# „Web of Linked Data“ (1)

- ⊙ Strukturovaná RDF data vzájemně „prolinkována“ odkazy
- ⊙ Koncept **Linked Data**
  - „zažehnuto“ iniciativou **Linking Open Data**
    - <http://esw.w3.org/topic/SweoIG/TaskForces/CommunityProjects/LinkingOpenData>

# Web of Linked Data (2)

- ⊙ Analogy
  - a global **database**
- ⊙ Designed for
  - **machines first, humans later**
- ⊙ Primary objects
  - **things (or descriptions of things)**
- ⊙ Links between
  - **things**
- ⊙ Degree of structure in (descriptions of) things
  - **high**
- ⊙ Semantics of content and links
  - **explicit**

# Linking Open Data project

- Překonat „chicken-egg“ problém na Sémantickém webu
  1. protože nikdo nevystavuje data, neexistují aplikace, které by je konzumovaly
  2. protože neexistují aplikace, nikdo data nevystavuje








# Linked Data

- ⦿ „a method of
  - **exposing,**
  - **sharing,**
  - and **connecting**  
pieces of data on the Semantic Web“
- ⦿ „emphasizes **Web access to data** using existing Web technologies such as **URIs** and **HTTP**“
- ⦿ „emphasizes **links** between related Web **resources**“

# Ukázka

▼ **Timothy Berners-Lee**   

Type ▶ ● Male  
▶ <http://xmlns.com/foaf/0.1/Person>



Label ○ **Tim Berners-Lee**

Assistant ▶ ● Amy van der Hiel

Home Page ▶ ● <http://www.w3.org/People/Berners-Lee/>

Office ▶ ...

Organization

▼ **World Wide Web Consortium**  


Type ▶ ● Social Entity  
▶ ● Organization

Label W3C

Public Home Page ▶ ● <http://www.w3.org/>

Part ▶ Architecture Domain  
▶ Interaction Domain  
▶ Technology and Society Domain  
▶ Ubiquitous Web Domain  
▶ Web Accessibility Initiative  
▶ Incubator, TAG, AB

Homepage ▶ ● <http://www.w3.org/>

Logo 

Name World Wide Web Consortium

Is organization of ▶ ● Timothy Berners-Lee


Public Home Page ▶ ● Tim Berners-Lee  
▶ ...

Based near ▶ ...

