

nemo

ontology & conceptual
modeling research group



Desenvolvimento Web

Ciclo de palestras de
Introdução à Computação

Vítor E. Silva Souza

(vitor.souza@ufes.br)

<http://www.inf.ufes.br/~vitorsouza>

Departamento de Informática

Centro Tecnológico

Universidade Federal do Espírito Santo



O que é uma aplicação Web?

- Um sistema distribuído;
- Acesso via navegador Web:
 - HyperText Transfer Protocol (HTTP);
 - HTML, JavaScript, CSS, ...;
 - Software cliente muito popular (ubíquo).
- Similar às aplicações comuns (não-Web):
 - Possuem interface com usuário, regras de negócio, dados, etc.;
 - Podem variar de um “Hello, world!” até sistemas muito complexos, milhões de usuários e LOCs.

De acordo com <http://www.informationisbeautiful.net/visualizations/million-lines-of-code/>, Facebook tem aprox. 60 MLOCs (incluindo backend)

Começou com a WWW

- Criada em 1989 como um sistema de comunicação para o CERN (European Organization for Nuclear Research).

```
GET /index.html HTTP/1.0  
Host: www.site.com  
[...]
```



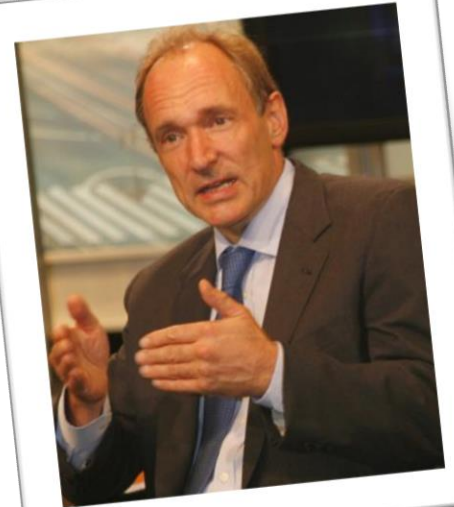
Requisição HTTP

Resposta HTTP



Procura por index.html, adiciona cabeçalhos e devolve.

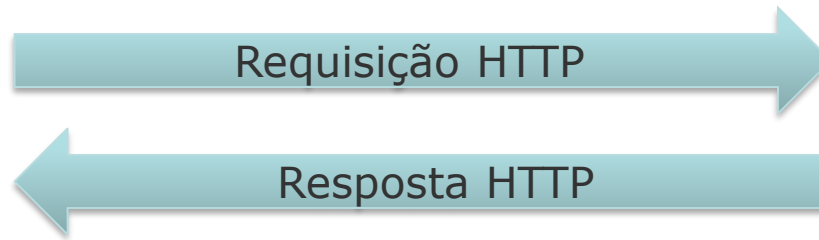
```
HTTP/1.1 200 OK  
Date: Fri, 15 Apr 2005 22:12:30 GMT  
Server: Apache/1.3.26 [...]  
Last-Modified: Wed, 23 Mar 2005 00:43:22 GMT  
Content-Length: 11379  
Content-Type: text/html  
[...]
```



Tim Berners-Lee,
criador da WWW

Ao invés de um arquivo, um programa

```
GET /index.cgi HTTP/1.0  
Host: www.site.com  
[...]
```



- 1: Roda um programa associado com index.cgi;
- 2: Programa retorna conteúdo em HTML;
- 3: Adiciona cabeçalhos e devolve.

```
HTTP/1.1 200 OK  
Date: Fri, 15 Apr 2005 22:12:30 GMT  
Server: Apache/1.3.26 [...]  
Last-Modified: Wed, 23 Mar 2005 00:43:22 GMT  
Content-Length: 11379  
Content-Type: text/html  
[...]
```

