

Banco de Dados

Fundamentos

André Santanchè e Luiz Celso Gomes Jr
Instituto de Computação - UNICAMP
Agosto 2014

Motivação

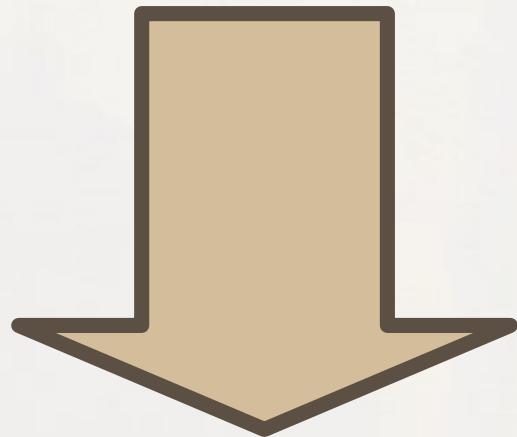
- Aplicações computacionais de todos os portes trabalham com grandes volumes de dados
 - Gerenciamento de uma farmácia
 - Sistema bibliotecário
 - Sistema bancário

Motivação

- Grandes volumes de dados e suas relações complexas justificam a criação de estratégias específicas para gerenciá-los

Motivação

- Grandes volumes de dados e suas relações complexas justificam a criação de estratégias específicas para gerenciá-los



Bancos de Dados

Aplicações Tradicionais

- Bancos de dados numéricos e tradicionais
- Exemplos:
 - Gerenciamento de uma farmácia
 - Sistema bibliotecário
 - Sistema bancário

Aplicações

- Aplicações mais recentes

- Bancos de Dados Multimídia
- Sistemas de Informação Geográfica (GIS)
- Data Warehouses
- etc.

Banco de Dados

Aplicação Exemplo

- Gerenciamento de uma biblioteca
- Serviços:
 - Cadastro de membros associados
 - Registro do acervo (ex.: livros, revistas etc.)
 - Controle de empréstimos

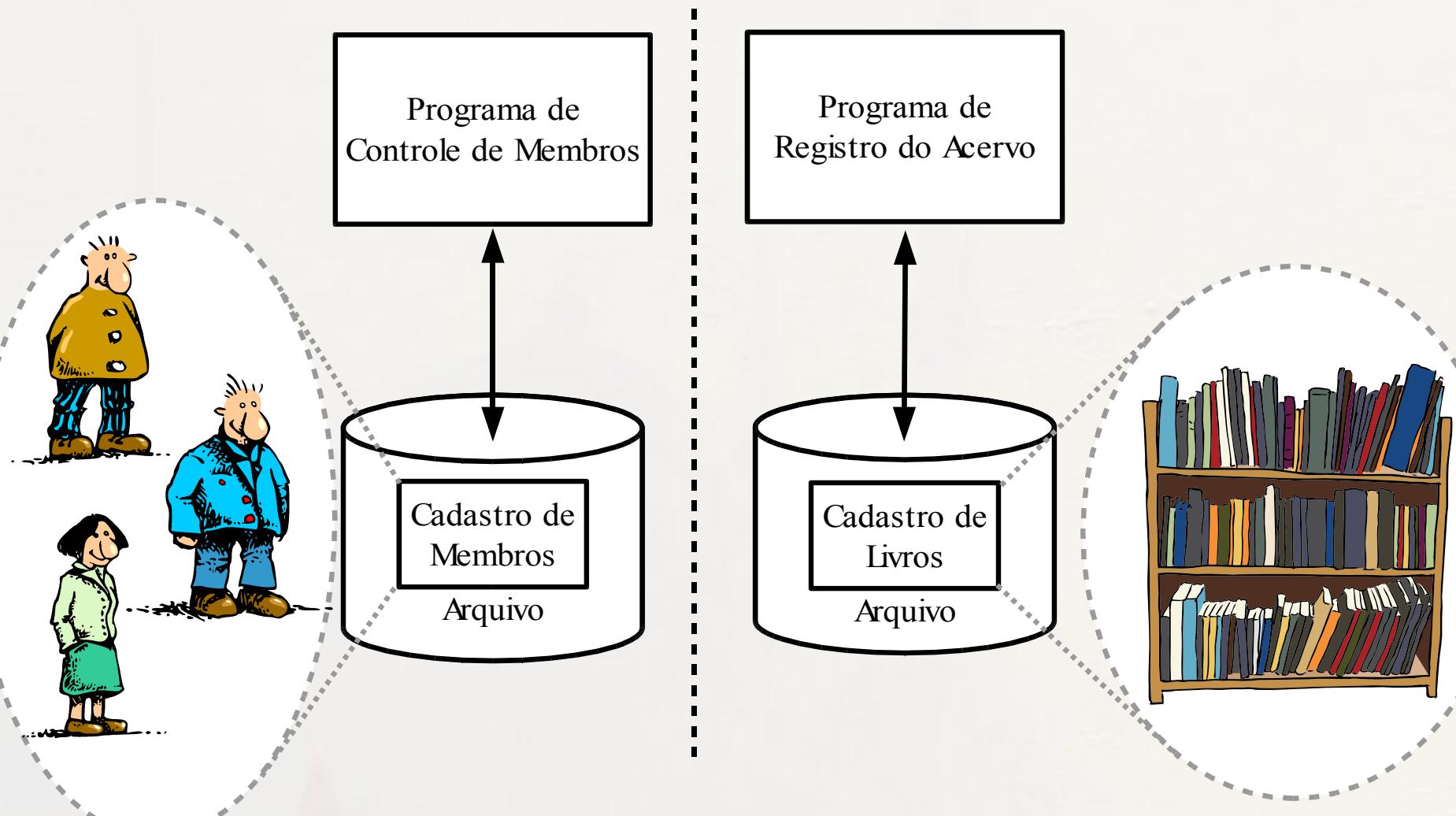
Banco de Dados

Perspectivas

- Arquitetura
- Abstração

Banco de Dados Arquitetura

Sistemas Isolados

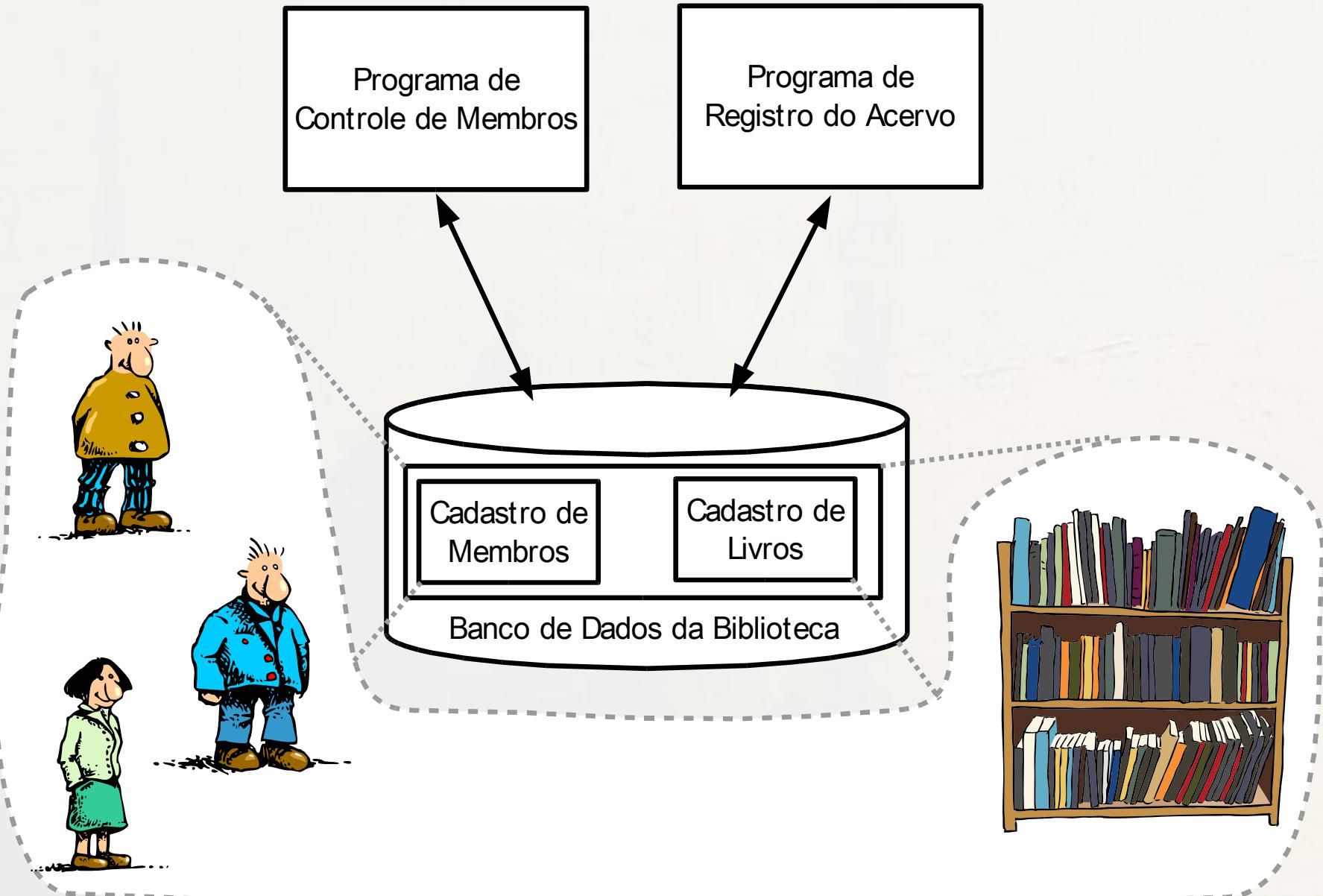


Sistemas/Arquivos Isolados

- Redundância não controlada
 - Repetição
 - Inconsistência
- Barreiras para relacionamento entre arquivos
- Dificuldades com:
 - acesso concorrente
 - integridade e recuperação em caso de crash
 - segurança e controle de acesso