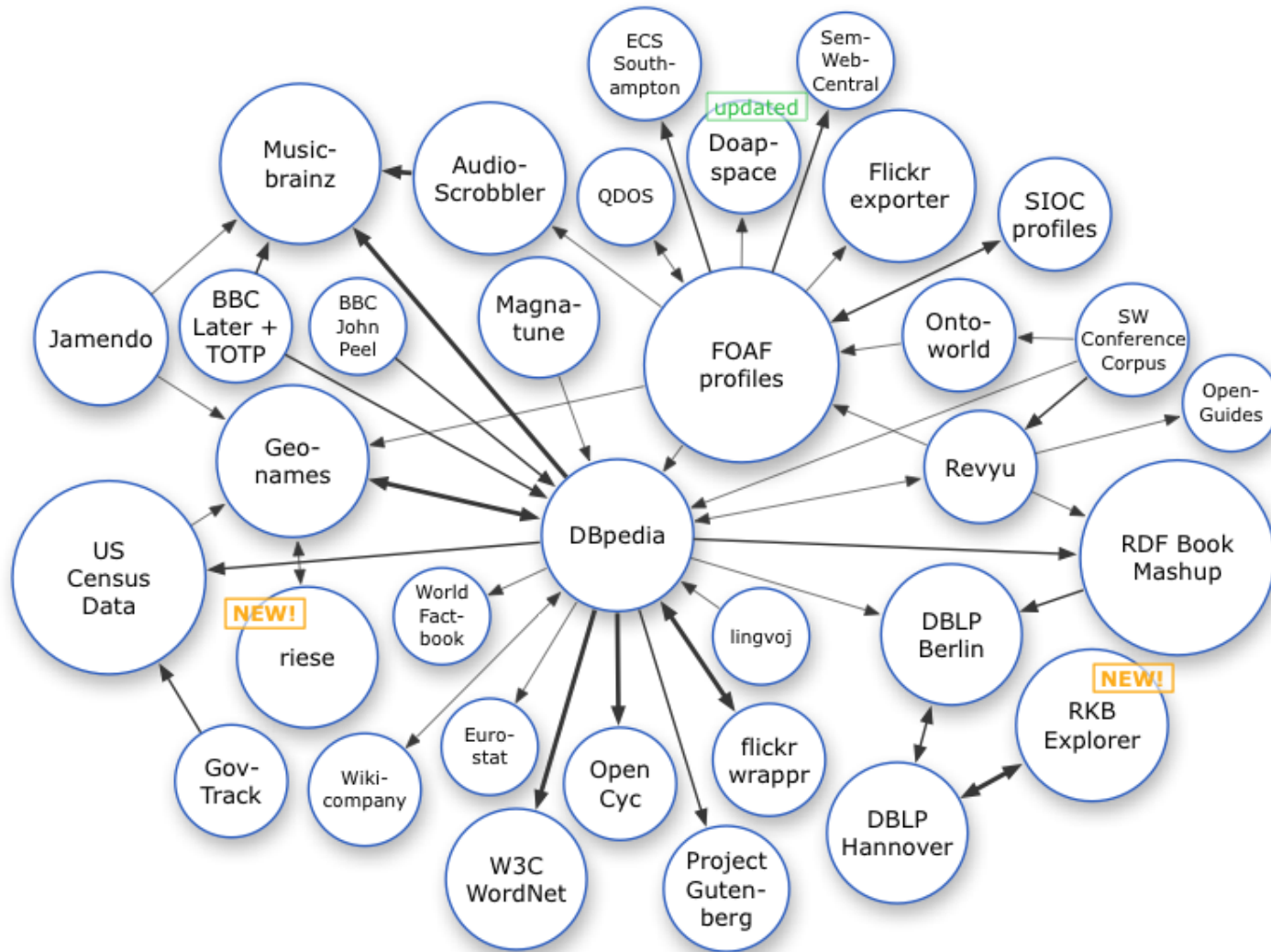
Family name Berners-Lee

Givenname Timothy

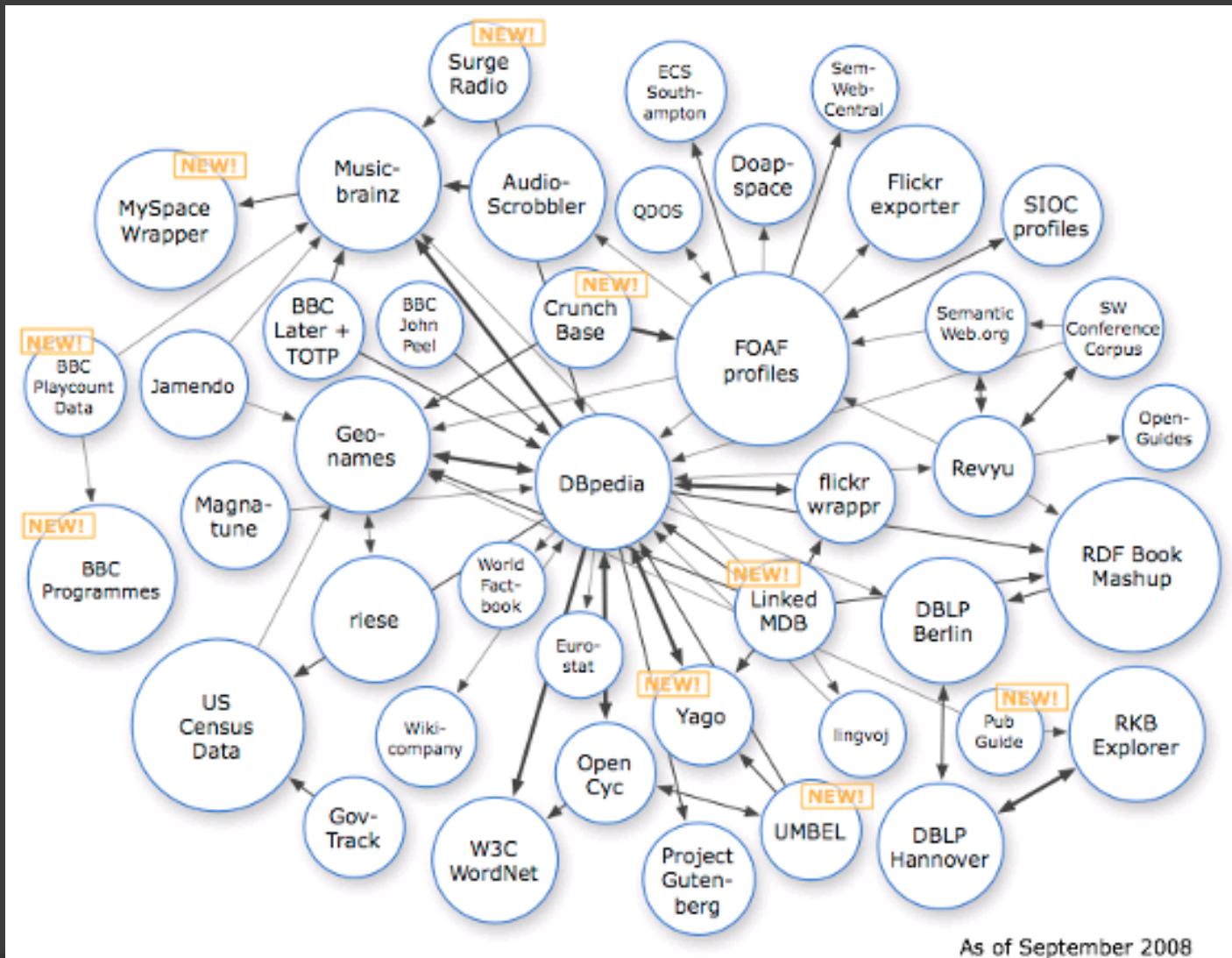
Homepage ▶ ● <http://www.w3.org/People/Berners-Lee/>