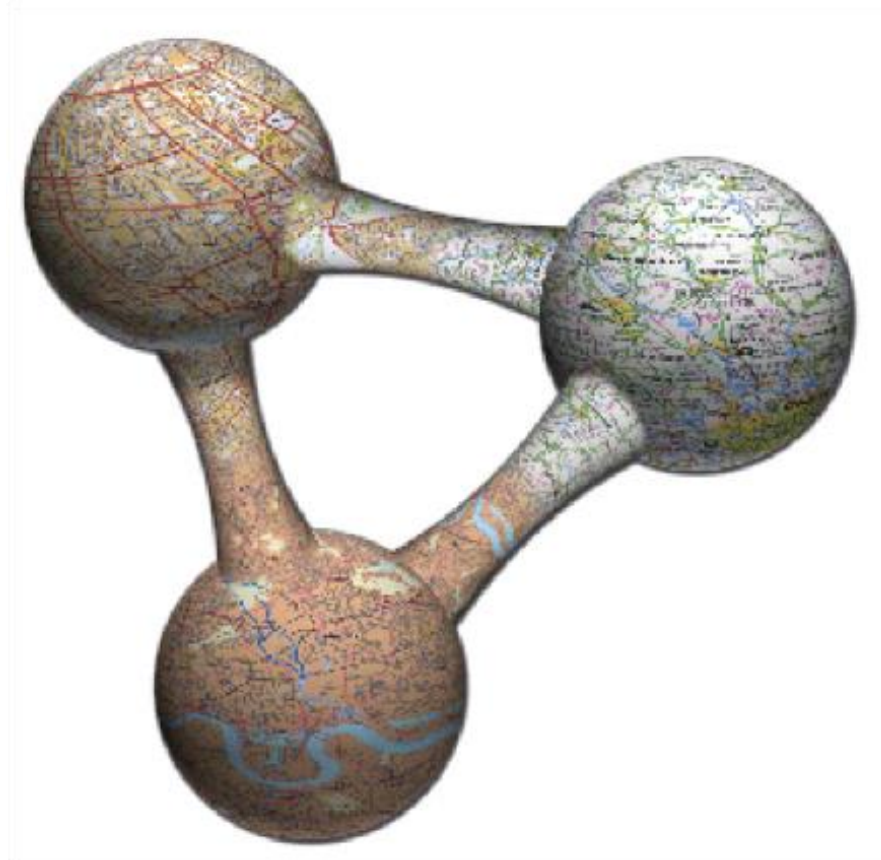
- Desenvolvimento Web consiste em escrever programas que:
 - Respondem à requisições usando HTTP;
 - Produzem resultados em uma linguagem compatível com um navegador Web (não só HTML!).
- Muitos componentes envolvidos:
 - O servidor Web;
 - Páginas Web, folhas de estilo, scripts, imagens, etc.
 - Código em uma linguagem de programação;
 - APIs, frameworks e bibliotecas existentes.

- 1993: CGI – Common Gateway Interface – C, C++, Fortran, Perl, etc.;
- 1994: Macromedia Coldfusion, PHP;
- 1995: Microsoft ASP;
- 1996: Java Servlets;
- 1999: JavaServer Pages (JSP);
- ...
- Hoje: aplicações ricas
 - AJAX;
 - Flash ActionScript;
 - HTML5, etc.



- 1995: lançamento oficial, no começo havia as Applets;
- 1996: Java Servlets;
- 1999: JavaServer Pages (JSP);
- 1999: plataforma J2EE (1.3 em 2001, 1.4 em 2003);
- 2000: Apache Struts, framework MVC pioneiro;
- 2004: JavaServer Faces;
- 2006: Java EE 5;
- 2009: Java EE 6;
- 2013: Java EE 7.





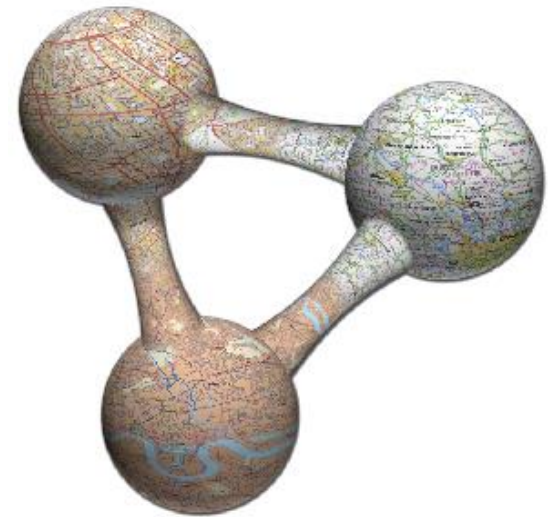
A Web Semântica

Novas possibilidades com dados abertos interligados.

- Novos dados são publicados todos os dias;
- Benefícios no consumo desses dados:
 - Amazon: dados de produtos disponíveis para terceiros, criando um ecossistema de afiliados;
 - Google/Yahoo!: consomem dados de vários sites e provêem resultados de busca melhores;
 - Projeto Genoma Humano: cooperação entre cientistas com troca de dados de pesquisa;
 - theyworkforyou.com: eleitores do Reino Unido podem avaliar o desempenho de seus representantes.

- Que forma de acesso a esses dados poderia facilitar ao máximo a reutilização dos mesmos?
- Como permitir que novos dados relevantes sejam descobertos no meio de tantos conjuntos de dados?
- Como fazer com que aplicações integrem dados de diferentes fontes, até mesmo inicialmente desconhecidas?