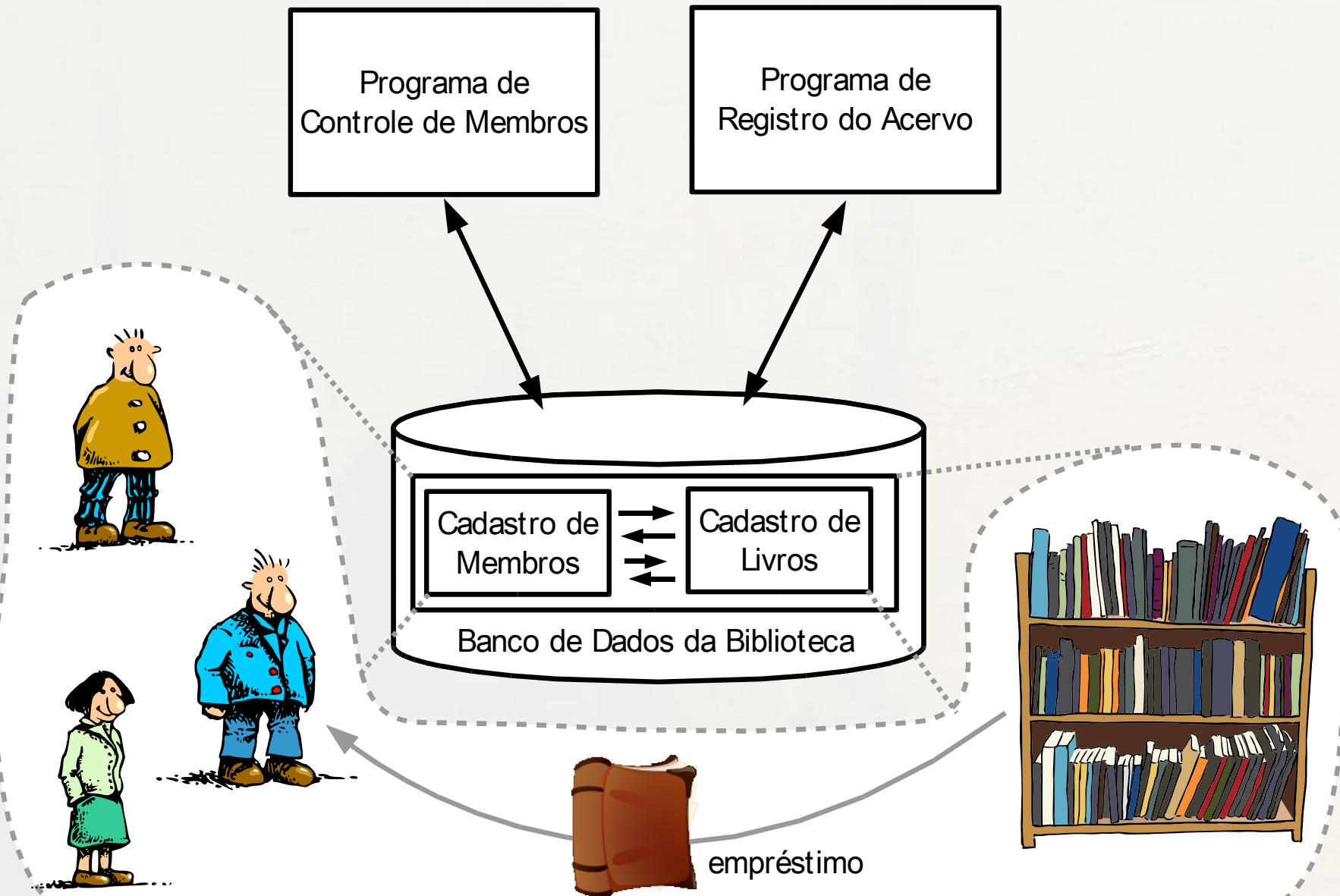
Banco de Dados

Compartilhamento de Dados



Banco de Dados

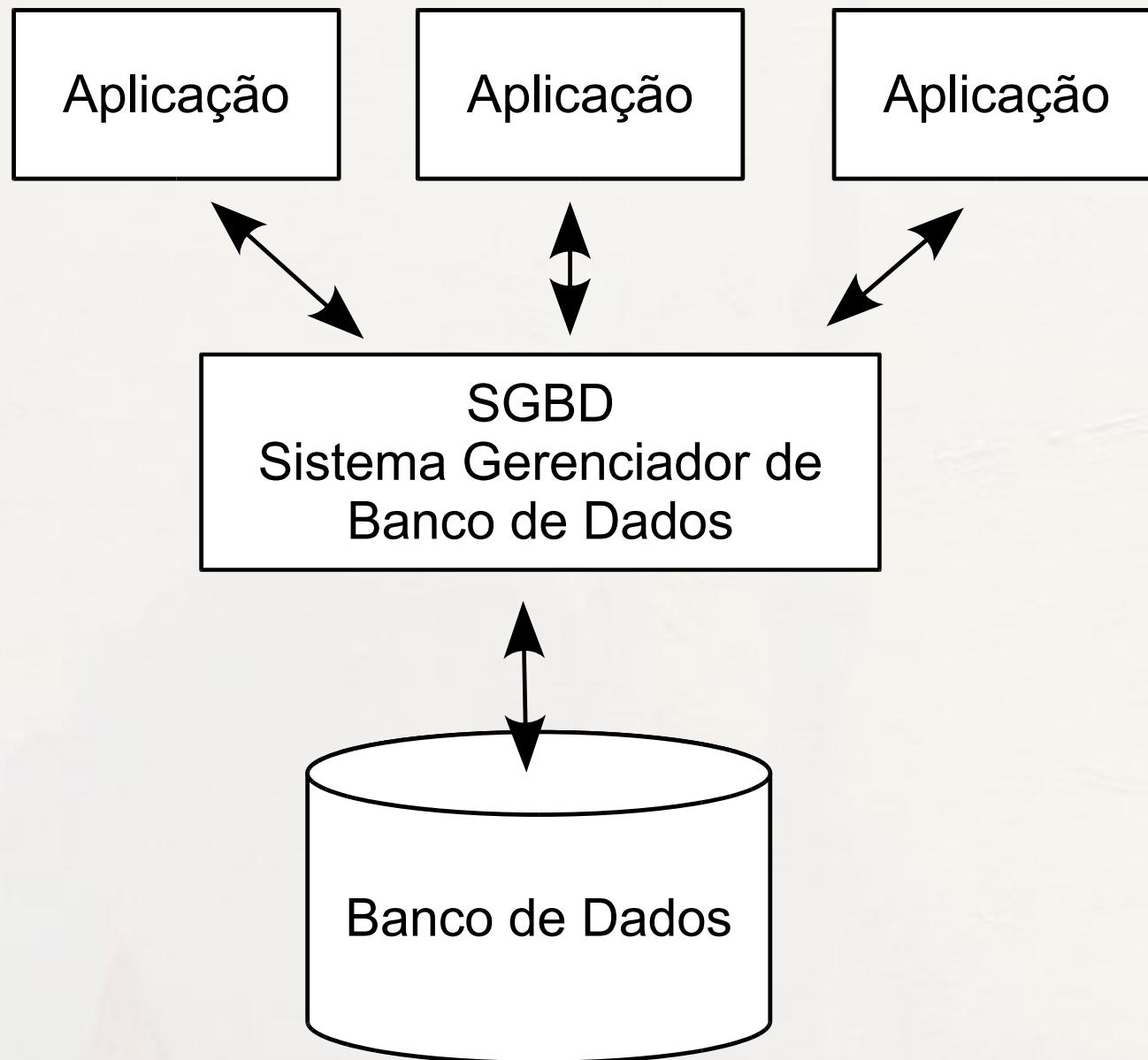
Relações entre os Dados



Sistema Gerenciador de Banco de Dados (SGBD)

- Sistema de software com finalidade genérica
- Projetado para a definição, construção e manipulação de bancos de dados
- Pode atender várias aplicações

SGBD



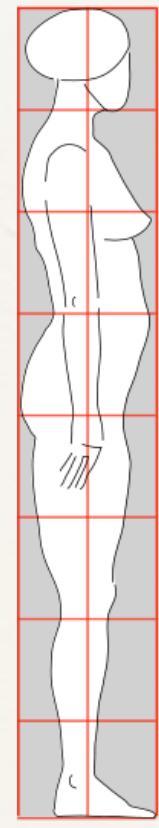
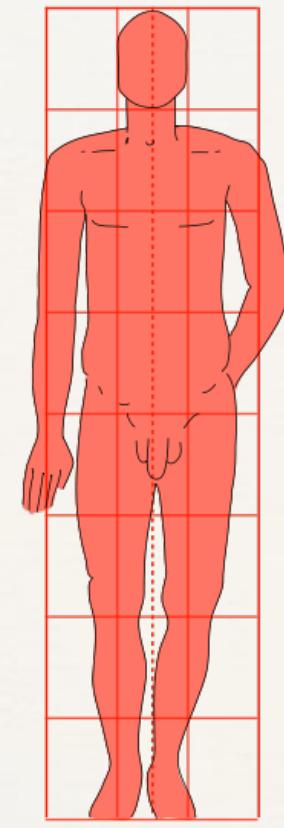
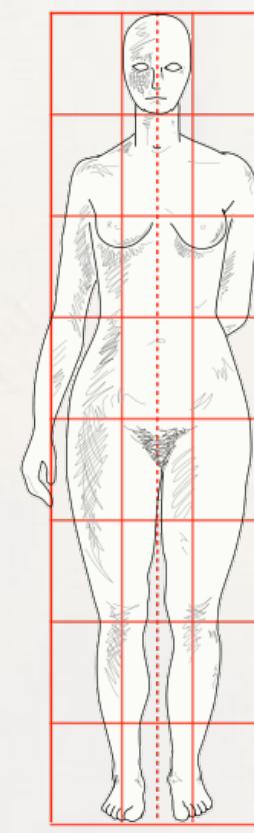
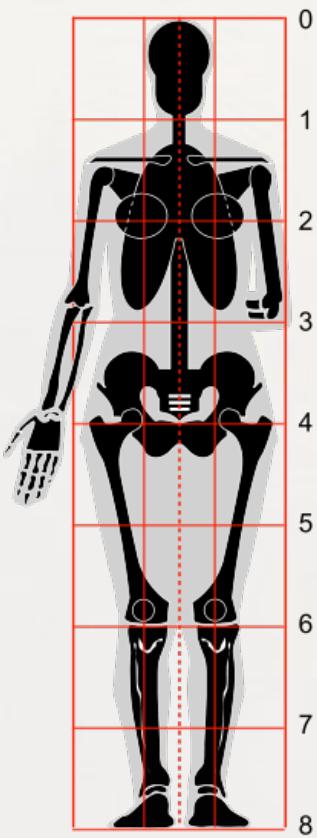
Vantagens de um SGBD

- Independência de dados
- Acesso eficiente
- Tempo reduzido no desenvolvimento de aplicações
- Segurança e integridade de dados
- Administração de dados uniforme
- Acesso concorrente
- Recuperação contra *crashes*

(Ramakrishnan, 2003b)

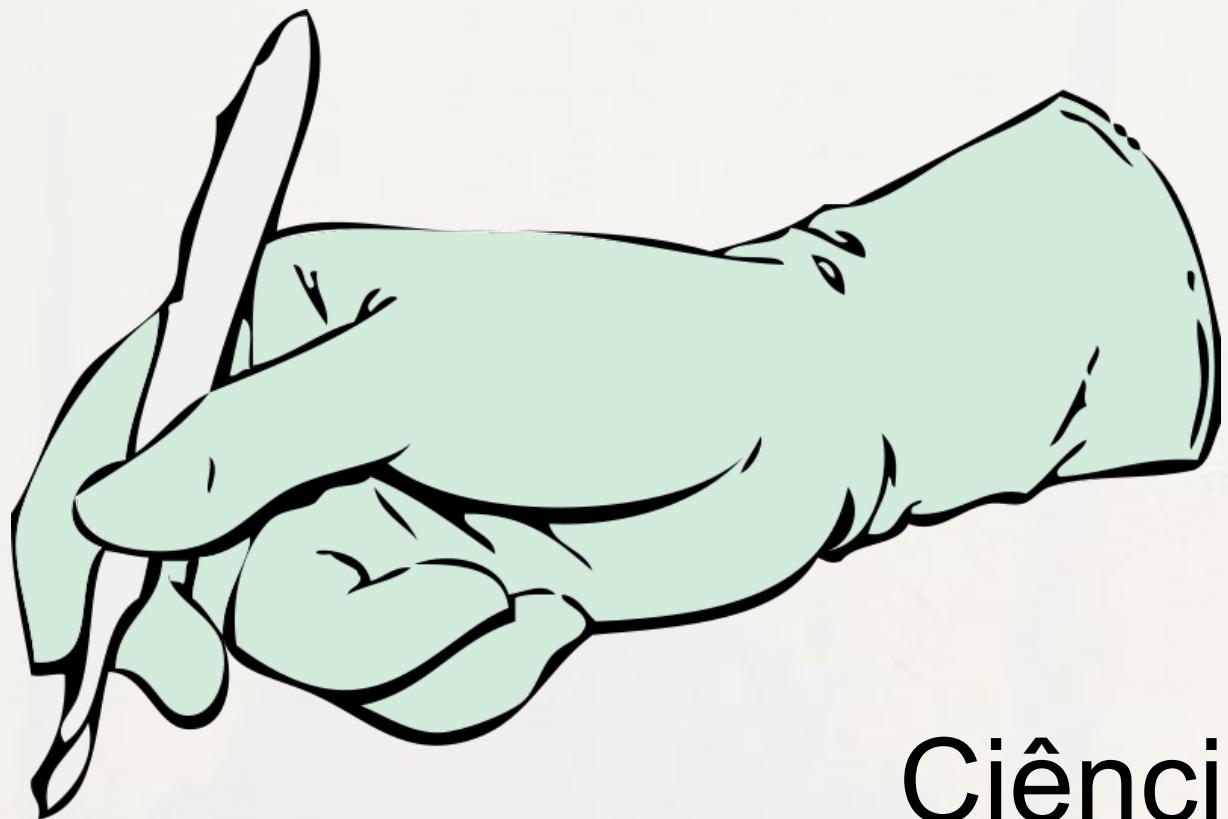
Como Aprendemos Computação?

Cirurgia



By Original by Schnorch retracted by [LadyofHats](#)

Cirurgia



Ciência da Faca

Ciência da Computação?

- Computer Science like Knife Science

(Dijkstra, 1986)

Banco de Dados

Data Science

Quarto Paradigma



The Fourth Paradigm: Data-Intensive Scientific Discovery
Editado por Tony Hey, Stewart Tansley, and Kristin Tolle
Microsoft Research
Redmond, 2009

Fourth Paradigm

- Thousand years ago:
 - science was empirical; describing natural phenomena
- Last few hundred years:
 - theoretical branch; using models, generalizations
- Last few decades:
 - a computational branch; simulating complex phenomena
- Today: data exploration (eScience)
 - unify theory, experiment, and simulation

(Jim Gray, 2007)

Trends

Web Search interest: h1n1. Worldwide, 2004 - present.