Img 

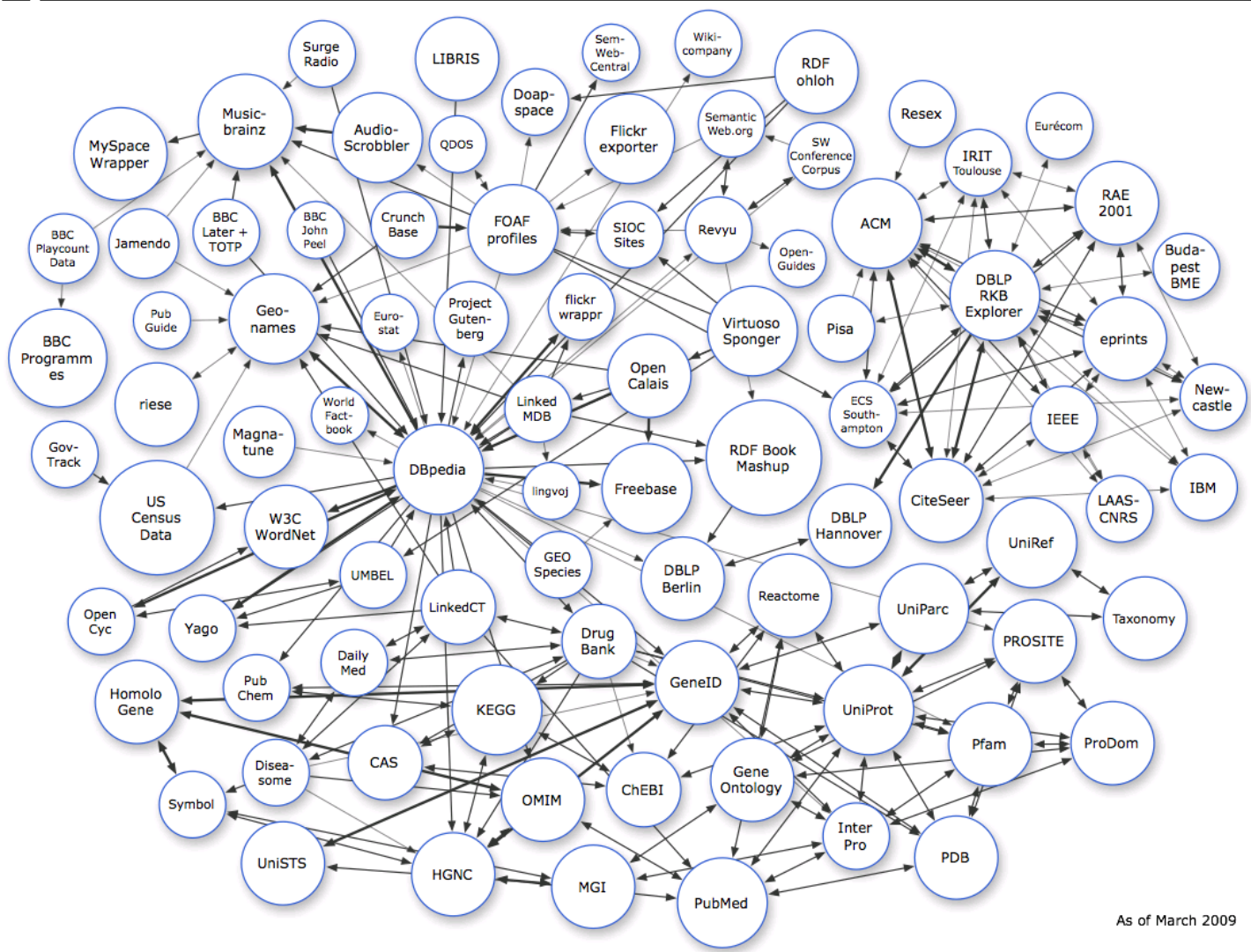
# Linked Data Map (Spring '08)



