



Arquivos e Web Semântica

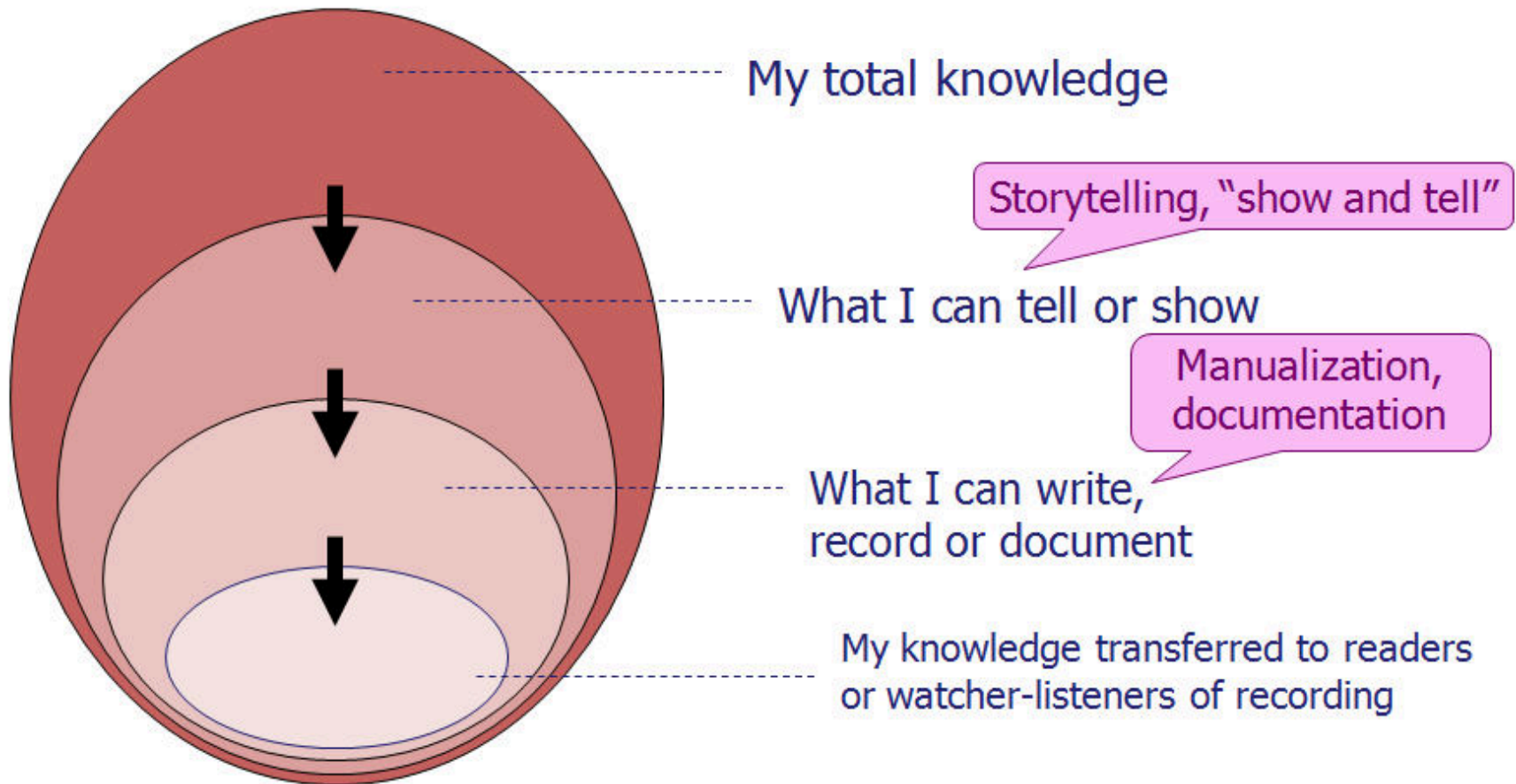
José Carlos Ramalho

jcr@di.uminho.pt

jcr@keep.pt

Ciclo de reflexões e debates promovido pelo Grupo de Documentos de
Arquivo (GT-GDA)

Conhecimento implícito



A nossa história

Um denominador comum:
geração automática de índices
(cronológico, toponímico,
antroponímico)

Facilidades: navegação
relacional e não hierárquica.

Período	
1989 - 1991	HiTeX – Transcrições
1998 - 2000	Ed. Eletrónica: “Memórias de José Inácio Gavetas do Cabido”; “Bulário Bracarense”; etc.
2000	Reverse Engineering da BD das “Inquirições de Género”
2003 – 2013	Digitarq: Gestão de metainformação, ODs, etc.
2006 – 2008	RODA: Preservação
2006 – 2013	CRAV: Consulta
2010 – 2014	SCAPE: projeto
2013 - ...	Archeevo: a nova
2013 – 2015	4C: projeto FP7
2014 - 2016	e-Ark (FP7): pan-European methodology for electronic document archiving

Milhões de registos
organizados
hierarquicamente:
ISAD, EAD, METS, etc.

Produção
massiva de
conteúdos

De regresso ao passado...

- HITEX (1989-1991): Um Sistema em Desenvolvimento para Historiadores e Arquivistas
 - Normalização: cada um deixar de fazer à sua maneira;
 - Reutilização: noção de componente reutilizável com interesse histórico;
 - Classificação: taxonomia padrão de conceitos históricos (classes) que exprime a ordem de subsunção sobre o conhecimento histórico;
 - Tolerância para com informação incompleta: permite a aquisição incremental de conhecimento histórico;
 - Resultados:
 - formato HiTeX: uma linguagem de fácil utilização para transcrição documental;
 - criação automática de índices: cronológico, toponímico e antroponímico.

3 focos recorrentes

- Toponímia;
- Antroponímia;
- Cronologia.

Um pequeno exemplo

“Certidão da doação que o arcebispo de Braga D.Martinho de Oliveira fez ao Cabido de Braga [: : :] Ano de 1300.”