Linked Data



- Estrutura:
 - HTML estrutura texto, não dados;
 - Mas até o HTML está ficando mais estruturado...

```
<!-- HTML 4 -->

<html>
  <head>
    <title></title>
  </head>
  <body>
    <div>
      <div> </div>
      <div> </div>
      <div> </div>
      <div> </div>
    </div>
  </body>
</html>
```



```
<!-- HTML 5 -->

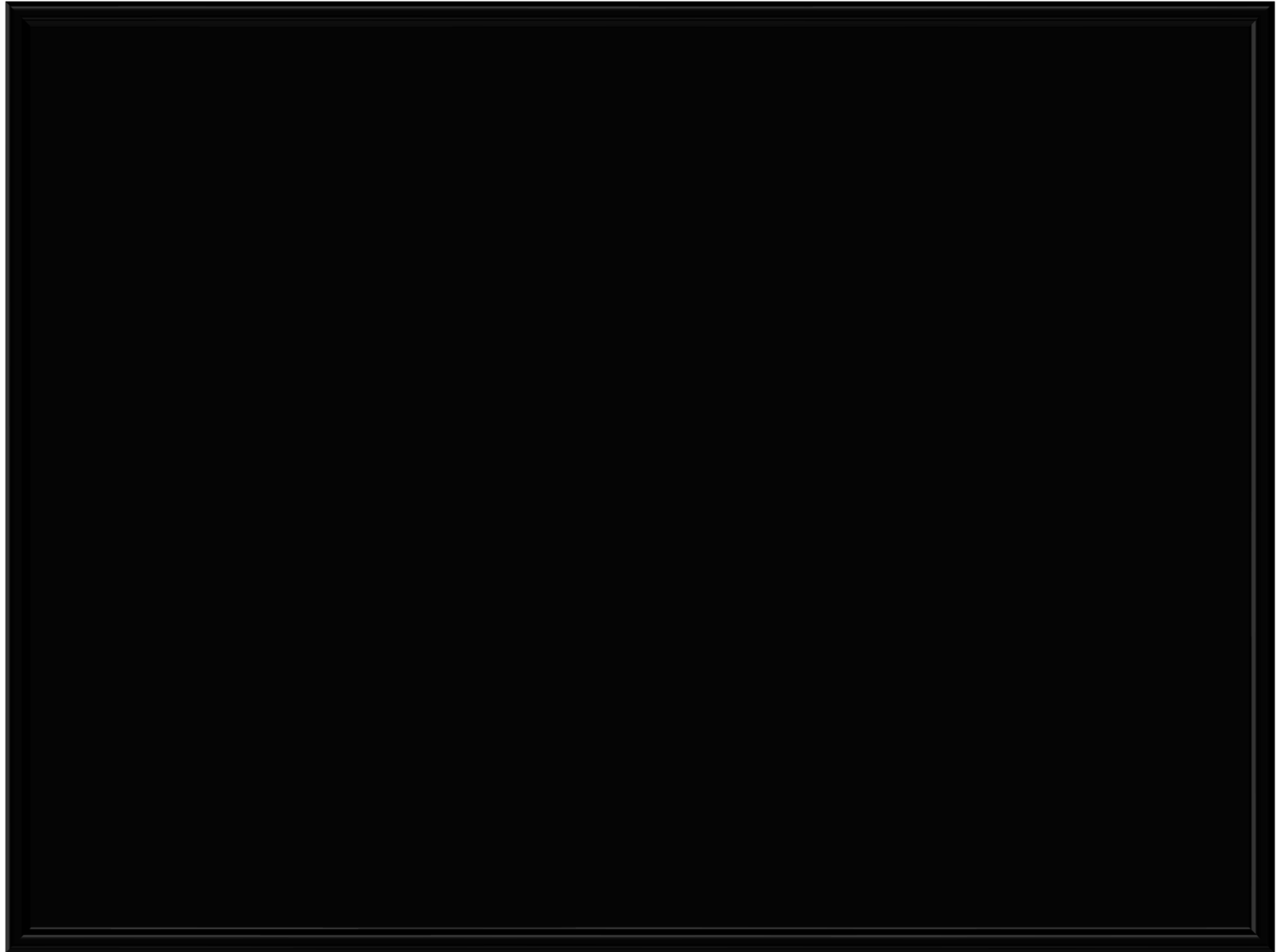
<html>
  <head>
    <title></title>
  </head>
  <body>
    <header> </header>
    <nav> </nav>
    <article> </article>
    <footer> </footer>
  </body>
</html>
```

- A proposta não é nova, já haviam algumas antes:
 - Microformatos (microformats.org): pequenos conjuntos de dados, com expressividade limitada;
 - APIs Web (programmableweb.com): XML, JSON, serviços REST... Não há padronização, maior esforço de integração.
- XML, JSON, etc. não têm algo que HTML tem há muito tempo: **hiperlinks!**

Humanos navegam,
softwares "rastejam"

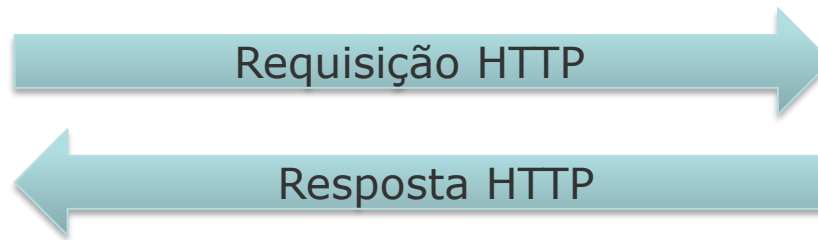


Linked Open Data (LOD)