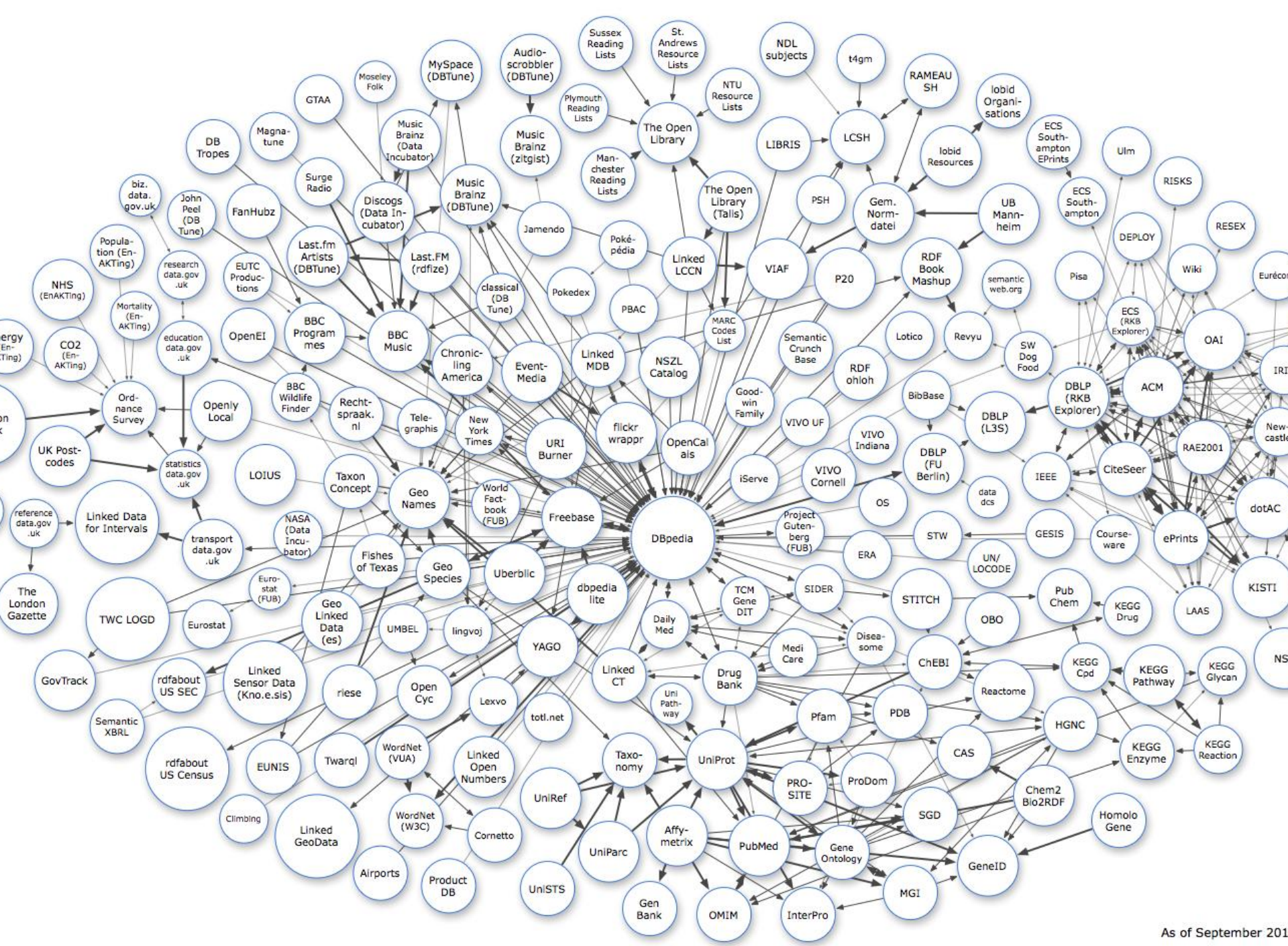
# Linked Data Map (Aut. '08)