Início do primeiro volume do Índice das Gavetas do Cabido de Braga

Asserções:

- D. Martinho de Oliveira era o Arcebispo de Braga em 1300;
- Este arcebispo doou qualquer coisa ao Cabido de Braga, nesse mesmo ano;
- O Cabido guardou uma certidão dessa doação;
- Essa certidão está arquivada no fundo documental designado Gavetas do Cabido;
- Uma referência a essa certidão é a que se pode encontrar no fol.1, vol.1 do correspondente Índice, compilado no século XVIII.

D. Martinho de Oliveira

Gavetas do Cabido

Certidão X

Índice

Modelo ontológico



Evolução da Web

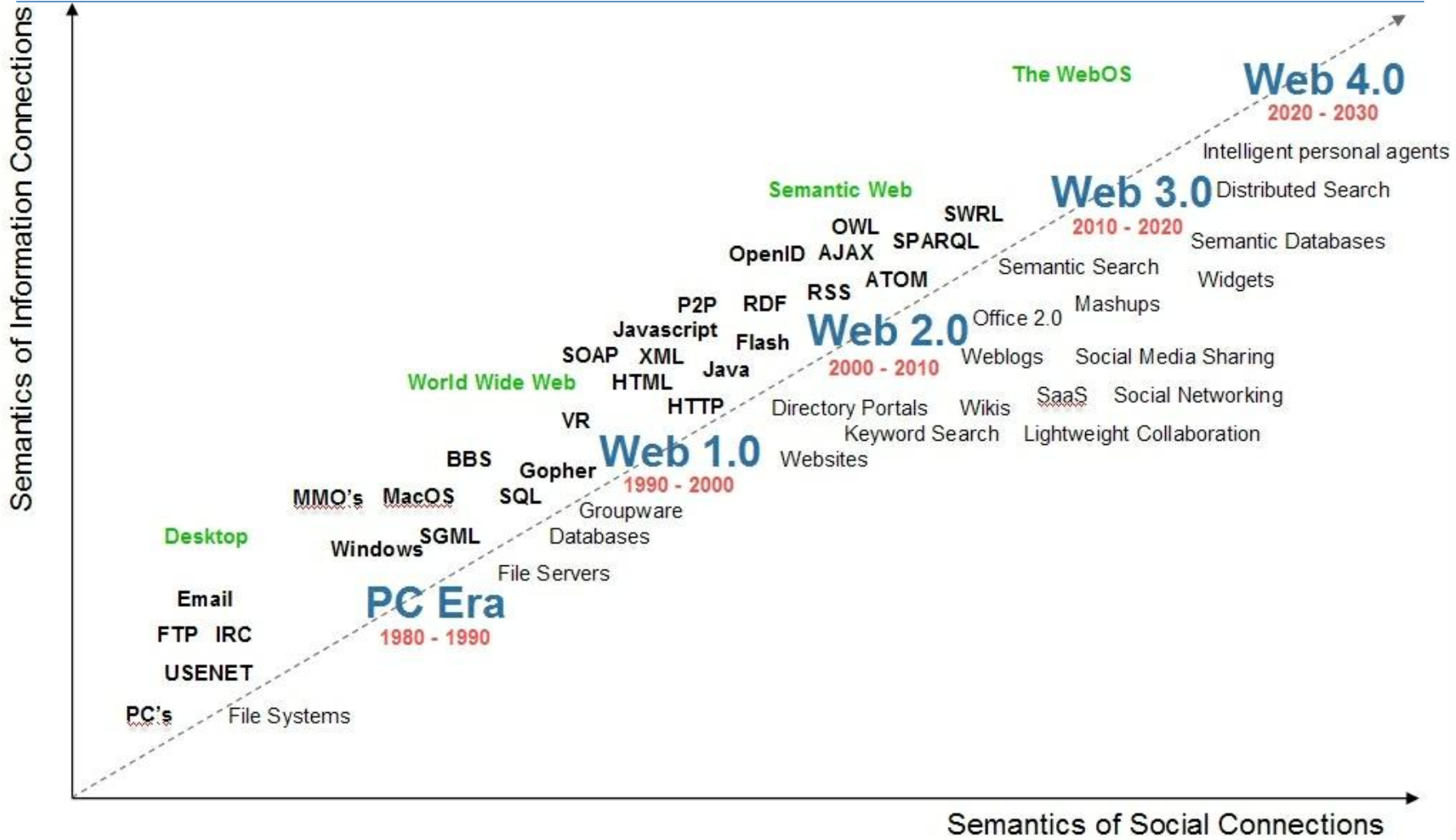
Web		Arquivos pt	
1ª geração	Páginas estáticas	1ª geração	Informatização
2ª geração	Páginas geradas a partir de conteúdos em SI	2ª geração	Acesso Web à informação em regime de acesso livre
3ª geração	Sites que se atualizam automaticamente mediante alterações de 3ªs partes: BBC • Integração automática de SI	3ª geração	<ul style="list-style-type: none"> • Integração • Dados com semântica • Novos modelos de acesso • Aplicações de valor acrescentado

É preciso aumentar o nível de percepção das “máquinas”...