Fonte: <http://en.wikipedia.org/wiki/File:Linked-open-data-Europeana-video.ogv>

WWW (Sintática)	Web Semântica
Documentos possuem URIs	Coisas (conceitos) possuem URIs
HTTP como mecanismo de acesso	Usa o mesmo mecanismo, permitindo que URIs sejam consultadas
HTML como formato padrão (importante para escalabilidade)	RDF como formato padrão
Links HTML conectam documentos e não possuem tipo	Links RDF conectam qualquer coisa e possuem tipos (moraEm e trabalhaEm entre Pessoa e Lugar)
Espaço global de informação (documentos interconectados)	Espaço global de dados (conceitos interconectados)



Conceitos possuem URIs

<http://biglynx.co.uk/people/matt-briggs>

<http://biglynx.co.uk/people/scott-miller>

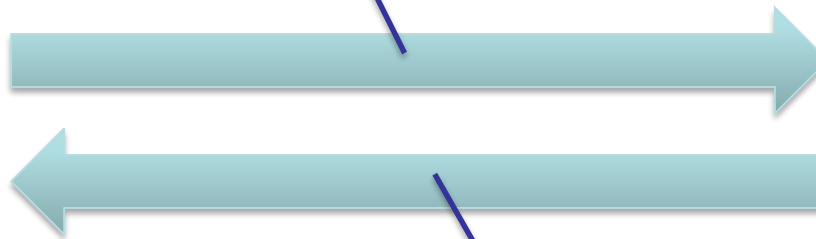
<http://xmlns.com/foaf/0.1/knows>

<http://biglynx.co.uk/people/linda-meyer>



Consumo de dados por software (1)

```
GET /people/matt-briggs HTTP/1.1  
Host: biglynx.co.uk  
Accept: text/html;q=0.5, application/rdf+xml
```



```
1 HTTP/1.1 303 See Other  
2 Location: http://biglynx.co.uk/people/matt-briggs.rdf  
3 Vary: Accept
```



Matt Briggs



matt-briggs



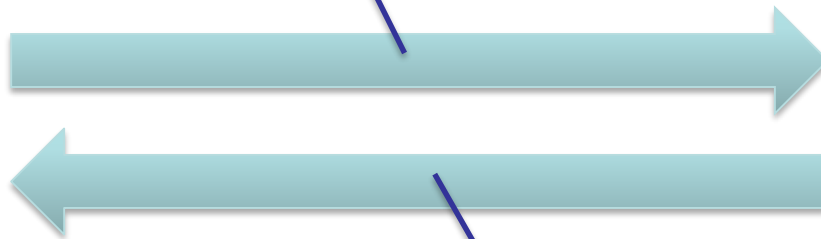
matt-briggs.rdf



matt-briggs.html

Consumo de dados por software (1)

```
GET /people/matt-briggs.rdf HTTP/1.1  
Host: biglynx.co.uk  
Accept: text/html;q=0.5, application/rdf+xml
```



```
HTTP/1.1 200 OK  
Content-Type: application/rdf+xml  
  
<?xml version="1.0" encoding="UTF-8"?>  
<rdf:RDF ...
```