# Linked Data Map (Spring '09)



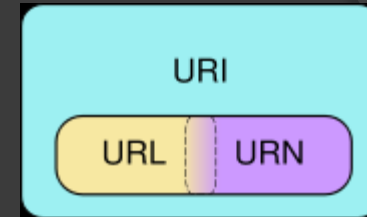




# 4 principy pro Linked Data

- ④ Use **URIs** to identify things that you expose to the Web as **resources**
- ④ Use **HTTP URIs** so that people can locate and look up (**dereference**) these things
- ④ Provide useful information about the **resource** when its **URI** is **dereferenced**
- ④ Include links to other, related **URIs** in the exposed data as a means of improving information discovery on the Web
- ④ [www.w3.org/DesignIssues/LinkedData.html](http://www.w3.org/DesignIssues/LinkedData.html)

# URI



- ⦿ Uniform Resource Identifier
- ⦿ „a URI is probably the most important part of the Web’s architecture“
- ⦿ URI
  - Identifikují webové stránky
  - Mohou identifikovat cokoli
    - „Zdroje“ (*resources*)
      - Osoby, věci, místa, abstraktní koncepty atd.
- ⦿ „A *resource* is simply something that can be identified with a URI.“



# Zdroje (Resources)

- ⦿ „Informační zdroje“
  - „**information resources**“
  - „things that can be expressed entirely as digital messages“
    - Webové stránky, PDF dokumenty, obrázky, atd.
- ⦿ „Neinformační zdroje“
  - „**non-information resources**“
  - V tomto pojetí je např. kniha neinformační zdroj 😊
  - Osoby, věci, místa, atd.

# „Dereferencování“ URIs

- ⦿ Information resource
  - A resource **with** representation (page, document, image)
  - Server should return **HTTP 200 OK**
- ⦿ Non-information resource
  - A resource **without** representation (vocabulary terms, human beings, etc.)
  - Server should return **HTTP 303 See Other**, and *Location* of related information resource headers
- ⦿ Demo
  - Non-information resource: <http://keg.vse.cz/resource/person/jan-zemanek>
  - Information resource: <http://keg.vse.cz/person/jan-zemanek>

# Browsing Linked Data

## ◎ The Tabulator

- „A generic RDF data browser and editor“
- Firefox extension
- <http://www.w3.org/2005/ajar/tab>

## ◎ Disco Hyperdata Browser

- „A simple browser for navigating the Semantic Web“
- Web-based
- [http://www4.wiwiss.fu-berlin.de/rdf\\_browser/](http://www4.wiwiss.fu-berlin.de/rdf_browser/)

# Zdroje

- ◉ <http://linkeddata.org/>
- ◉ Linked Data (ESW W3C)
  - <http://esw.w3.org/topic/LinkedData>
- ◉ Linked Data
  - [http://en.wikipedia.org/wiki/Linked\\_Data](http://en.wikipedia.org/wiki/Linked_Data)
- ◉ Howto publish Linked Data on the Web
  - <http://www4.wiwiss.fu-berlin.de/bizer/pub/LinkedDataTutorial/>
- ◉ Bootstrapping the Web of Data (slidy)
  - <http://linkeddata.org/slides/2008-02-amsterdam-catch.pdf>
- ◉ Linked Data – The essentials for Semantic Web application development (slidy)
  - [http://keg.vse.cz/\\_slides/2008-linked-data.pdf](http://keg.vse.cz/_slides/2008-linked-data.pdf)