Hot Searches

Top Charts **New!****Explore**

Search terms

h1n1[+ Add term](#)

Limit to

Web Search

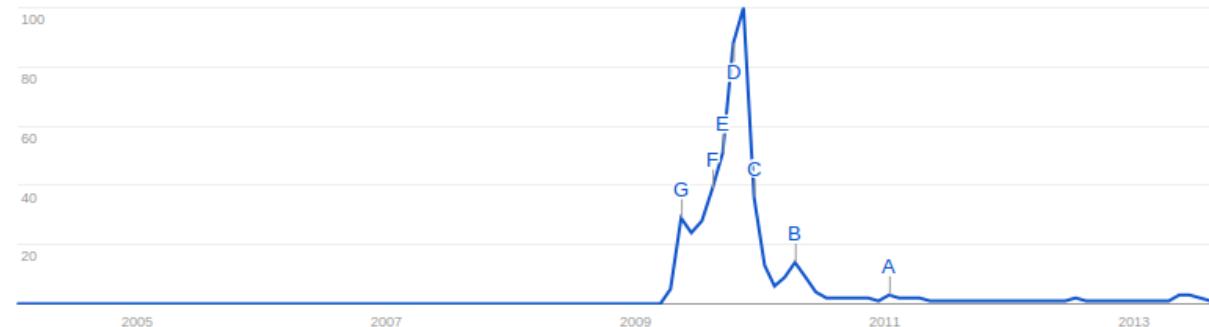
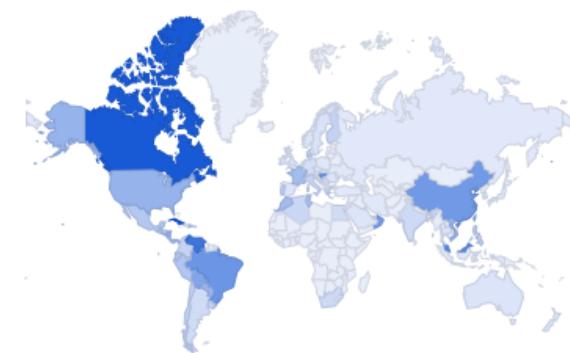
Worldwide

2004 - present

All categories

Interest over time

The number 100 represents the peak search interest

 News headlines Forecast [Embed](#)**Regional interest** 

0 100

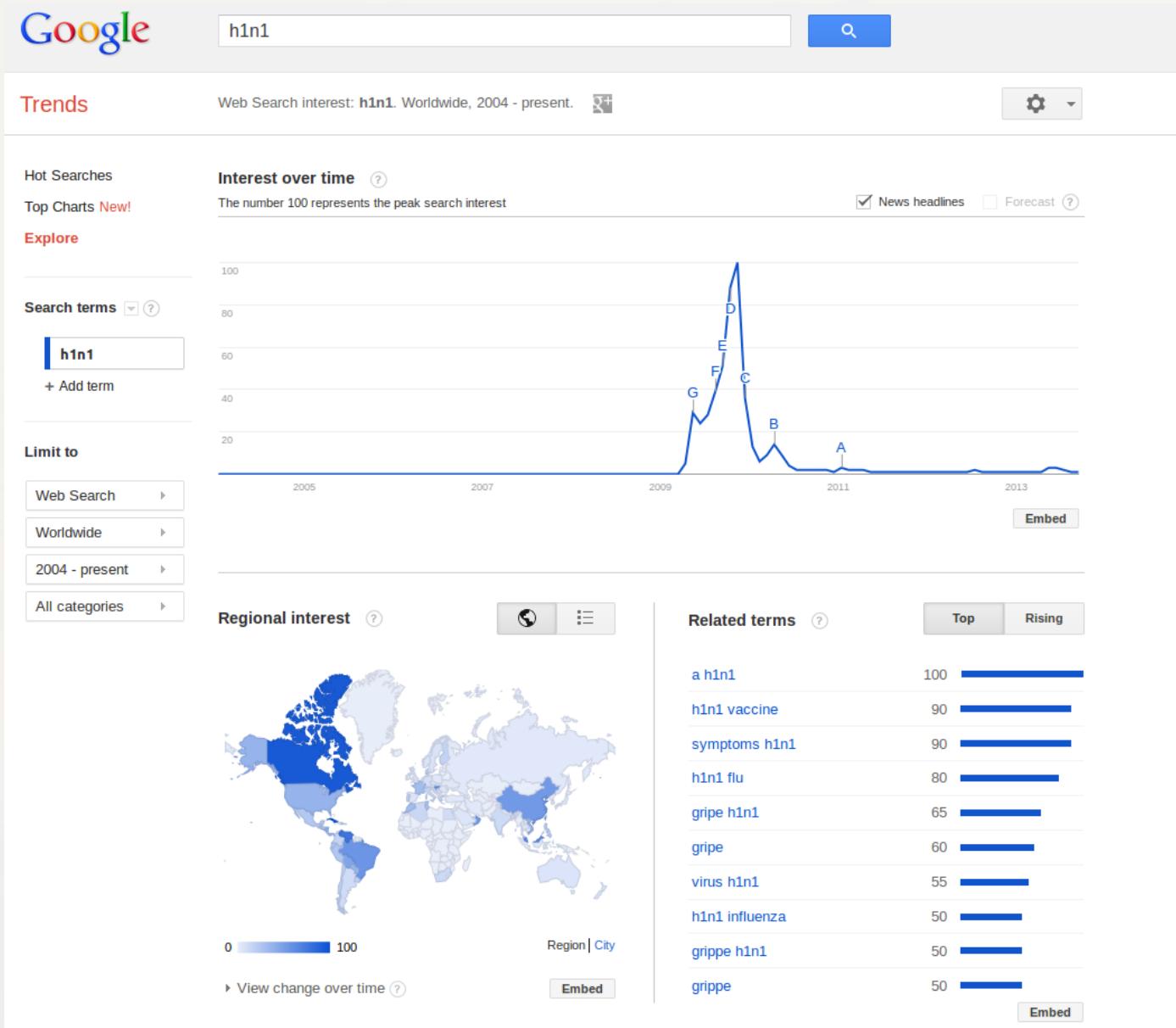
Region | City

[View change over time](#) [Embed](#)**Related terms** [Top](#)[Rising](#)

a h1n1	100	
h1n1 vaccine	90	
symptoms h1n1	90	
h1n1 flu	80	
gripe h1n1	65	
gripe	60	
virus h1n1	55	
h1n1 influenza	50	
grippe h1n1	50	
grippe	50	

[Embed](#)