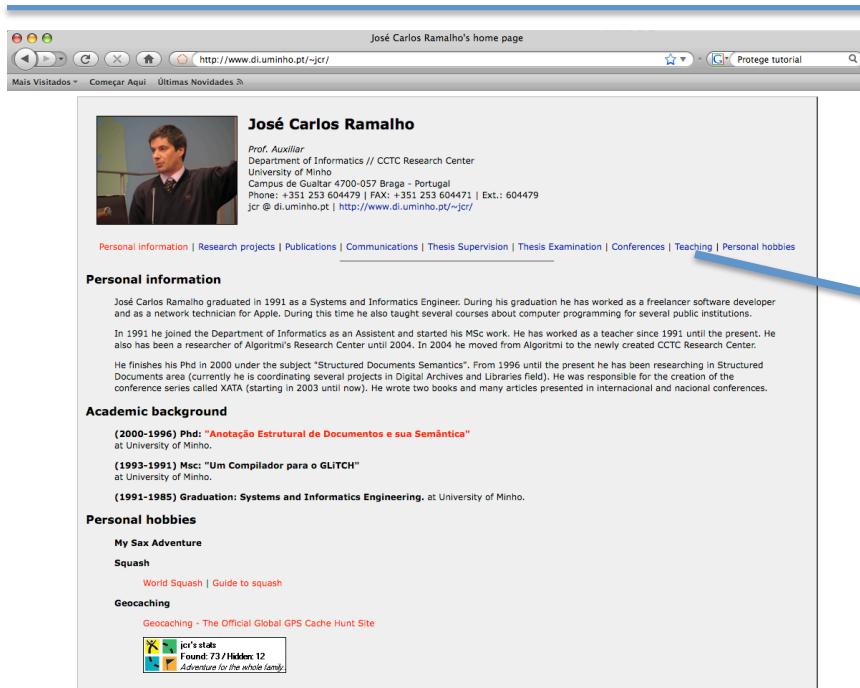
persistentes

• ...

Evolução



Hoje: a Web Sintáctica



Browser: José Carlos Ramalho's home page
URL: http://www.di.uminho.pt/~jcr/

José Carlos Ramalho
Prof. Auxiliar
Department of Informatics // CCTC Research Center
University of Minho
Campus de Gualtar 4700-057 Braga - Portugal
Phone: +351 253 604479 | FAX: +351 253 604471 | Ext.: 604479
jcr @ di.uminho.pt | http://www.di.uminho.pt/~jcr/

[Personal information](#) | [Research projects](#) | [Publications](#) | [Communications](#) | [Thesis Supervision](#) | [Thesis Examination](#) | [Conferences](#) | [Teaching](#) | [Personal hobbies](#)

Personal information

José Carlos Ramalho graduated in 1991 as a Systems and Informatics Engineer. During his graduation he has worked as a freelancer software developer and as a network technician for Apple. During this time he also taught several courses about computer programming for several public institutions.

In 1991 he joined the Department of Informatics as an Assistant and started his MSc work. He has worked as a teacher since 1991 until the present. He also has been a researcher of Algorithm's Research Center until 2004. In 2004 he moved from Algorithm to the newly created CCTC Research Center.

He finishes his PhD in 2000 under the subject "Structured Documents Semantics". From 1996 until the present he has been researching in Structured Documents area (currently he is coordinating several projects in Digital Archives and Libraries field). He was responsible for the creation of the conference series called XATA (starting in 2003 until now). He wrote two books and many articles presented in internacional and nacional conferences.

Academic background

(2000-1996) Phd: "Anotação Estrutural de Documentos e sua Semântica" at University of Minho.
(1993-1991) Msc: "Um Compilador para o GLITCH" at University of Minho.
(1991-1985) Graduation: **Systems and Informatics Engineering**, at University of Minho.

Personal hobbies

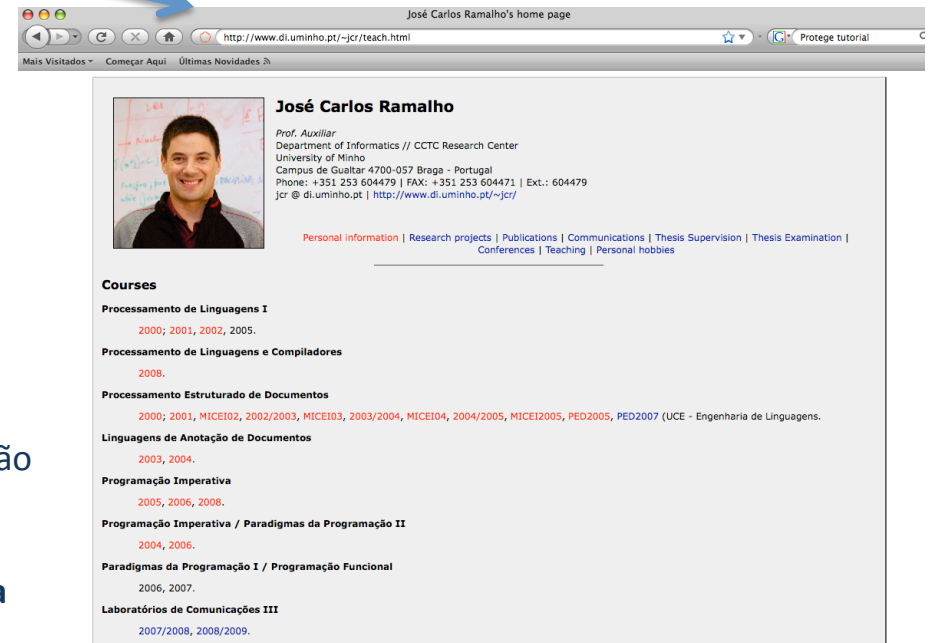
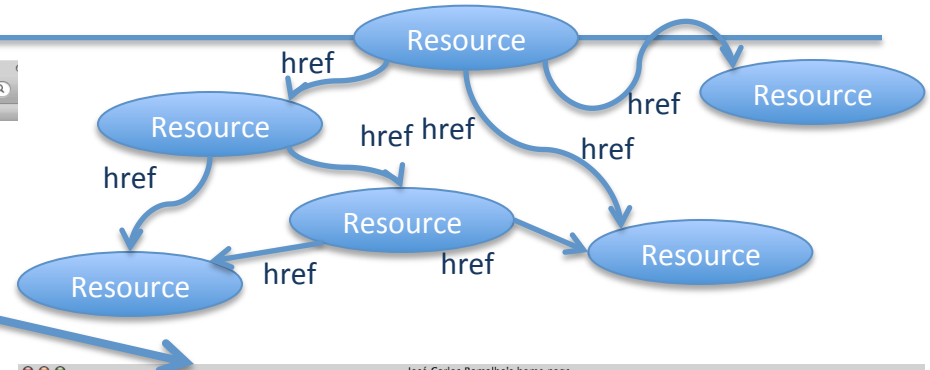