Matt Briggs



matt-briggs

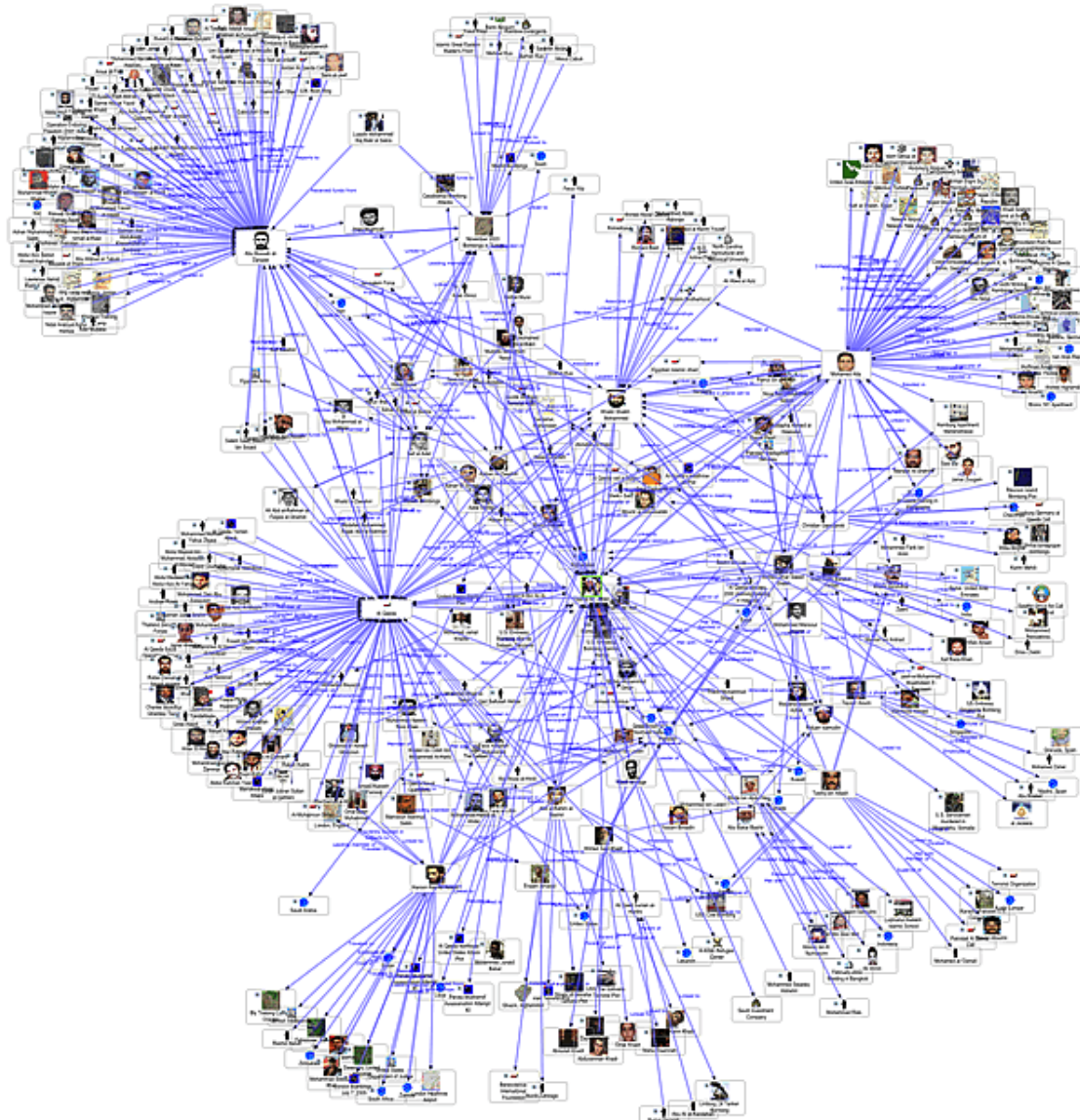


matt-briggs.rdf



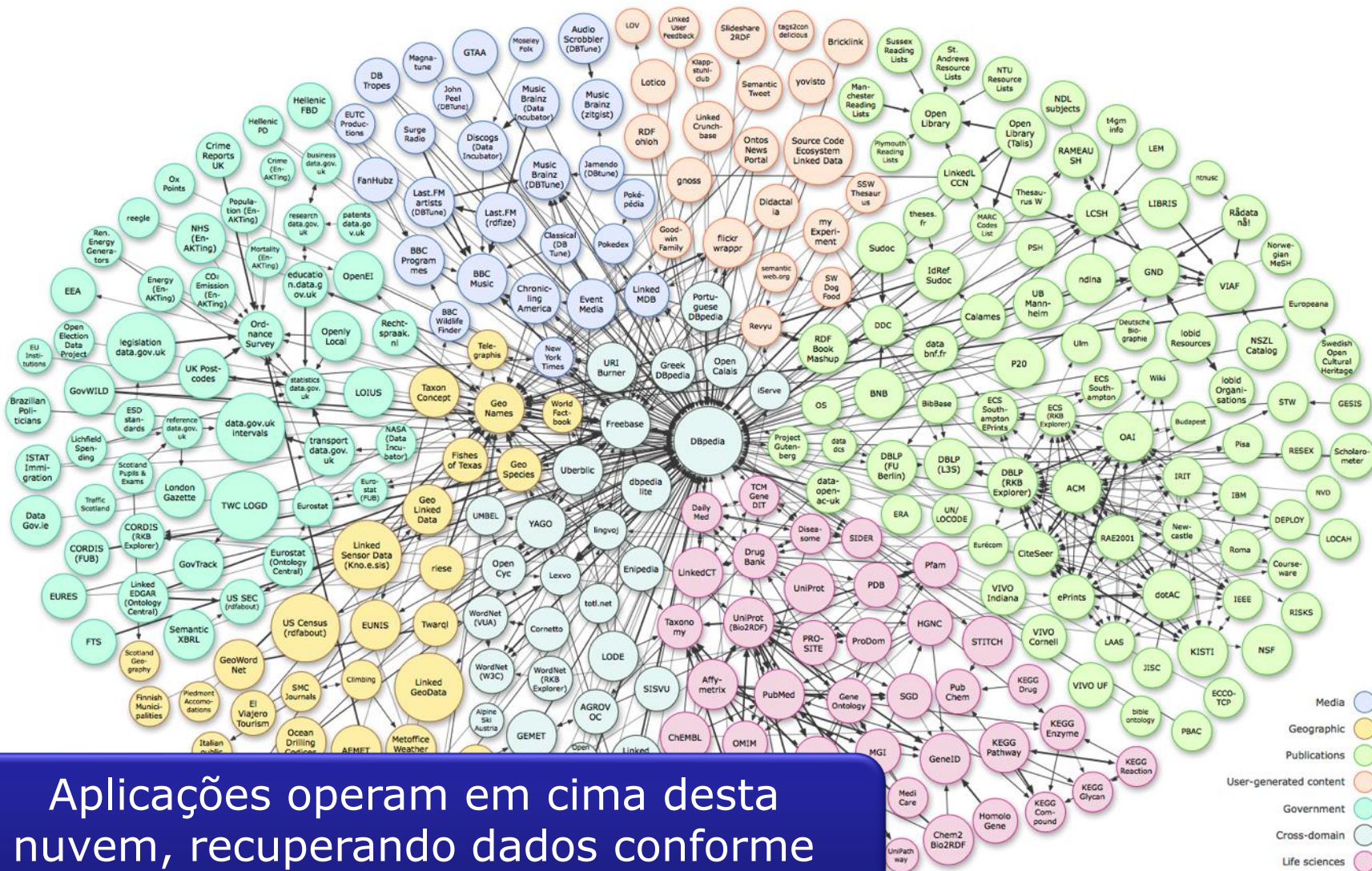
matt-briggs.html


Formam-se redes de dados



- Conjunto de predicados (e suas URIs) prontos pra uso;
- Exemplo: FOAF (<http://xmlns.com/foaf/spec/>):
 - Classes: Agent, Document, Group, Image, LabelProperty, OnlineAccount, OnlineChatAccount, OnlineEcommerceAccount, OnlineGamingAccount, Organization, **Person**, PersonalProfileDocument, Project
 - Propriedades: account, accountName, accountServiceHomepage, age, aimChatID, **based_near**, birthday, currentProject, depiction, depicts, dnaChecksum, familyName, family_name, firstName, focus, fundedBy, geekcode, gender, givenName, givenname, holdsAccount, homepage, icqChatID, img, interest, isPrimaryTopicOf, jabberID, **knows**, lastName, logo, made, maker, mbox, mbox_sha1sum, member, membershipClass, msnChatID, myersBriggs, **name**, nick, openid, page, pastProject, phone, plan, primaryTopic, publications, schoolHomepage, sha1, skypeID, status, surname, theme, thumbnail, tipjar, title, topic, topic_interest, weblog, workInfoHomepage, workplaceHomepage, yahooChatID

E forma-se a nuvem de LOD



As of September 2011 

Aplicações operam em cima desta nuvem, recuperando dados conforme necessário.

Fonte: <http://lod-cloud.net>

Exemplos de vocabulários / conj. dados



 Freebase

2,506,735,867

Facts
(and counting)

43,905,156

Topics
(and counting)



E muitos outros...



Linked Open Vocabularies (LOV)