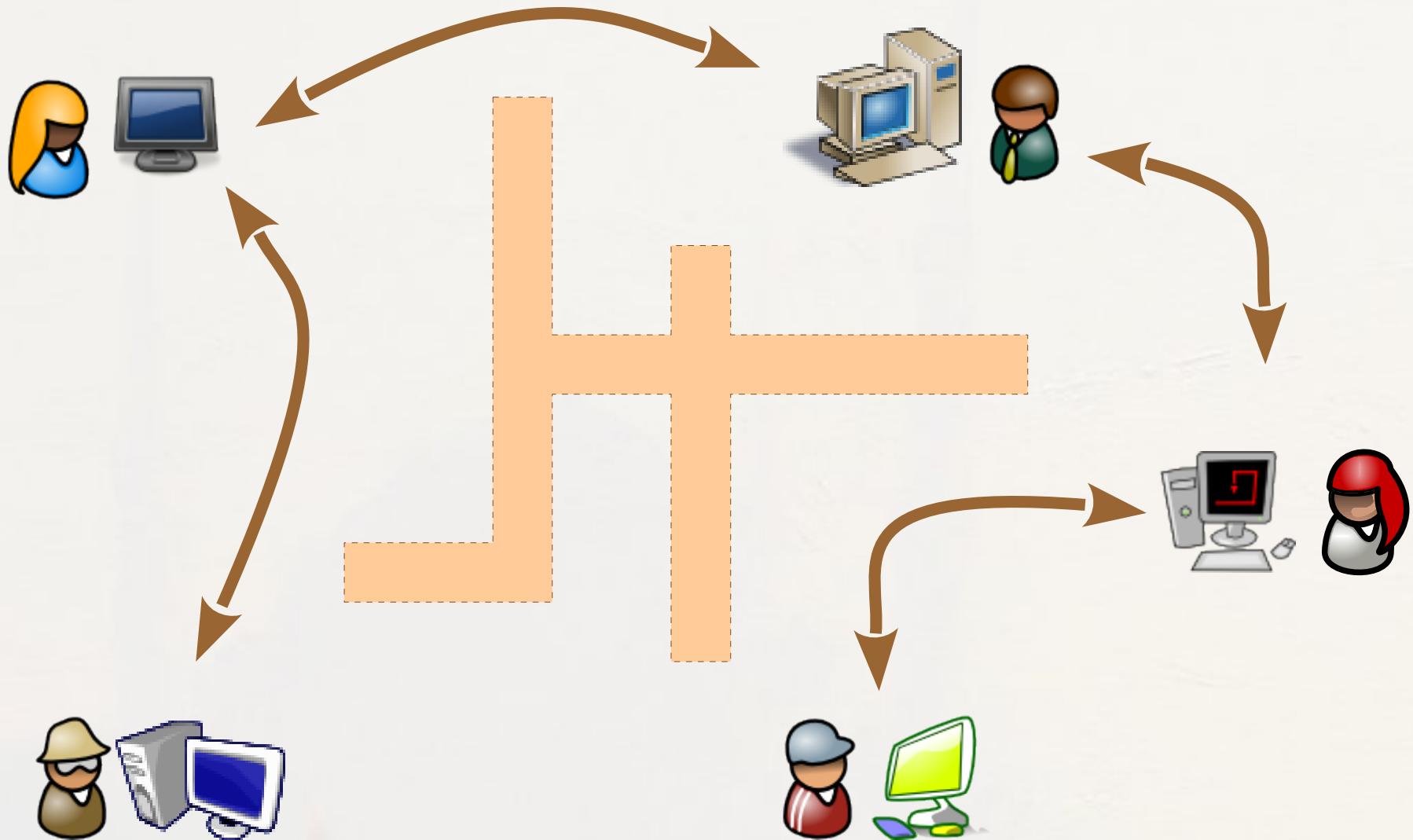
Google Trends



Mundo Físico



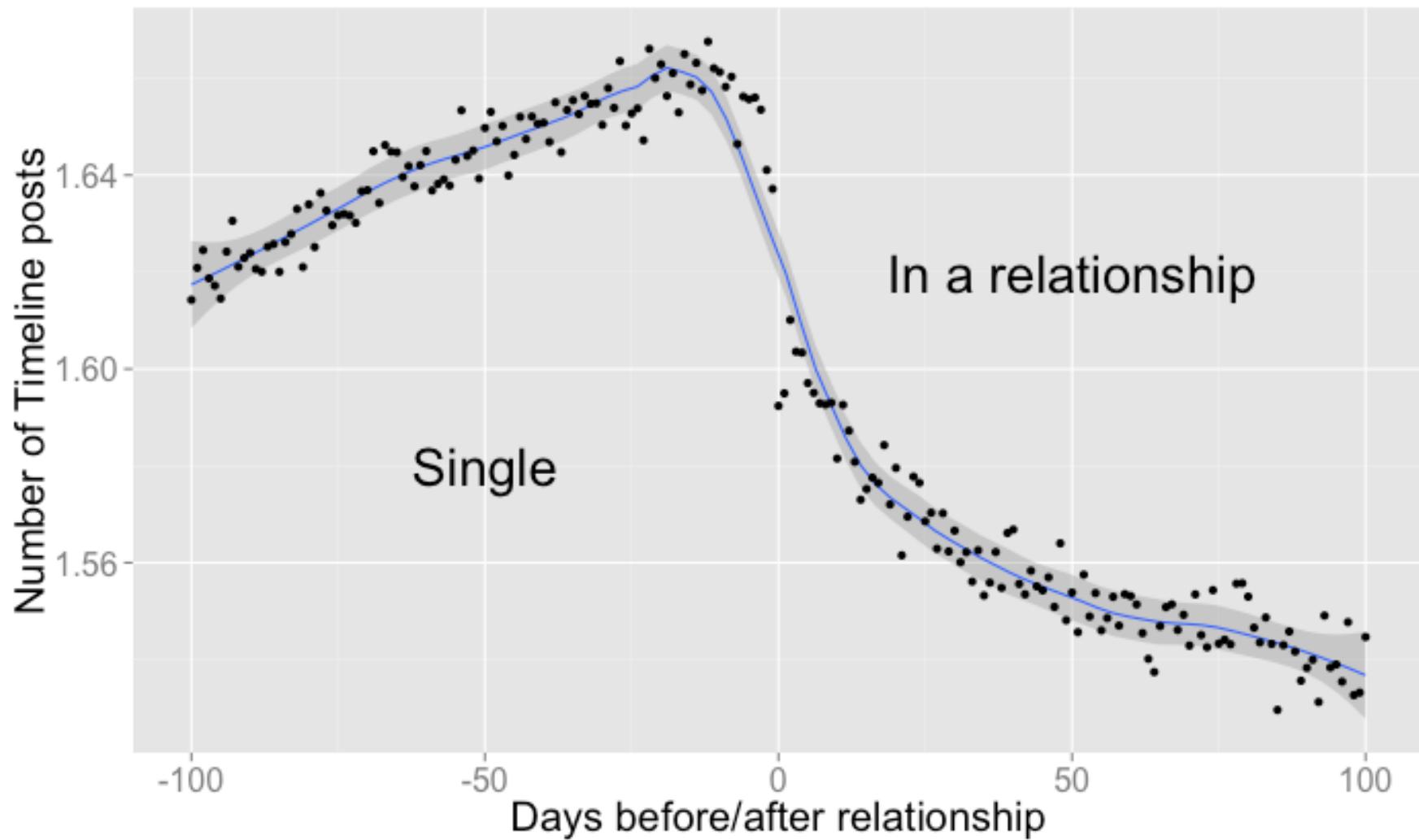
Mundo Virtual



Facebook

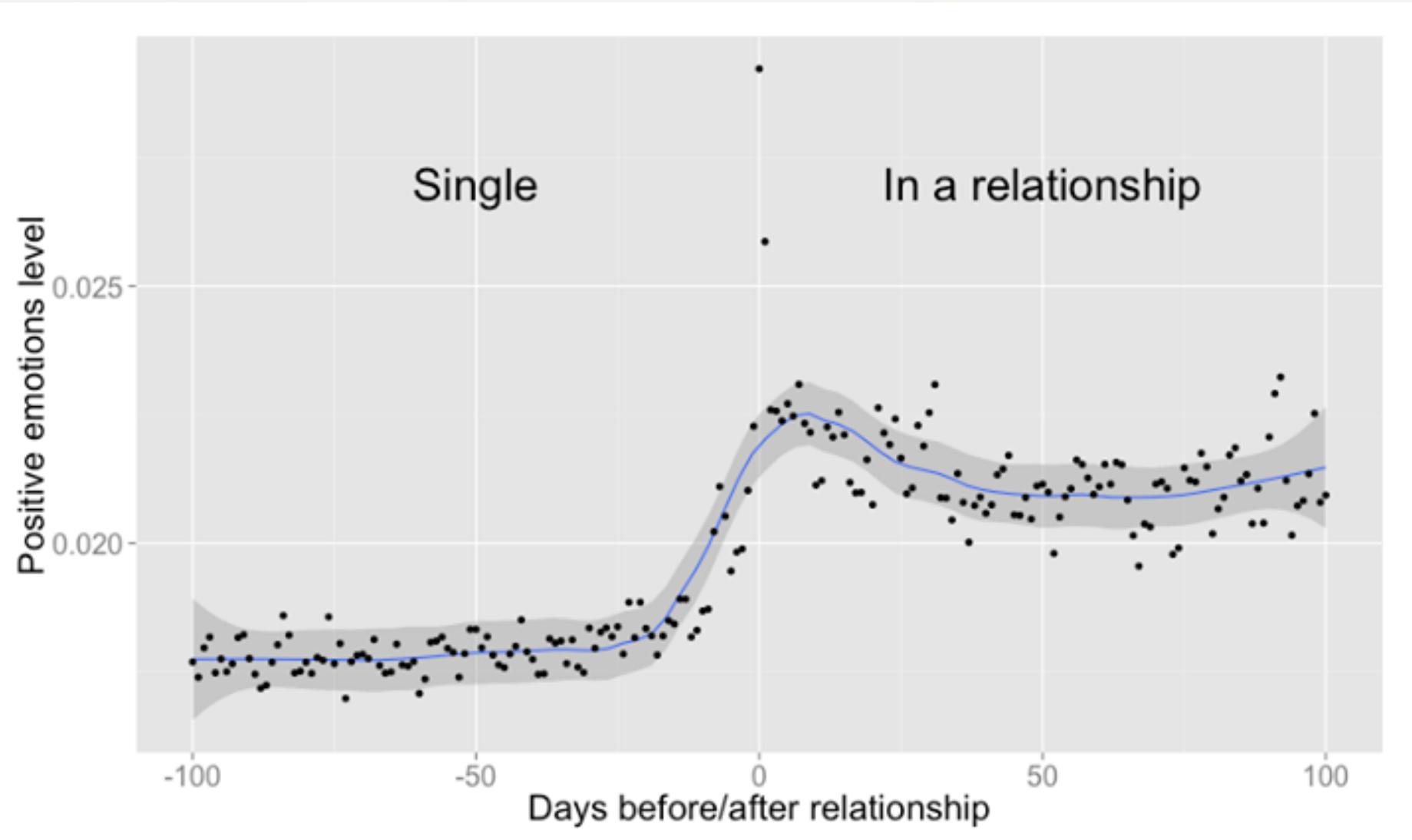
May 2013 - 1.11 billion people

Facebook



(Diuk, 2014)

Facebook



(Diuk, 2014)

The Formation of Love

The Formation of Love

By Carlos Greg Diuk on Friday, February 14, 2014
at 3:59pm

by Carlos Diuk, Facebook Data Science

<https://www.facebook.com/notes/facebook-data-science/the-formation-of-love/10152064609253859>

Genoma Humano

3.3 bilhões base-pairs

Genoma Humano

3.3 bilhões base-pairs

BioInformática

What Wal-Mart Knows About Customers' Habits

Constance L. Hays
The New York Times, 2004

What Wal-Mart Knows About Customers' Habits

- "start predicting what's going to happen, instead of waiting for it to happen"
- "We didn't know in the past that strawberry Pop-Tarts increase in sales, like seven times their normal sales rate, ahead of a hurricane"
- "And the pre-hurricane top-selling item was beer"

Linda M. Dillman - Wal Mart

Why Graphs? Linked Data

Wikipedia

Firefox ▾ W Paris - Wikipedia, the free encyclopedia + en.wikipedia.org/wiki/Paris

Article Talk Read View source View history Search

Paris

From Wikipedia, the free encyclopedia

This article is about the capital of France. For other uses, see [Paris \(disambiguation\)](#).

Paris (English /pærɪs/, /pɛrɪs/; French: [paʁi] (listen)) is the capital and most populous city of France. It is situated on the River Seine, in the north of the country, at the heart of the Île-de-France region. Within its administrative limits (the 20 arrondissements), the city had 2,234,105 inhabitants in 2009 while its metropolitan area is one of the largest population centres in Europe with more than 12 million inhabitants.

Country France
Region Île-de-France
Department Paris
Subdivisions 20 arrondissements
Government

- **Mayor (2008–14)** Bertrand Delanoë (PS)

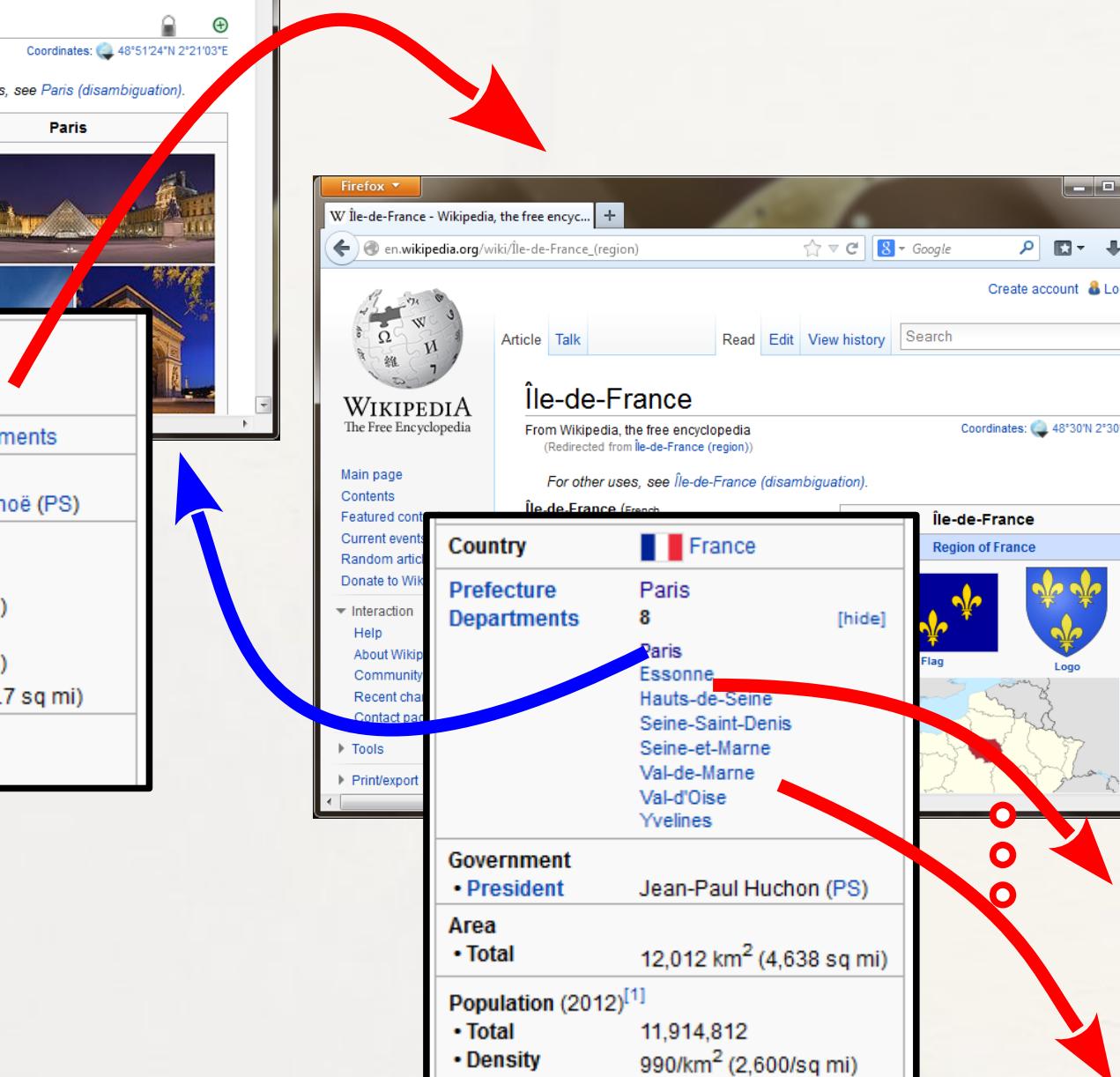
Area^[1]

- **Urban (2010)** 2,844.8 km² (1,098.4 sq mi)
- **Metro (2010)** 17,174.4 km² (6,631.1 sq mi)
- **Land¹** 105.4 km² (40.7 sq mi)

Population (2010)^[5]

- **Rank** 1st in France

Coordinates: 48°51'24"N 2°21'03"E



Firefox ▾ W Île-de-France - Wikipedia, the free encyclopedia + en.wikipedia.org/wiki/Île-de-France_(region)

Article Talk Read Edit View history Search

Île-de-France

From Wikipedia, the free encyclopedia (Redirected from [Île-de-France \(region\)](#))

For other uses, see [Île-de-France \(disambiguation\)](#).

Île-de-France (French)

Country France
Prefecture Paris
Departments 8

Regions

- Paris
- Essonne
- Hauts-de-Seine
- Seine-Saint-Denis
- Seine-et-Marne
- Val-de-Marne
- Val-d'Oise
- Yvelines

Government

- **President** Jean-Paul Huchon (PS)

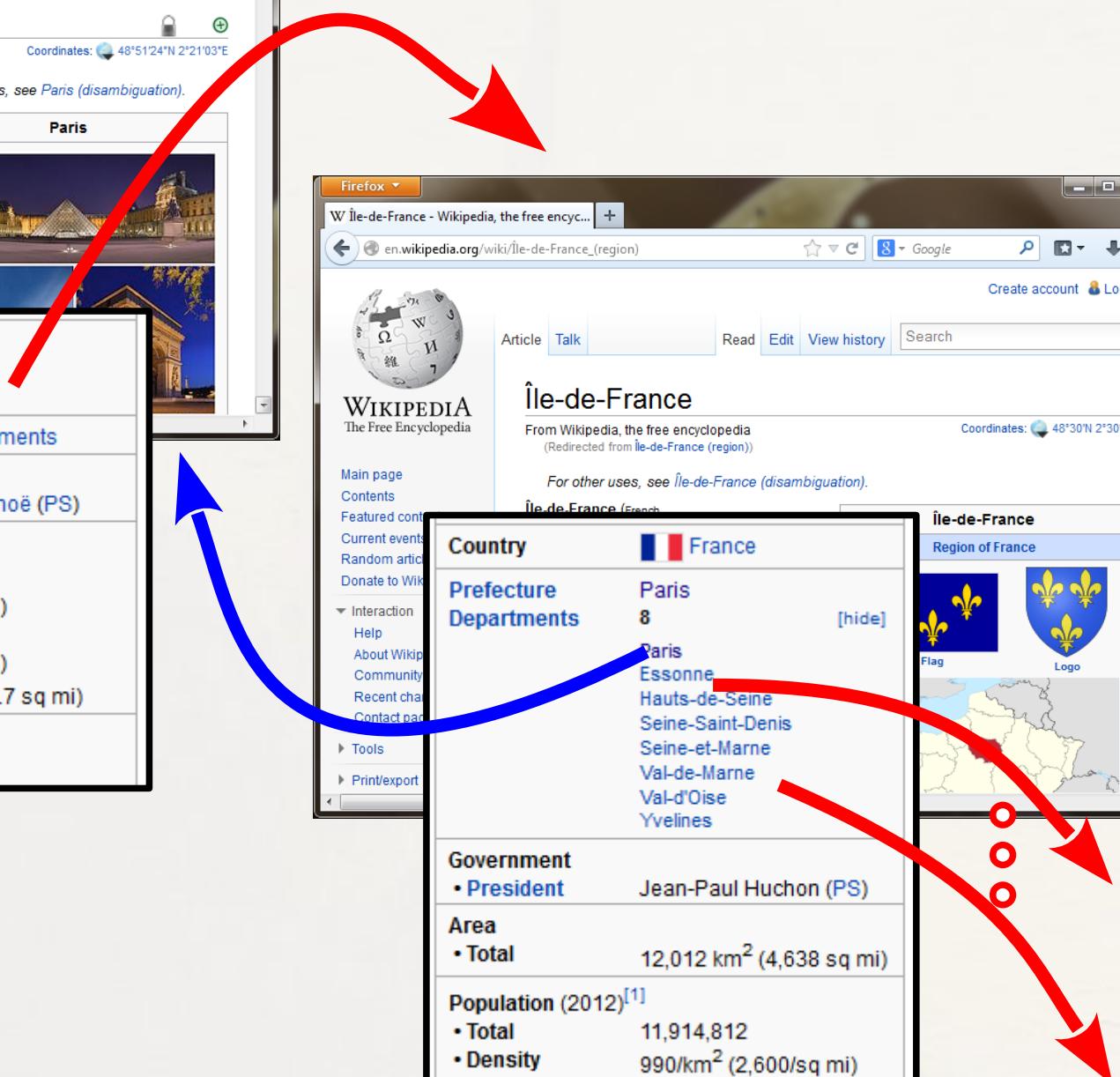
Area

- **Total** 12,012 km² (4,638 sq mi)