My Sax Adventure

Squash

[World Squash](#) | [Guide to squash](#)

Geocaching

[Geocaching - The Official Global GPS Cache Hunt Site](#)



Browser: José Carlos Ramalho's home page
URL: http://www.di.uminho.pt/~jcr/teach.html

José Carlos Ramalho
Prof. Auxiliar
Department of Informatics // CCTC Research Center
University of Minho
Campus de Gualtar 4700-057 Braga - Portugal
Phone: +351 253 604479 | FAX: +351 253 604471 | Ext.: 604479
jcr @ di.uminho.pt | http://www.di.uminho.pt/~jcr/

[Personal information](#) | [Research projects](#) | [Publications](#) | [Communications](#) | [Thesis Supervision](#) | [Thesis Examination](#) | [Conferences](#) | [Teaching](#) | [Personal hobbies](#)

Courses

Processamento de Linguagens I
2000; 2001, 2002, 2005.

Processamento de Linguagens e Compiladores
2008.

Processamento Estruturado de Documentos
2000; 2001, MICEI02, 2002/2003, MICEI03, 2003/2004, MICEI04, 2004/2005, MICEI2005, PED2005, PED2007 (UCE - Engenharia de Linguagens).

Linguagens de Anotação de Documentos
2003, 2004.

Programação Imperativa
2005, 2006, 2008.

Programação Imperativa / Paradigmas da Programação II
2004, 2006.

Paradigmas da Programação I / Programação Funcional
2006, 2007.

Laboratórios de Comunicações III
2007/2008, 2008/2009.

- Um lugar onde os computadores fazem a apresentação da informação (fácil) e as pessoas a interpretação e a navegação (difícil).
- **Ideia: Porque não colocar o computador a fazer uma maior parte do trabalho difícil?**

Exemplo: a mesma página...

José Carlos Ramalho's home page

Prof. Auxiliar
Department of Informatics // CCTC Research Center
University of Minho
Campus de Gualtar 4700-057 Braga - Portugal
Phone: +351 253 604479 | FAX: +351 253 604479
jcr @ di.uminho.pt | http://www.di.uminho.pt/~jcr

RDFa Play Documentation Tools Developers

Personal information | Research projects | Publications | Communications | Theses

Personal information

José Carlos Ramalho graduated in 1991 as a Systems and Informatics Engineering and as a network technician for Apple. During this time he also taught several courses. In 1991 he joined the Department of Informatics as an Assistant and started his research in the area of Algorithms' Research Center until 2004. In 2004 he finishes his PhD in 2000 under the subject "Structured Documents Semantics Documents area (currently he is coordinating several projects in Digital Archive conference series called XATA (starting in 2003 until now)). He wrote two books:

Academic background

(2000-1996) Phd: "Anotação Estrutural de Documentos e sua Semântica" at University of Minho.
(1993-1991) Msc: "Um Compilador para o GLITCH" at University of Minho.
(1991-1985) Graduation: Systems and Informatics Engineering, at University of Minho.

Personal hobbies

My Sax Adventure

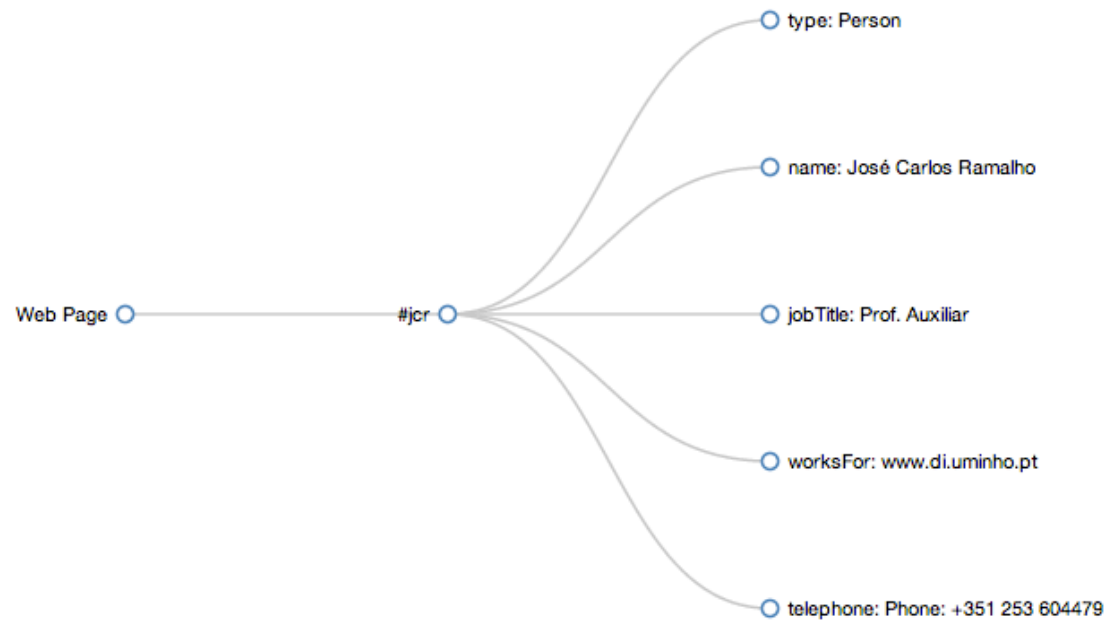
Squash

World Squash | Guide to squash

Geocaching

Geocaching - The Official Global GPS Cache Hunt Site

Ice's stats
Found: 73 / Hidden: 12
Adventure to the whole family!