				<pre><rdf:RDF> <!-- Ontology head - <owl:Ontology rdf:al <swivt:creationDat <owl:imports rdf:r </owl:Ontology></pre>

- Pesquisa: quero a URL de sites (ex.: blogs) de pessoas que Tim Berners-Lee conhece;
- Resultado de pesquisa sintática:
 - Uma página sobre Tim Berners-Lee;
 - Outra página sobre Tim Berners-Lee;
 - Etc.
- Resultado de pesquisa semântica:
 - Um recurso Person p1 (dado que TBL knows p1);
 - Outro recurso p2 (TBL knows p2);
 - Etc.

Mecanismos de busca não ficam para trás!

- Google provê “rich snippets” para pessoas, organizações, produtos, eventos, música, etc.



The screenshot shows a Google search interface. The search bar contains the text "what is the birth date of barack obama". Below the search bar, the "Web" tab is selected. The search results show "About 1,360,000 results (0.59 seconds)". The main result is a rich snippet for Barack Obama's birth date: "August 4, 1961 (age 52 years)". Below this, it says "Barack Obama, Date of birth". To the right of the main snippet is a photo of Barack Obama. Below the main snippet are three smaller snippets for George Washington (February 22, 1732), Vladimir Putin (October 7, 1952), and Abraham Lincoln (February 12, 1809). On the right side of the page, there is a profile card for Barack Obama with 4,340,780 followers on Google+ and a "Follow" button. Below the profile card, there is a detailed description of Barack Obama, including his birth date, full name, nationality, spouse, education, and parents.