Population (2012)^[1]

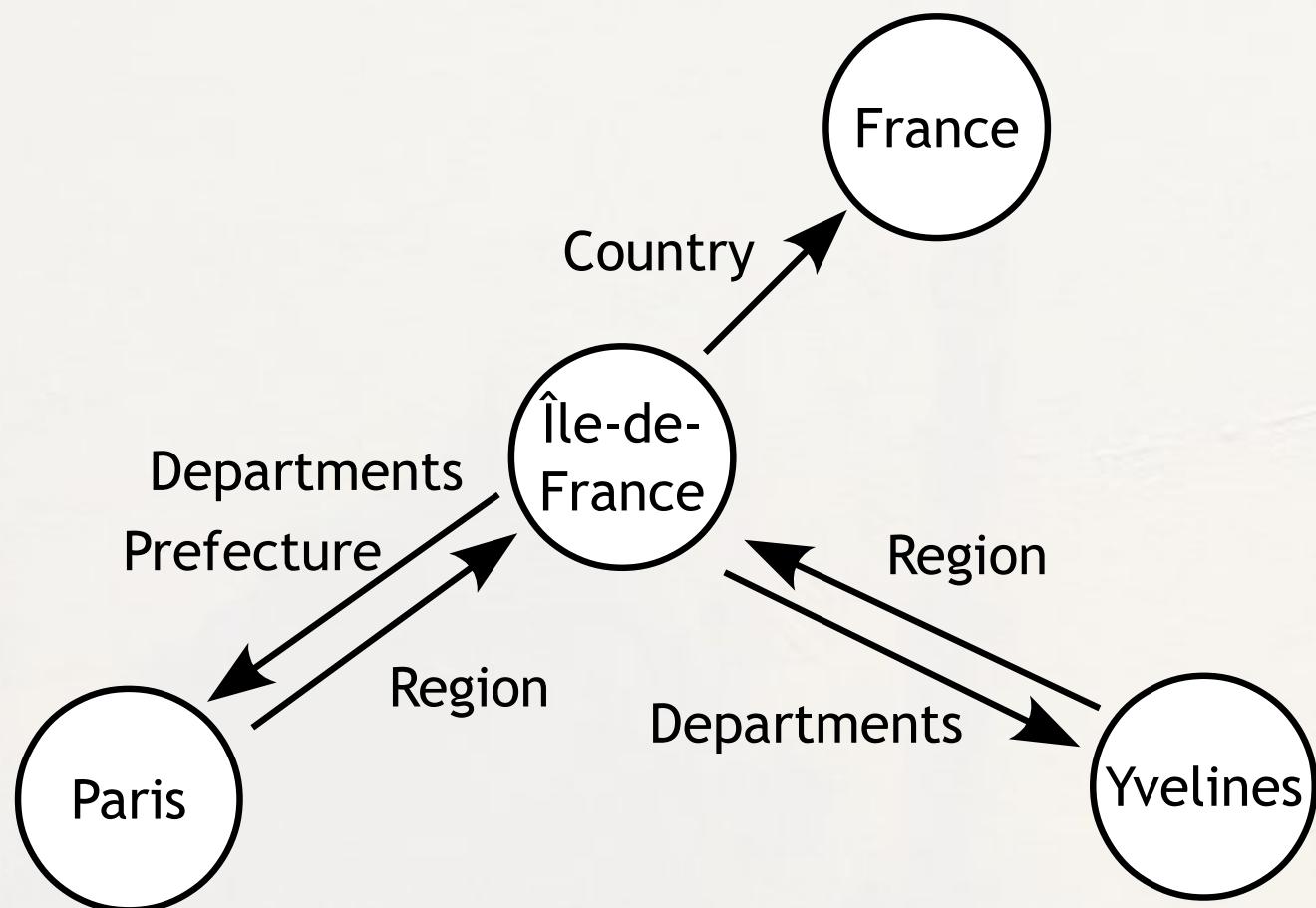
- **Total** 11,914,812
- **Density** 990/km² (2,600/sq mi)

Coordinates: 48°30'N 2°30'E

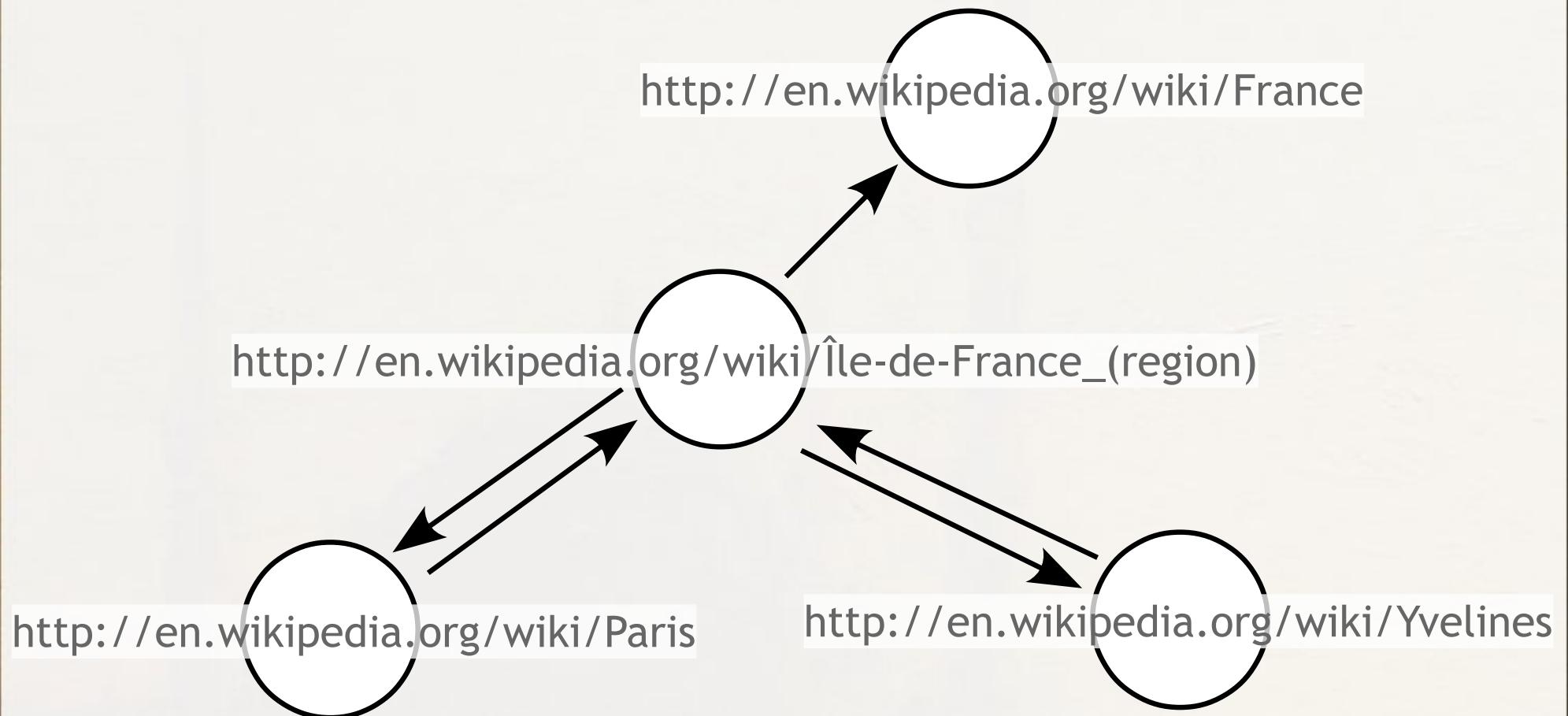


Infobox

DBpedia



DBpedia (URIs)



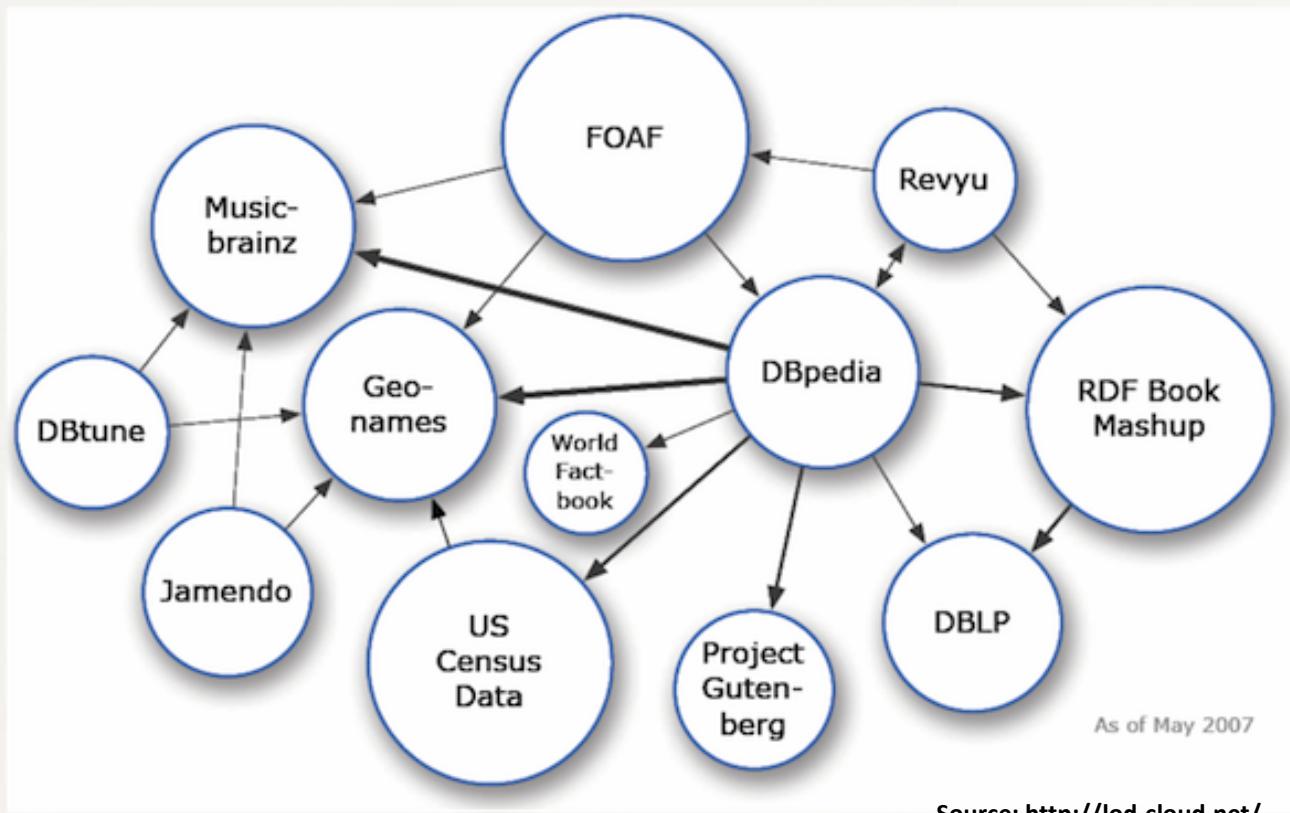
DBpedia - English

- 4 million things
- 3.22 million classified in a consistent ontology
 - 832,000 persons
 - 639,000 places (427,000 populated)
 - 372,000 creative works
 - 116,000 music albums; 78,000 films; 18,500 video games
 - 209,000 organizations
 - 226,000 species
 - 5,600 diseases.