A informação semântica

Extracted structured data

rdfa-node

type:

Person

relationship:

name:

worksFor

value:

Department of In
4700-057 Braga
jcr @ di.uminho.

href:

<http://www.di.uminho>

name:

workLocation

value:

Campus de Gual

href:

<http://www.geonames.org/2742032/braga.html>

property:

name:

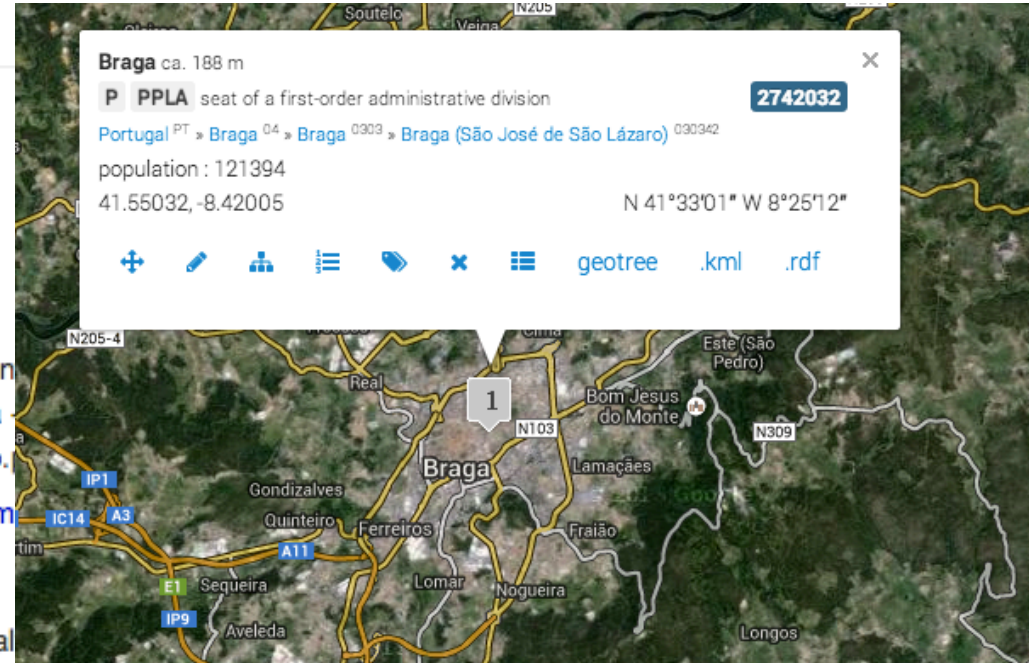
José Carlos Ramalho

jobTitle:

Prof. Auxiliar

telephone:

Phone: +351 253 604479



Consequências



phone josé carlos ramalho

Web

Images

Videos

Maps

News

More ▾

Search tools

About 2,540,000 results (0.58 seconds)

José Carlos Ramalho's home page

www.di.uminho.pt/~jcr ▾

Phone: +351 253 604479 | FAX: +351 253 604471 | Ext.: 604479 ... **José Carlos Ramalho** graduated in 1991 as a Systems and Informatics Engineer. During his ...

Exemplo: a Europeana

Forest SPARQL Search

Europeana Graph CHO Subject Object All Download in: JSON | RDF | N3/



[Carta régia de D. Afonso IV a Lisboa]

Source: <http://data.europeana.eu/item/09409a/12DB173F60974FB60DC4717C360E60F3C1CEFD63>
Description: D. Afonso IV confirma à cidade de Lisboa os foros e privilégios concedidos pelos reis seus antecessores.

Inference

Subject	Predicate	Object
http://data.europeana.eu/aggregation/provider/09409a...EFD63	rdf:type	ore:Aggregation
http://data.europeana.eu/aggregation/provider/09409a...EFD63	edm:aggregatedCHO	http://data.europeana.eu/item/09409a/12DB173F60974FB60DC4717C360E60F3C1CEFD63
http://data.europeana.eu/aggregation/provider/09409a...EFD63	edm:dataProvider	Arquivo Municipal de Lisboa

Apenas Semântica?!



SIMPLE SEARCH

ADVANCED SEARCH

FINDING AIDS

ONLINE SERVICES

HELP

- AC/FFH Fundo de Fomento da Habitação 1963-04/1987-07-
- SAALN Serviço de Apoio Ambulatório Local/Zona Norte 15
- GSM Gestão Social de Moradores 1974-06-01/1979-07-
- 014 Correspondência 1976/1976-07-23
- 24.042.001 Solicitações aos moradores do exterior

SOLICITAÇÕES AOS MORADORES DO EXTERIOR

DESCRIPTION LEVEL

Installation unit

REFERENCE CODE

PT/ADPRT/AC/FFH-SAALN/GSM/014/24.042.001

DATE RANGE

1976-05-07 ✓ to 1976-07-23 ✓

SCOPE AND CONTENT: Contém correspondência dirigida ao SAALN proveniente do exterior : "Liga para a Protecção da Natureza", "A Voz do Porto", "projecto omnicooper", sobre assuntos vários.

Inquirições de Género



RELATED MATERIAL

Série Inquirições de genere: Joao Araujo Sampaio, Irinao. Proc.19456.