Google

what is the birth date of barack obama

Web News Images Shopping Videos More Search tools

About 1,360,000 results (0.59 seconds)

August 4, 1961 (age 52 years)

Barack Obama, Date of birth

George Washington February 22, 1732

Vladimir Putin October 7, 1952

Abraham Lincoln February 12, 1809

Feedback

Barack Obama

4,340,780 followers on Google+

Follow

Barack Hussein Obama II is the 44th and c States, and the first African American to ho

Born: August 4, 1961 (age 52), Honolulu,

Full name: Barack Hussein Obama II

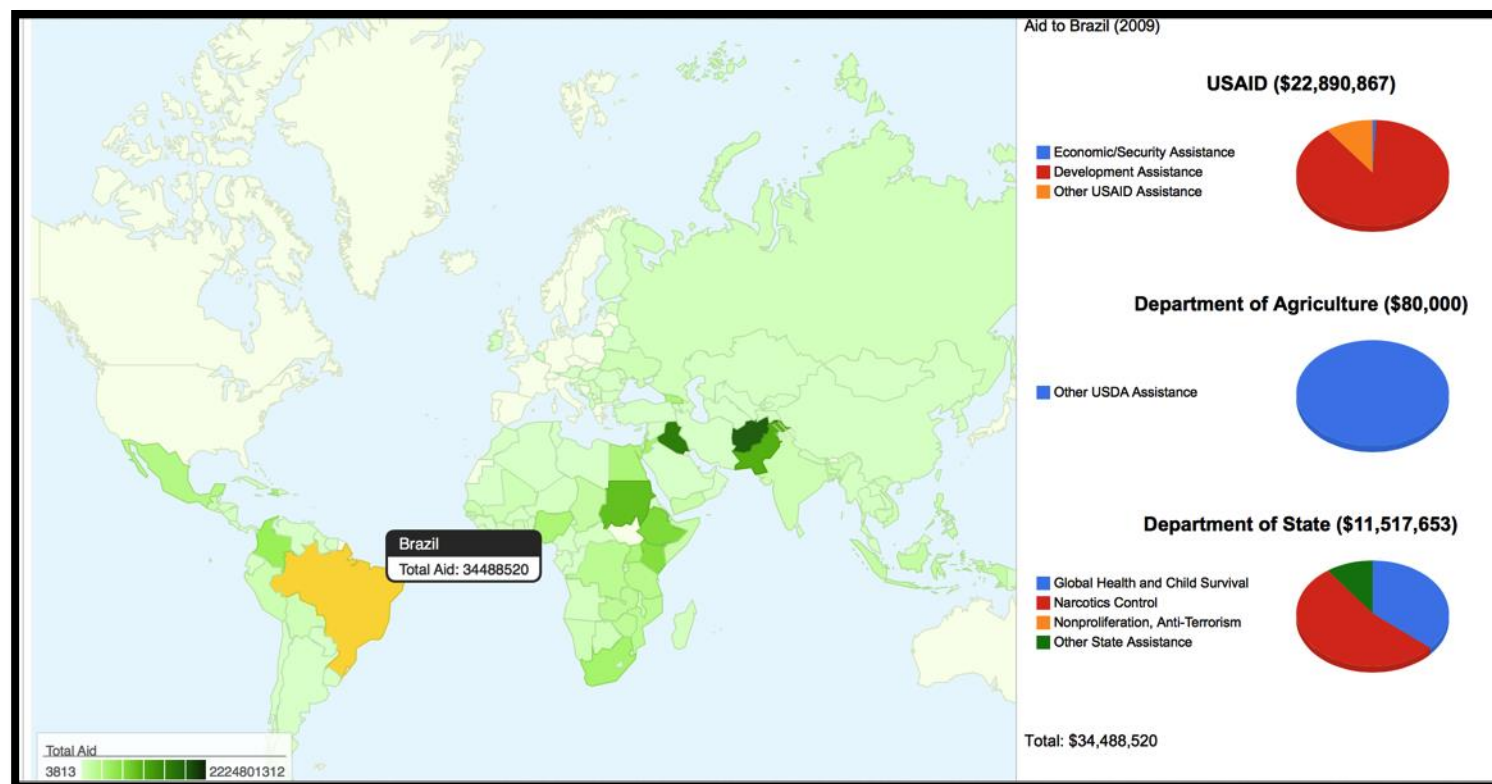
Nationality: American

Spouse: Michelle Obama (m. 1992)