DBpedia - International

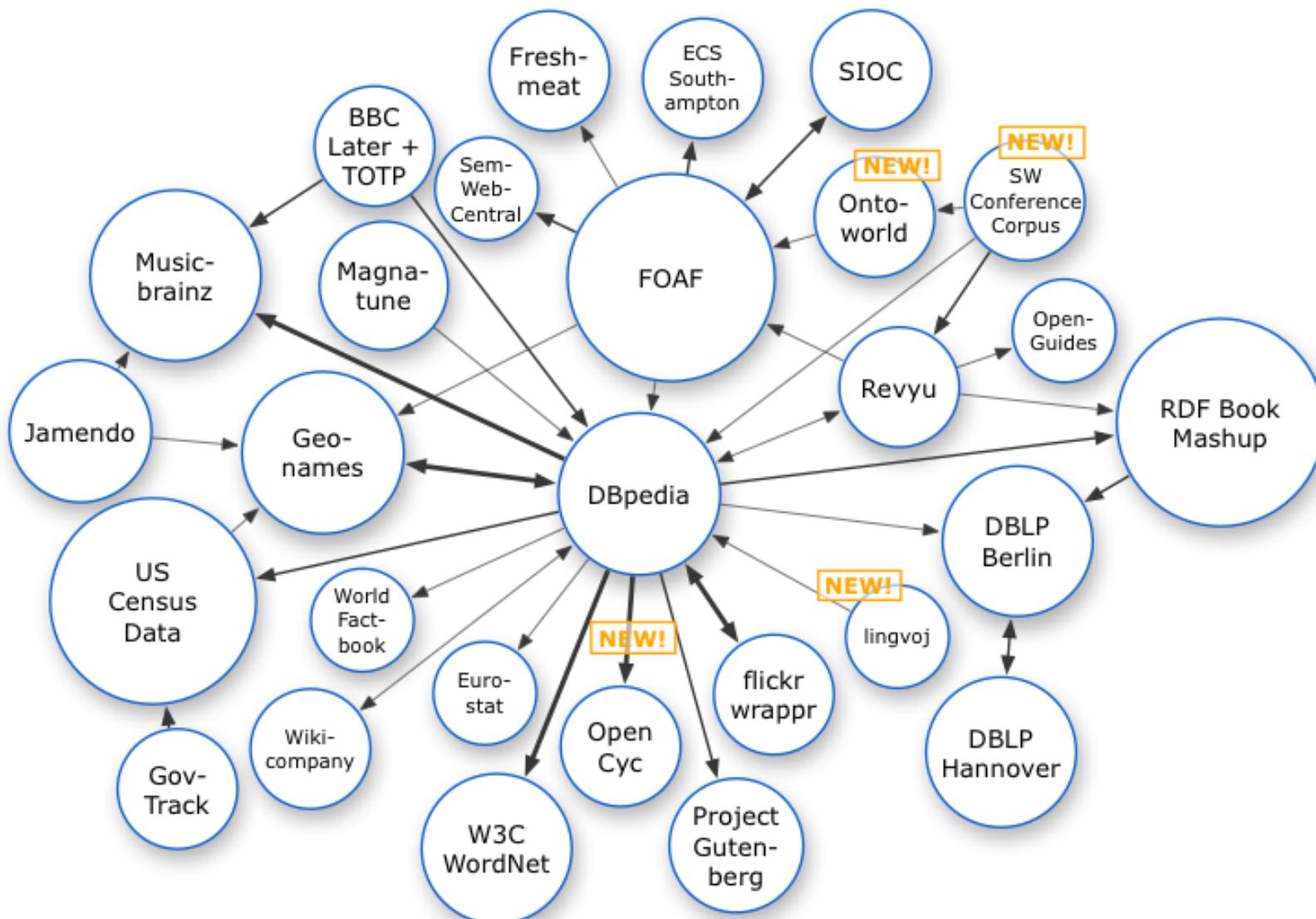
- 119 languages
- 24.9 million things
- 16.8 million interlinked with English
- 12.6 million unique things

Linked Data



Datasets published following Linked Data ‘format’: **05/2007**

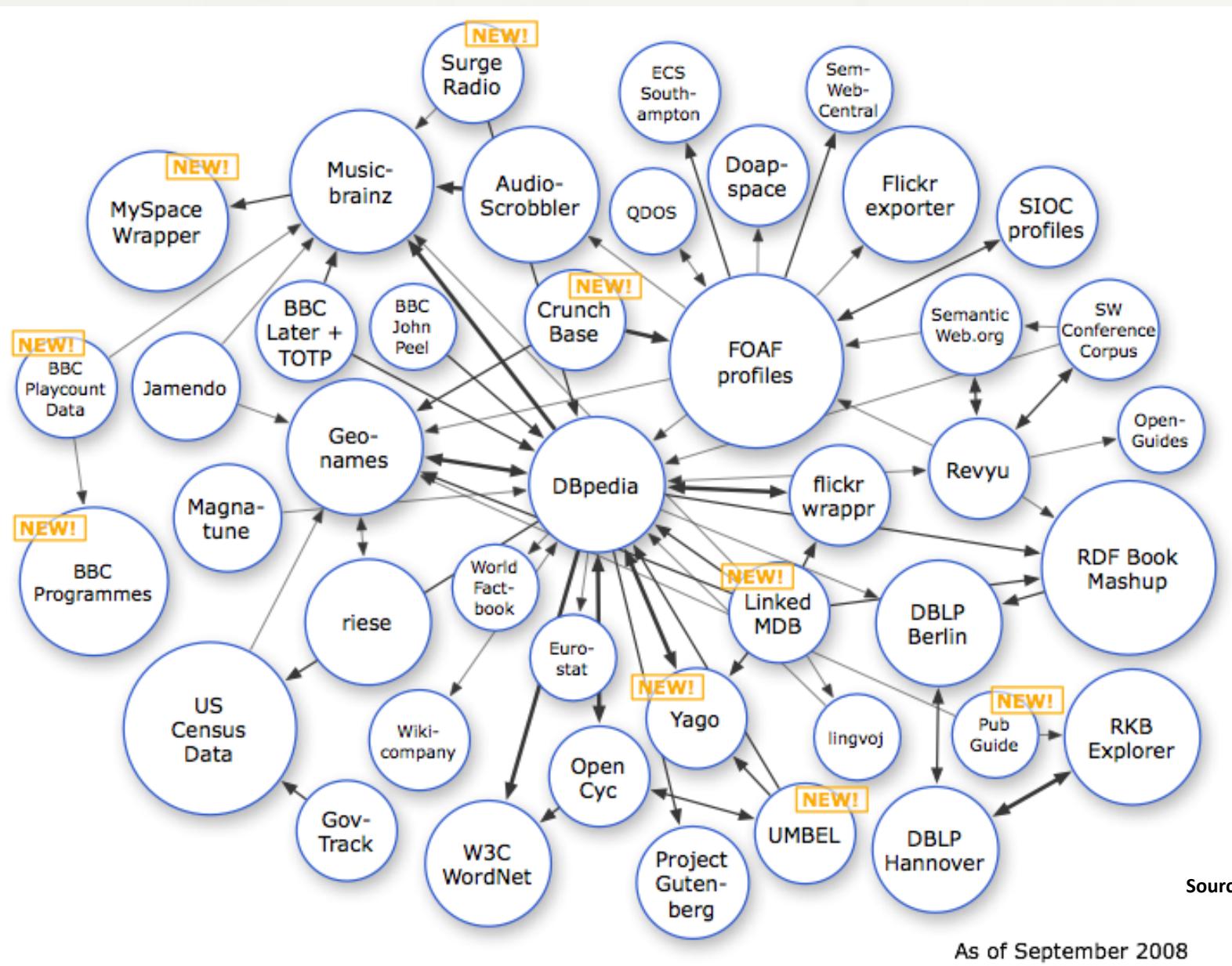
Linked Data



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

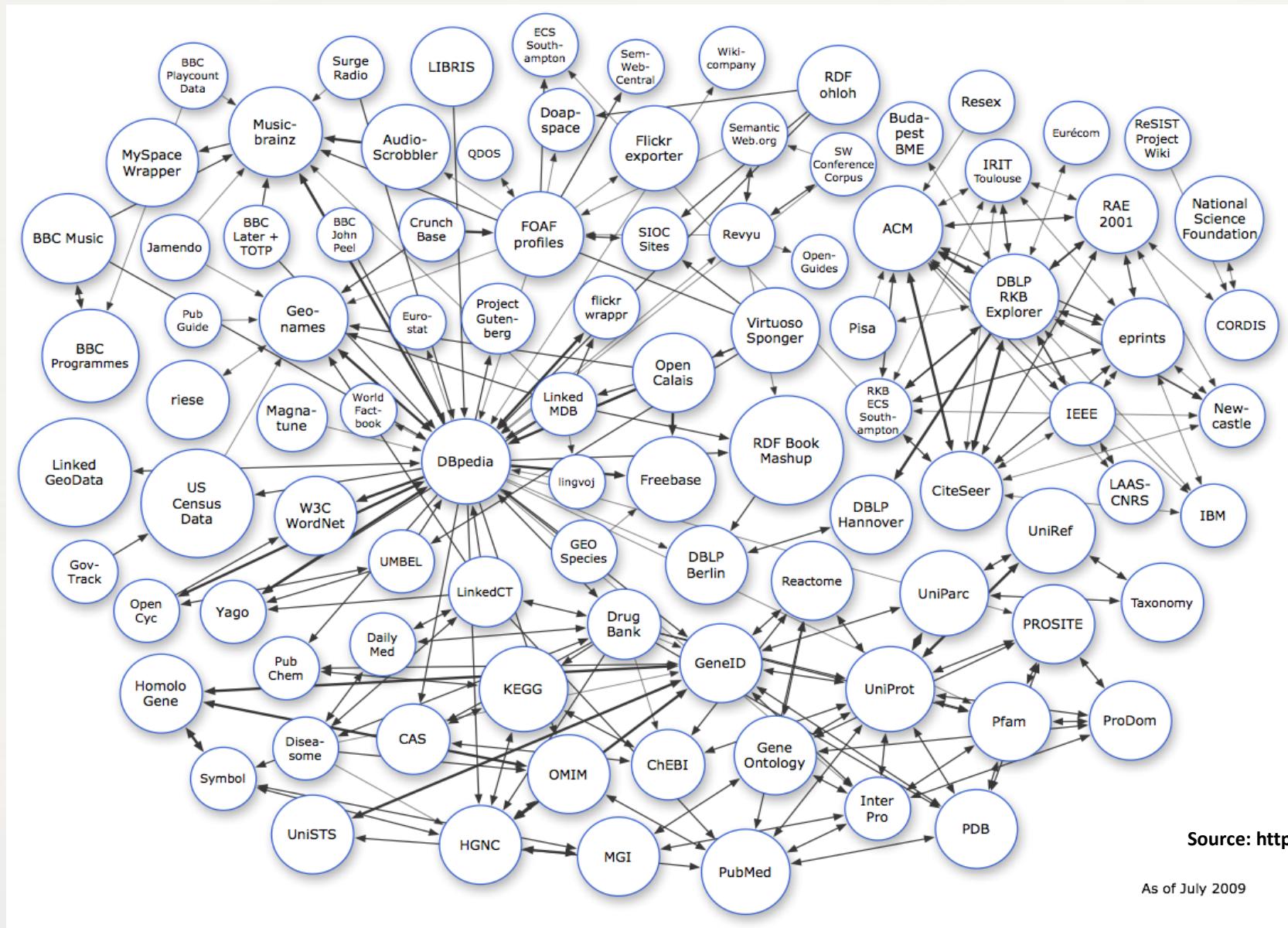
Datasets published following Linked Data 'format': **11/2007**

Linked Data



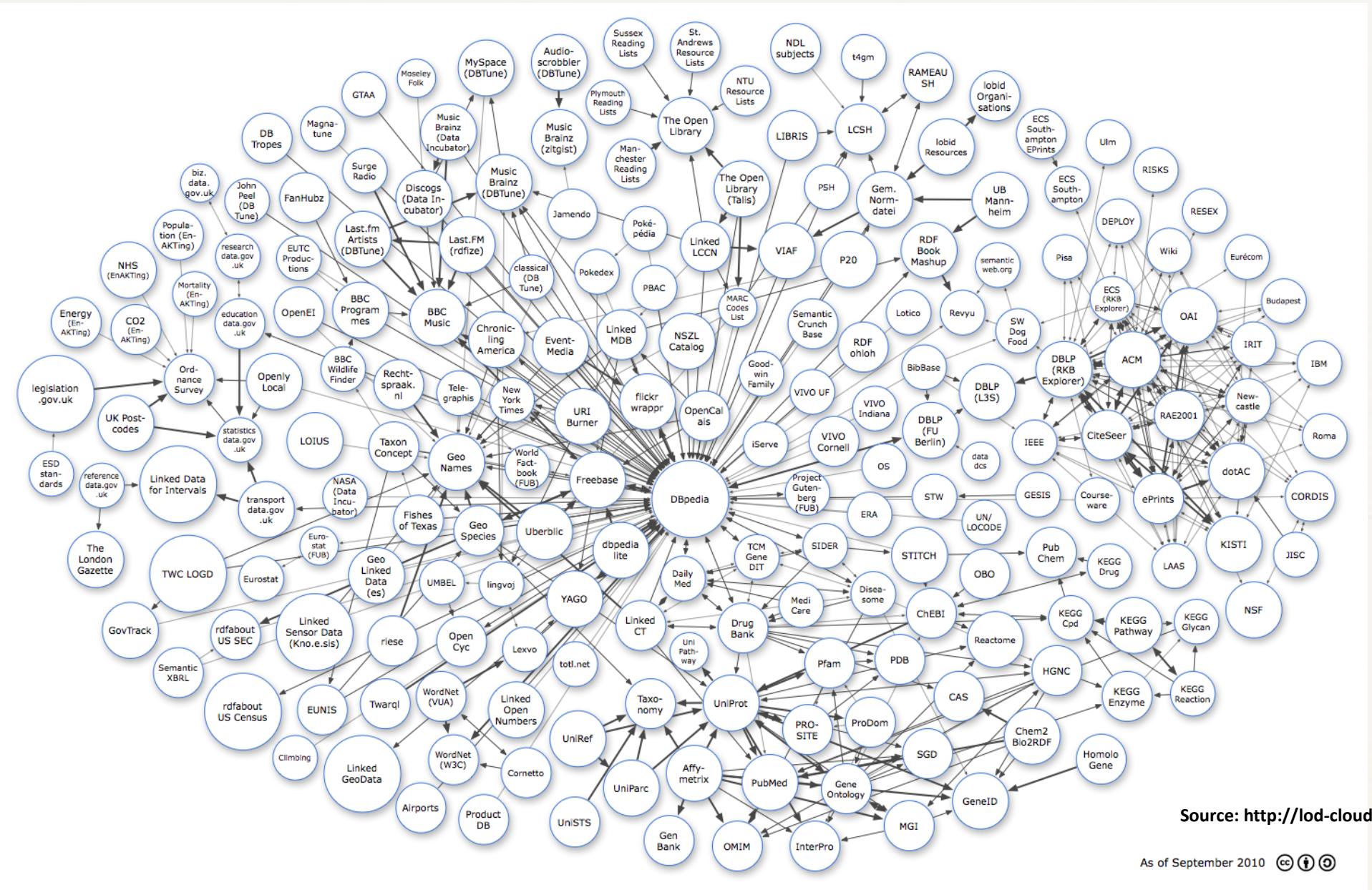
Datasets published following Linked Data ‘format’: 2008