DESCRIPTION LEVEL

dc File

REFERENCE CODE

PT/UM-ADB/DIO/MAB/006/00008

dc 00007 Inquirição de genere de Antonio Pereira 1680-10-

dc 00008 Inquirição de genere de Antonio Araujo Sampaio 1

dc 00009 Inquirição de genere de Antonio Pereira 1680-04-

TITLE TYPE

Atribuído

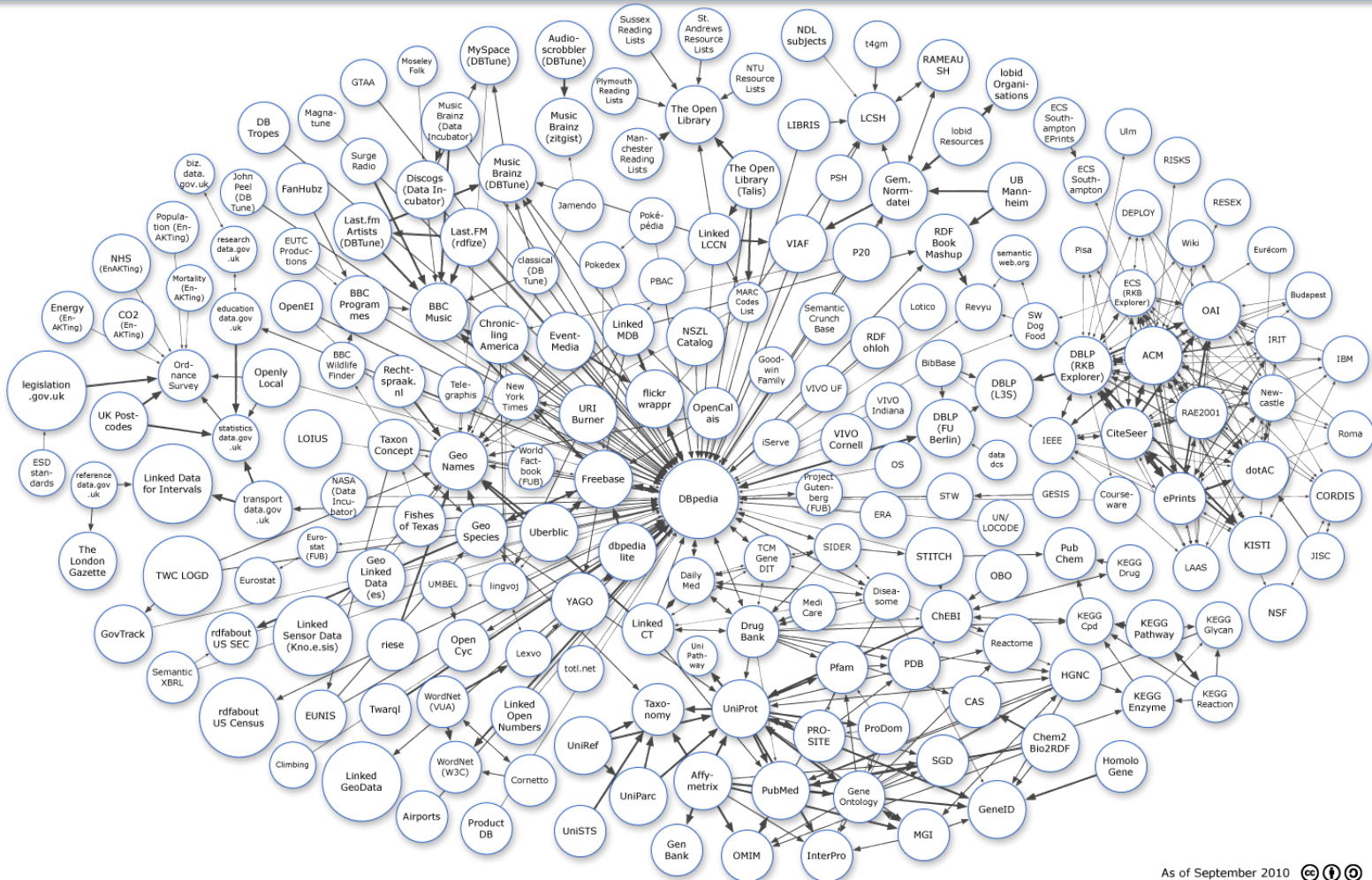
Semântica & “Linking”



Informação disponível
na Web não é
suficiente...

- É necessário que:
 - a informação esteja disponível na Web
 - acessível via tecnologias Web standard;
 - a informação esteja interligada através da Web;
 - ie, a informação pode ser integrada através da Web.

Linked Open Data, Set. 2010



Por áreas específicas

Gostaria que os sistemas de arquivo portugueses fizessem parte desta área nos próximos anos...



As opções da Europeiaana

- **Geonames**: 10 milhões de nomes/designações geográficos;
- **Dbpedia**: base de conhecimento que encerra a descrição de 3,64 milhões de itens;
- **Catalog of Life (CoL)**: catálogo completo de todos os organismos vivos conhecidos;
- **Uniprot**: banco de dados sobre sequenciação de proteínas;
- **GEMET**: thesaurus que define um vocabulário controlado para a temática ambiental;
- **FOAF**: “Friend of a Friend” – ontologia para descrever pessoas e respetivas relações.

Geonames

The screenshot displays the Geonames web application interface. On the left, a hierarchical tree view shows administrative divisions for Portugal (PT). The tree is expanded to show the following structure:

- Portugal (PT)
 - Autonomous Region of the Algarve (01)
 - Aveiro (02)
 - Bragança (05)
 - Distrito da Guarda (11)
 - Distrito de Beja (03)
 - Distrito de Braga (04)
 - Distrito de Castelo Branco (06)
 - Distrito de Coimbra (07)
 - Distrito de Faro (09)
 - Albufeira Municipality (0801)
 - Alcoutim (0802)
 - Aljezur (0803)
 - Castro Marim (0804)
 - Faro (0805)
 - Lagoa (0806)
 - Lagos (0807)
 - Loulé (0808)
 - Almancil (080801)**
 - Alte (080802)
 - Ameixial (080803)
 - Benafim (080810)
 - Boliqueime (080809)

- Guam (GU) [GUM]

On the right, a map shows the location of Almancil. A pop-up window provides the following information:

- Almancil** ca. 68 m
- ADM3** third-order administrative division
- 8012002
- Portugal PT » Faro 09 » Loulé 0808 » Almancil 080801
- Layers 4, -8.03186
- Coordinates: N 37°05'23" W 8°01'55"
- Export options: geotree, .kml, .rdf

Descrição arquivística com semântica

1. É preciso identificar univocamente os registos descritivos: utilizando URIs persistentes;
2. Os registos descritivos devem estar num formato normalizado de modo a que recursos e respetivas relações possam ser automaticamente reconhecidos (RDF);
3. Incluir nas descrições o maior número de links (mínimo = 50 links externos / dataset) possíveis para outros recursos de informação.

Linguagens para esp. semântica

- RDF;
- [SKOS](#);
- OWL;
- RDFa (anotações leves para serem usadas em páginas Web).