Education: Harvard Law School (1988–19

Parents: Ann Dunham, Barack Obama Sr

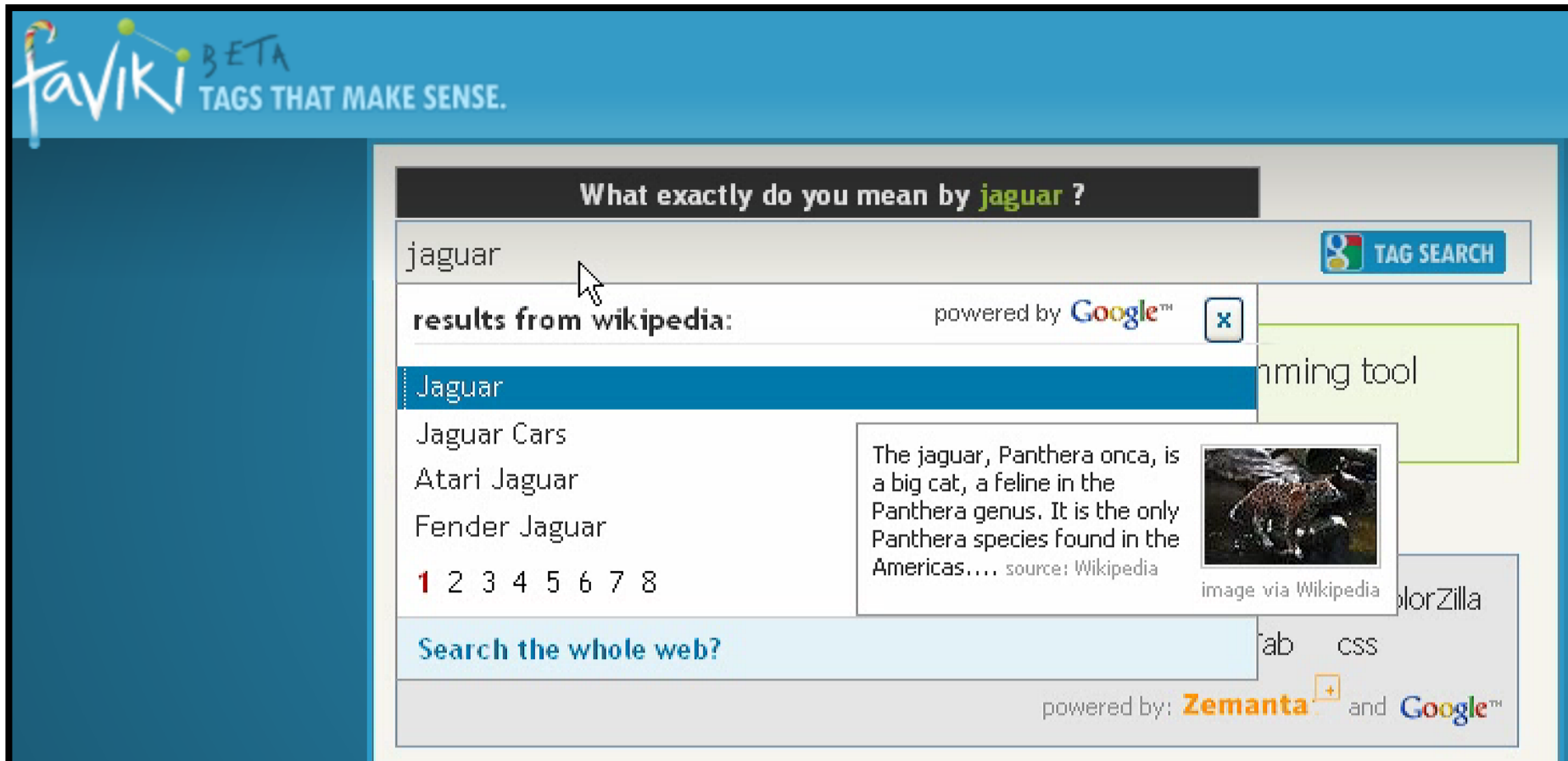
- U.S. Global Foreign Aid Mashup: <http://data-gov.tw.rpi.edu/demo/USForeignAid/demo-1554.html>;
 - Combina gastos no exterior de diferentes departamentos do governo americano.



- DBPedia Mobile (wiki.dbpedia.org/DBpediaMobile):
 - Ajudam turistas a explorar uma cidade;
 - Mostra informações sobre locais próximos;
 - Permite que usuários publiquem check-ins, fotos e avaliações em cima de seus dados.



- Faviki: <http://www.faviki.com>;
 - Bookmarking social;
 - Permite usar conceitos da Wikipedia como tags.



The screenshot shows the Faviki website interface. At the top left, the logo reads "faviki BETA TAGS THAT MAKE SENSE." The main search area has a text input field containing "jaguar" and a "TAG SEARCH" button. Below the input field, it says "results from wikipedia: powered by Google™". A list of search results is displayed, with "Jaguar" selected. The results include "Jaguar Cars", "Atari Jaguar", and "Fender Jaguar". A preview box for the selected "Jaguar" result shows a description: "The jaguar, Panthera onca, is a big cat, a feline in the Panthera genus. It is the only Panthera species found in the Americas.... source: Wikipedia" and an image of a jaguar. At the bottom, it says "Search the whole web?" and "powered by: Zemanta and Google™".

- Curso “Desenvolvimento Web e Web Semântica” em 2014/1 – será repetido em 2015/1, 2016/1, ...;
- Alunos produziram aplicações Web com algo da LD:
 - Código-fonte: <https://github.com/dwws-ufes/>
 - Aplicações: <http://dev.nemo.inf.ufes.br:8180/dwws/>



<http://nemo.inf.ufes.br/>