Linked Data



Datasets published following Linked Data ‘format’: 2009

Linked Data

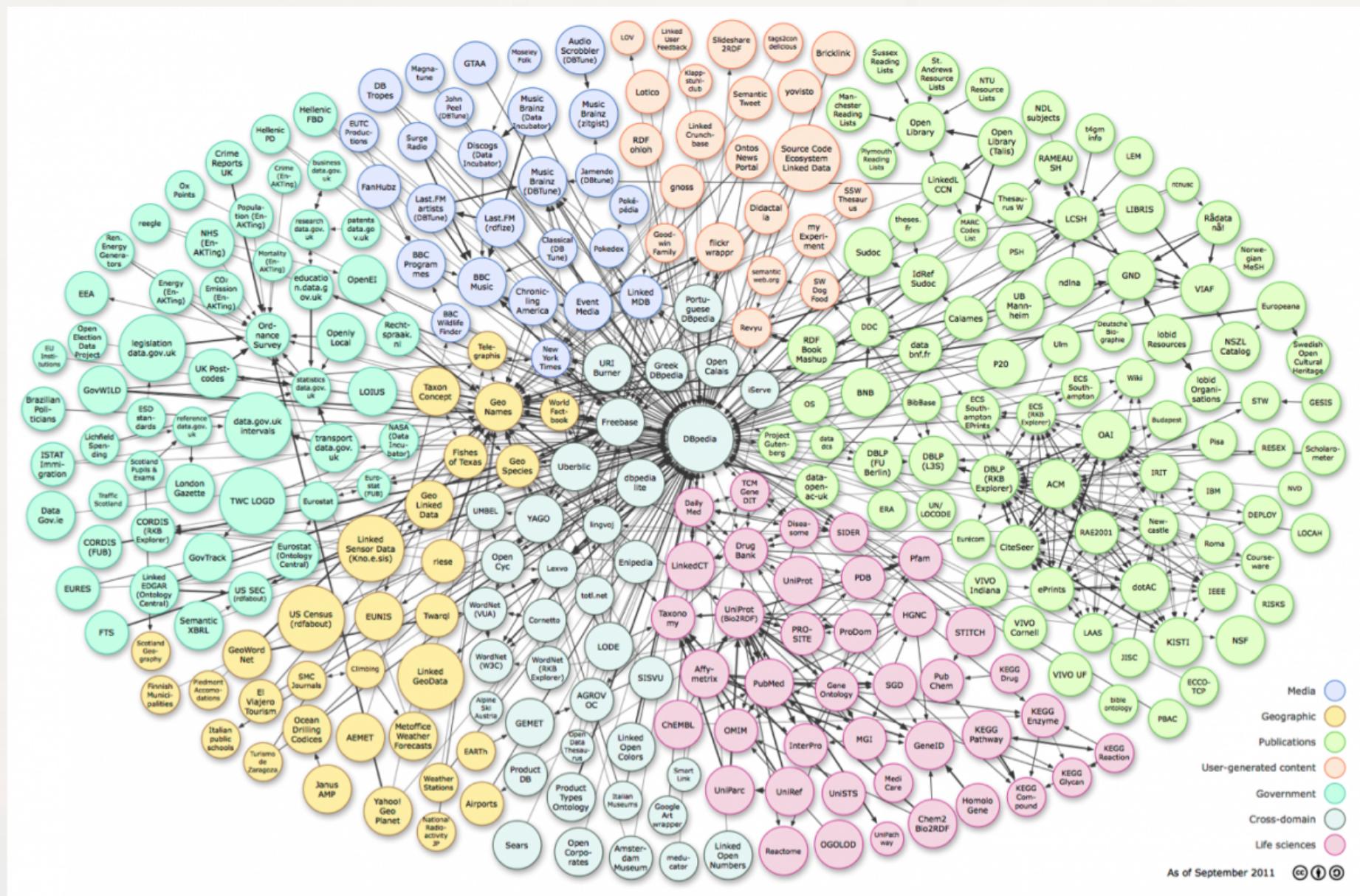


Source: <http://lod-cloud.net/>

As of September 2010

Datasets published following Linked Data ‘format’: 2010

Linked Data

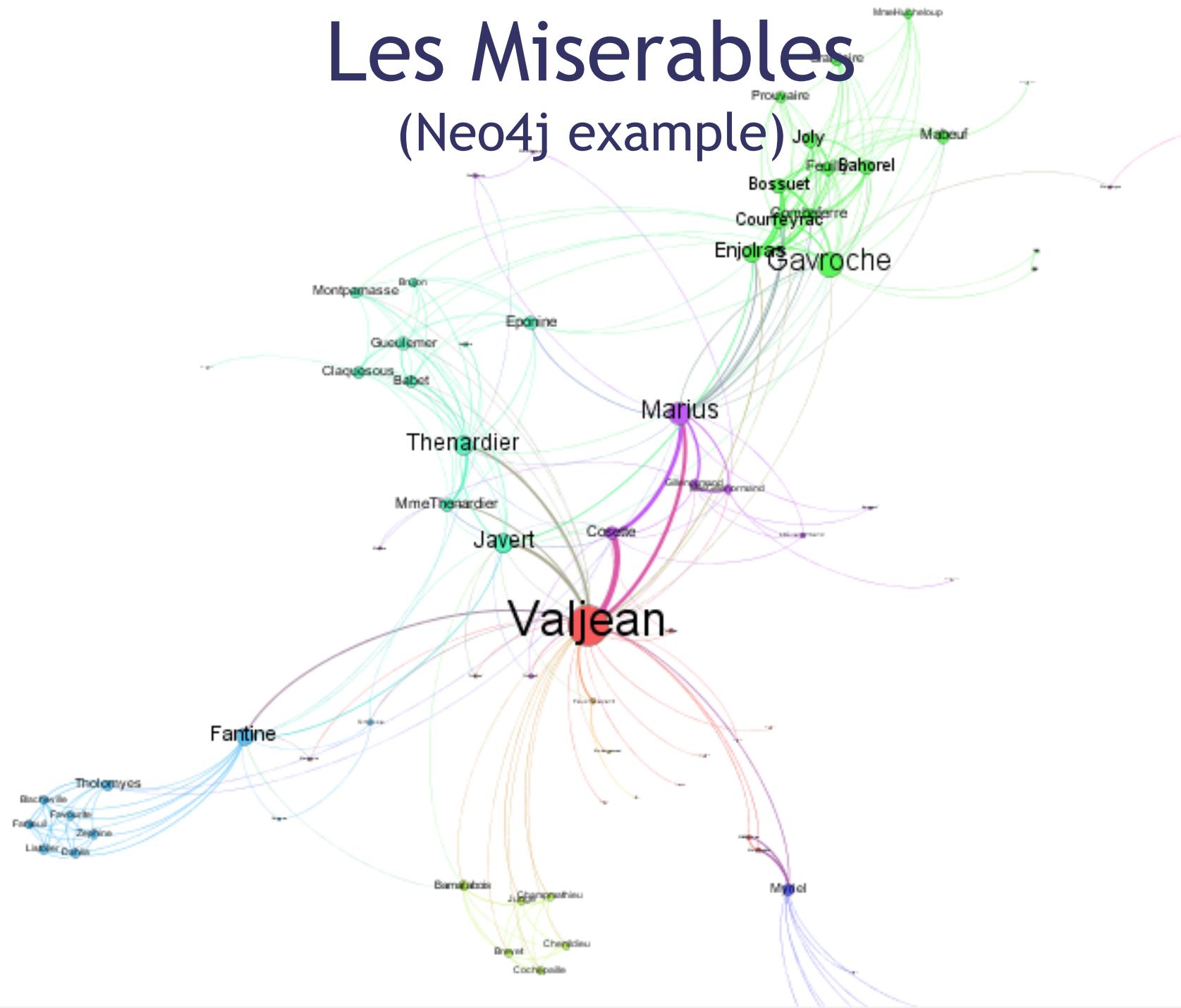


Datasets published following Linked Data ‘format’: 2011

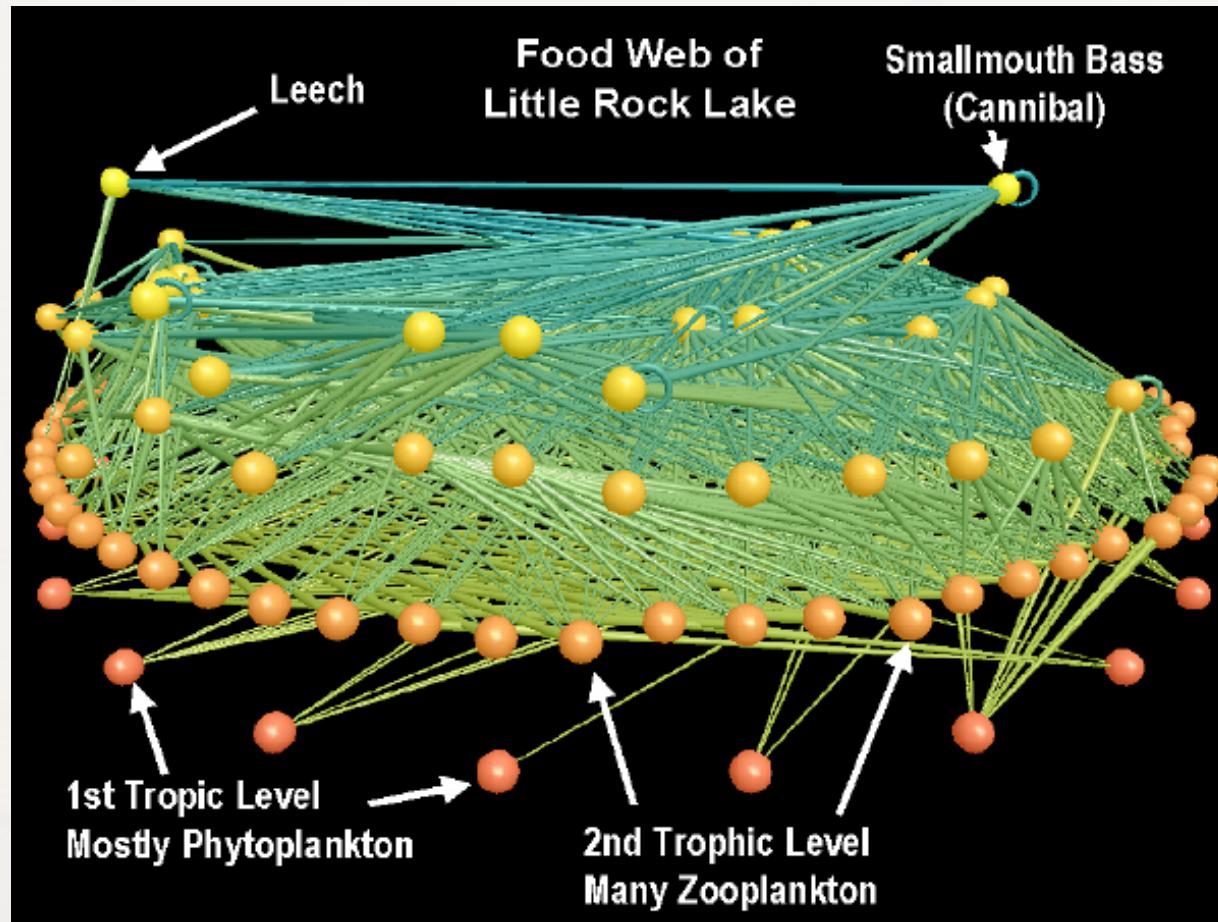
Les Misérables

Proudhon Joly

(Neo4j example)