Vocabulários

- Para integrar dados é necessário acordar
 - nos termos a usar:
 - “translator”, “author”
 - nas categorias a usar:
 - “Person”, “literature”
 - nas relações entre aqueles:
 - “an author is also a Person...”, “historical fiction is a narrower term than fiction”
 - novas relações podem ser deduzidas.

O que é preciso?

- Acordo num vocabulário;
- Acordo numa ontologia;
- Embeber esta semântica no “front-end” dos sistemas de informação atuais (experiências em curso);
- No fim: vontade e força política (difícil mas não impossível).

Vocabulários

- É necessária uma linguagem de especificação formal para estes vocabulários;
- Para os definir;
- Para lhes associar semântica de modo a tornar claro como é que novas relações podem ser deduzidas.

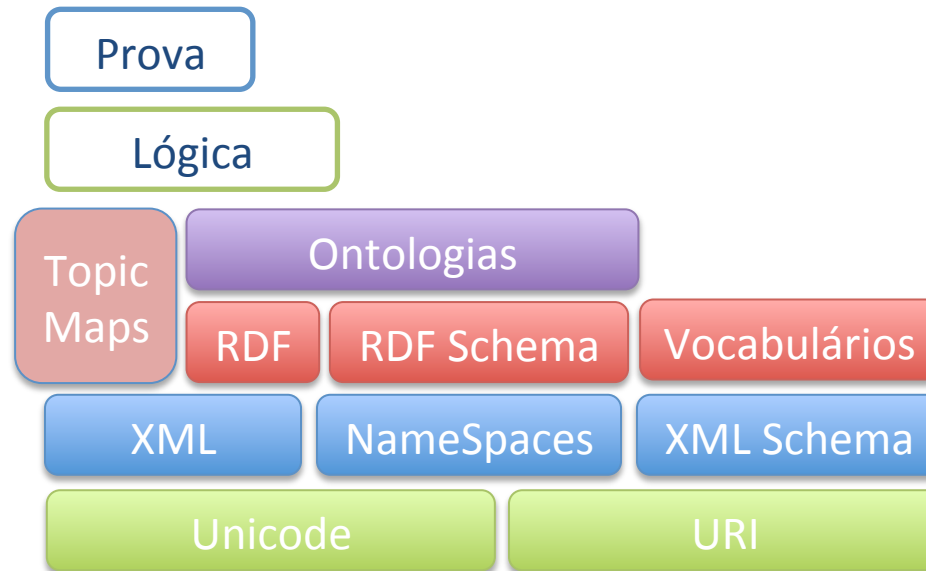
Ontologias

- Definição:

“Especificação formal de uma área de conhecimento através da definição dos conceitos que lhe são subjacentes e relações entre estes.”

- Conhecimento explícito (fácil);
- Conhecimento implícito (difícil, é preciso torná-lo explícito).

Do bit à Semântica:



Último exemplo: Inquirições de “Génere”

- Motor SPARQL: <http://librdf.org/query>
- Ontologia light: <http://www.di.uminho.pt/~jcr/XML/sparql/inquiricoes-light.owl>

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>  
PREFIX owl: <http://www.w3.org/2002/07/owl#>  
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema>  
PREFIX inq: <http://jcr.di.uminho.pt/ontologias/rc2012/inq.owl#>
```

Interrogações: Nomes dos indivíduos?

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX inq: <http://jcr.di.uminho.pt/ontologias/rc2#>

SELECT ?na
WHERE {
  ?a inq:nome ?na.
}
```

Count	na
1	Amaro Alvares Carvalho
2	Antonio Garcia Correia
3	Antonio Brandao Coelho
4	Antonio Alvares Carvalho
5	Agostinho Machado
6	Antonio Barbosa Veloso
7	Antonio Silva Marinho
8	Antonio Magalhaes
9	Antonio Pereira
10	Andre Vieira Carvalho
11	Agostinho Silva
12	Antonio Jorge
13	Antonio Goncalves
14	Antonio Macedo
15	Manuel Araujo Pereira
16	Alexandre Carneiro
17	Antonio Carneiro Felgueiras Gaio
18	Alexandre Barros
19	Antonio Rocha Araujo
20	Andre Coelho

Os pares de nomes dos que são Irmãos?