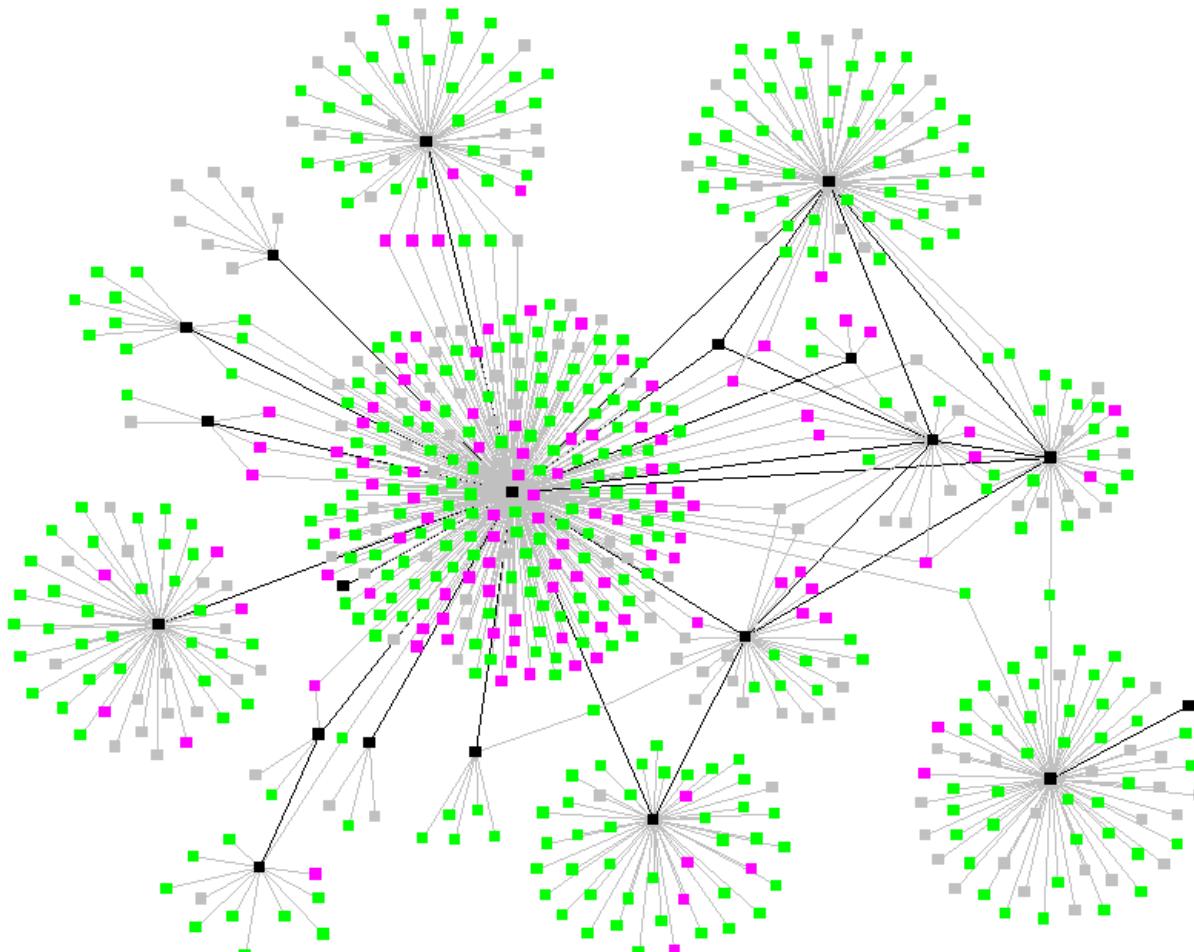


Freshwater food web



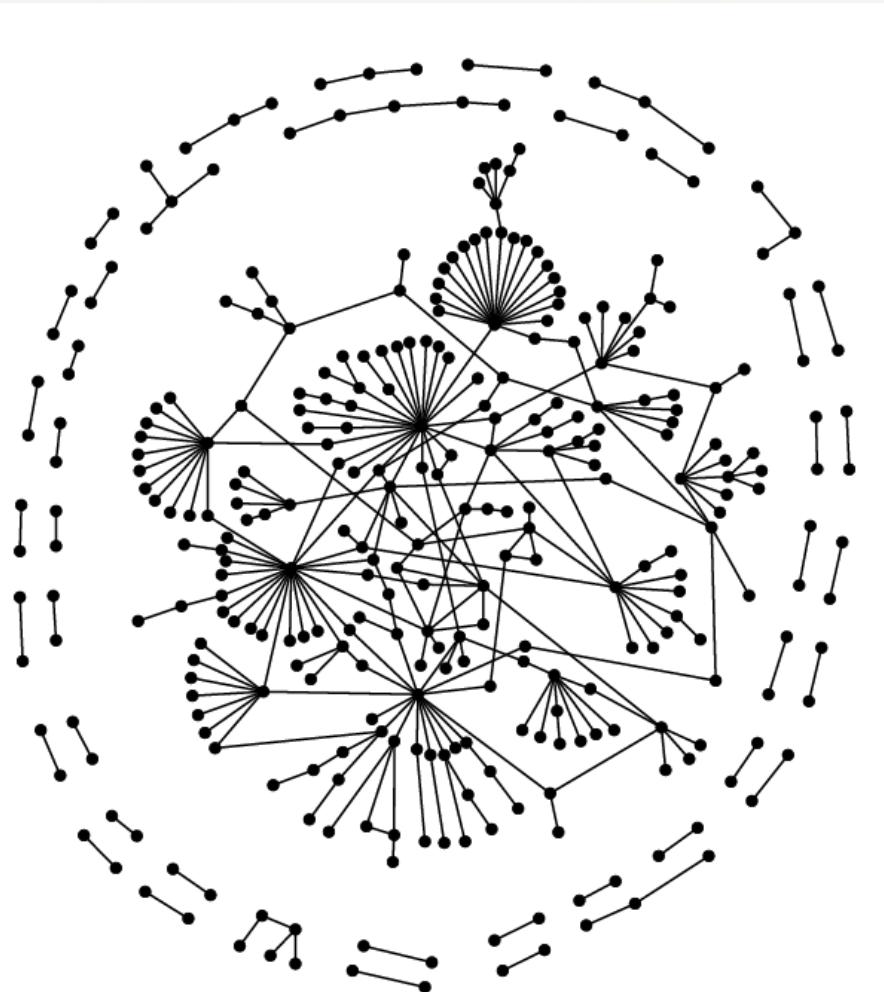
Freshwater food web: Neo Martinez and Richard Williams.

Contágio da TB



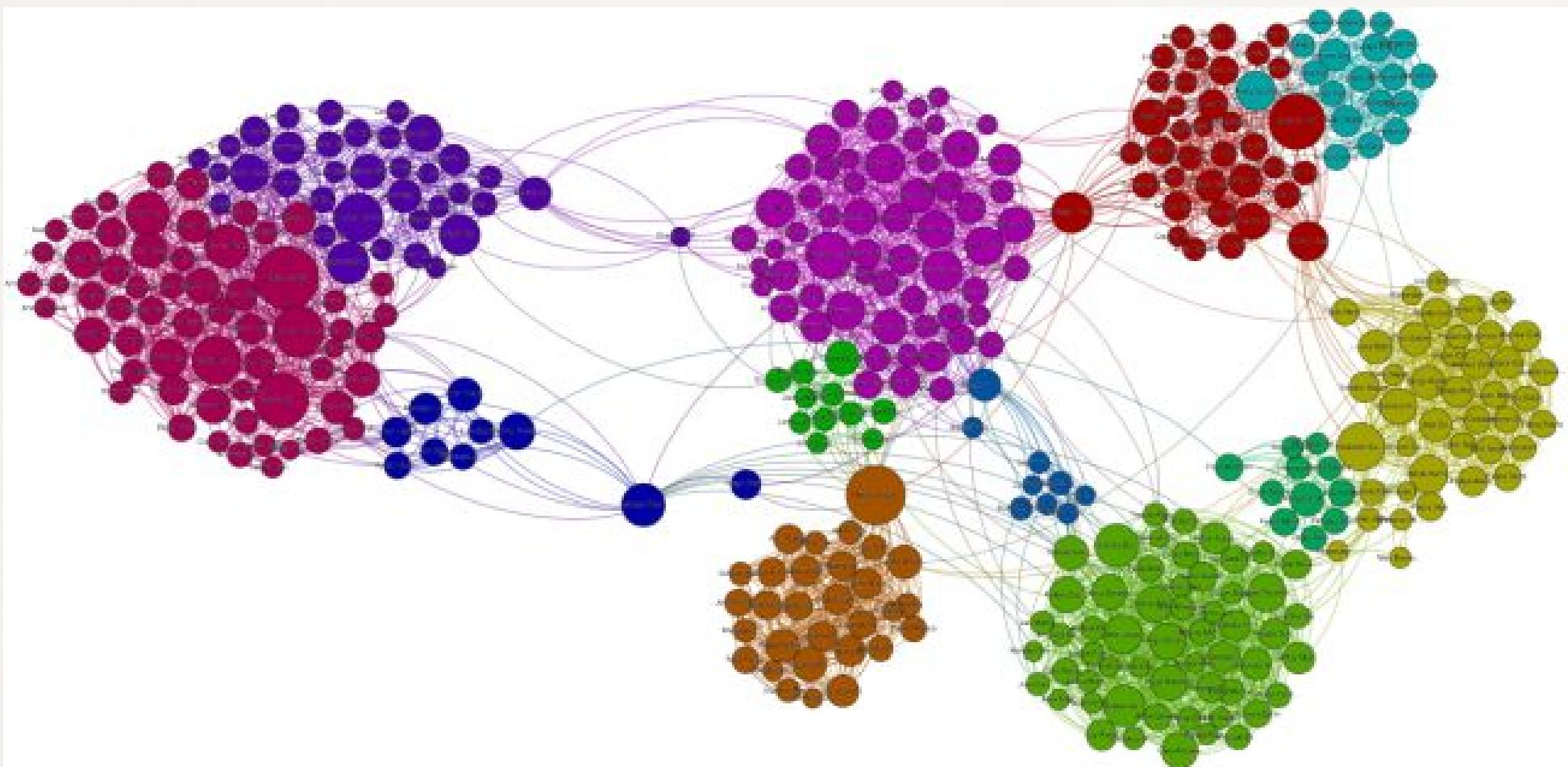
Contagion of TB, books on politics: Valdis Krebs, www.orgnet.com.

Proteínas da Levedura

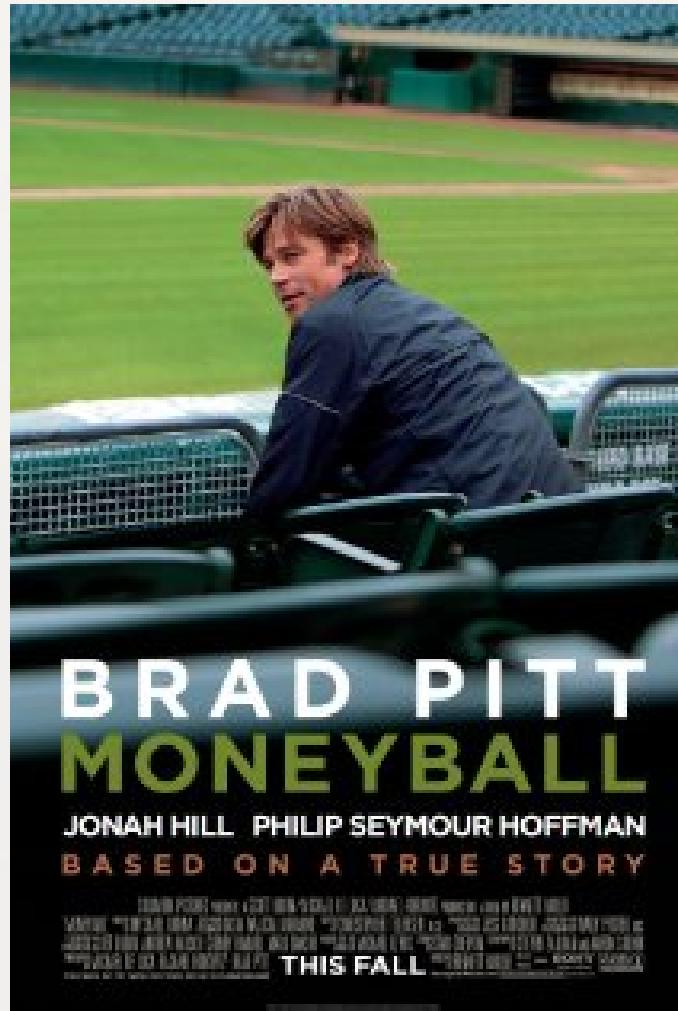


Yeast proteins: Sergei Maslov and Kim Sneppen,
Specificity and stability in topology of protein networks,
Science 296, 910-913 (2002).

Redes Sociais



Dados e Estratégia



Alemania e Big Data

SAP and Germany Make a Big Data Team at the World Cup

July 8, 2014 By Ben Hammonds

Sporttechie

<http://www.sporttechie.com/2014/07/08/sap-and-germany-make-smart-big-data-choices-at-world-cup/>

Alemanha e Big Data

- SAP is using Big Data to help the German coaching staff **make smart decisions** on tactics, player fitness, scouting, preparation as well as in game management. SAP has introduced a new concept called **SAP Match Insights** that assists players and coaches to prepare themselves for upcoming matches by dissecting key situations that may present themselves throughout the course of the match.

Segundo Bloco

- Data Mining
- Data Warehousing
- Information Retrieval
- Distributed and Parallel Databases
- BigData
- NoSQL

Tópicos

- Engenharia de Modelos
- BDs na Web e XML
 - XQuery
- Bds de Grafos
- Ontologias e Web Semântica
- O Quarto Paradigma
- BDs na cloud
- BDs em memória
- Bibliotecas Digitais
- Visual Analytics
- BDs Ativos
- Datalog

Referências

- Dijkstra, E. W. (1986) On a cultural gap. *The Mathematical Intelligencer*. vol. 8, no. 1, pp. 48-52.
- Ramakrishnan, Raghu; Gehrke, Johannes (2003) **Database Management Systems**. McGraw-Hill, 3rd edition.

André Santanchè

<http://www.ic.unicamp.br/~santanche>

Licença

- Estes slides são concedidos sob uma Licença Creative Commons. Sob as seguintes condições: Atribuição, Uso Não-Comercial e Compartilhamento pela mesma Licença.
- Mais detalhes sobre a referida licença Creative Commons veja no link:
<http://creativecommons.org/licenses/by-nc-sa/3.0/>
- Fotografia da capa feita por André Santanchè no Petit Palais (Paris) em 17/02/2013 do quadro: Fantasia à Constantinople de Felix Ziem