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX inq: <http://jcr
```

```
SELECT ?na, ?nb
WHERE {
  ?a inq:Irmao ?b.
  ?a inq:nome ?na.
  ?b inq:nome ?nb
}
```

Count	na	nb
1	Joao Pereira Pinto Tavora	Joao Pereira Pinto Tavora
2	Joao Pereira Pinto Tavora	Diogo Pereira Pinto
3	Joao Pereira Pinto Tavora	Antonio Pereira Pinto Tavora
4	Diogo Pereira Pinto	Joao Pereira Pinto Tavora
5	Diogo Pereira Pinto	Diogo Pereira Pinto
6	Diogo Pereira Pinto	Antonio Pereira Pinto Tavora
7	Antonio Pereira Pinto Tavora	Joao Pereira Pinto Tavora
8	Antonio Pereira Pinto Tavora	Diogo Pereira Pinto
9	Antonio Pereira Pinto Tavora	Antonio Pereira Pinto Tavora
10	Manuel Carvalho	Manuel Carvalho

Algumas recomendações

- Publica a tua informação primeiro, preocupa-te com as interfaces/aparência depois!
 - a “informação crua” pode ser útil assim mesmo e outros poderão usá-la;
 - podes acrescentar-lhe valor, mais tarde, criando-lhe acessos sofisticados.
- Se possível, publica a tua informação em RDF caso contrário, outros poderão ajudar-te na conversão
 - confia na comunidade...
- *Acrésceta links para outras fontes.* “Apenas” publicar não é suficiente...

Posso começar pela minha homepage



José Carlos Ramalho

Prof. Auxiliar
Departamento de
Informática
Universidade de
Algarve
Campus de Faro
Phone: +351 289 330 200
jcr@di.uam.faro.pt

Personal

Publications

[2013] [2012] [2011] [2010] [2009] [2008]
[1996] [1995] [1993]

2013 (1)

- [inproceedings, archiving2013] - " **Guidelines for legacy repository migration**"; IS&T Archiving 2013; Washington DC; <http://hdl.handle.net/1822/23978>;
- [inbook, IGI2013] - *Ricardo Freitas, José Carlos Ramalho* - " **Innovations in XML Applications and Metadata Management** "; Database Preservation: Using Ontologies and Semantic Web; IGI Global; 2013.

Extracted structured data

rdfa-node

type: Publication

property:

hasTitle: Guidelines for legacy repository migration

hasAuthors: Miguel Ferreira

hasAuthors: José Carlos Ramalho

hasAuthors: Luis Faria

rdfa-node

type: Publication

property:

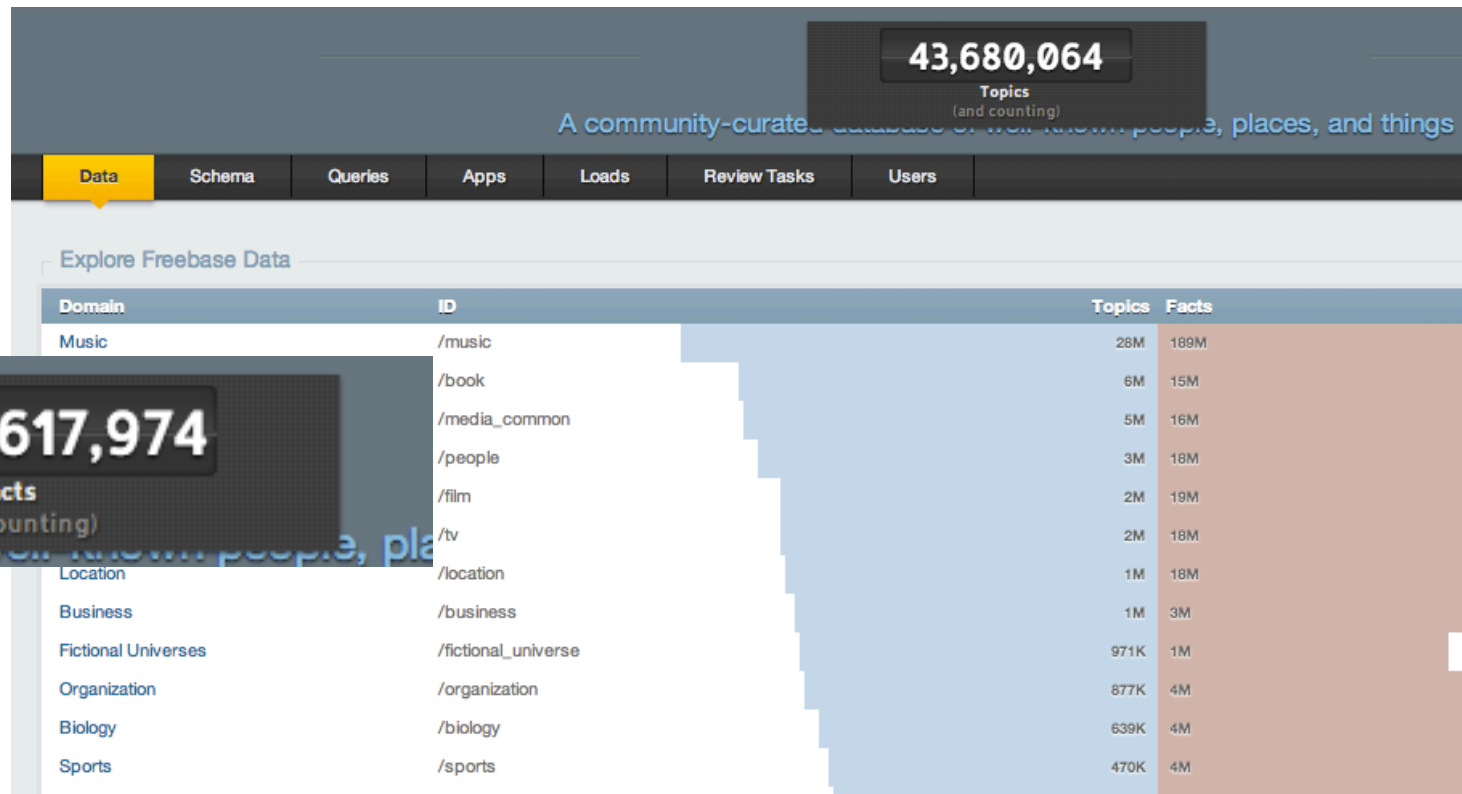
hasAuthors: Ricardo Freitas

hasAuthors: José Carlos Ramalho

hasTitle: Innovations in XML Applications and Metadata Management

Há sempre “vozes” divergentes...

- Google: www.freebase.com



The screenshot shows the Freebase website interface. At the top right, a large black box displays the number **43,680,064** with the text **Topics** and *(and counting)* below it. Below this, a navigation bar contains tabs for **Data**, **Schema**, **Queries**, **Apps**, **Loads**, **Review Tasks**, and **Users**. The main content area is titled "Explore Freebase Data" and features a table with the following columns: **Domain**, **ID**, **Topics**, and **Facts**. A large black box on the left side of the table displays the number **2,493,617,974** with the text **Facts** and *(and counting)* below it. The table lists various domains and their corresponding statistics:

Domain	ID	Topics	Facts
Music	/music	28M	189M
	/book	6M	15M
	/media_common	5M	16M
	/people	3M	18M
	/film	2M	19M
	/tv	2M	18M
Location	/location	1M	18M
Business	/business	1M	3M
Fictional Universes	/fictional_universe	971K	1M
Organization	/organization	877K	4M
Biology	/biology	639K	4M
Sports	/sports	470K	4M

Acessível e “integrável”

- Estrutura simples de 3 níveis: domínio, tipo e propriedade;
- Query by example with a simple language.

Freebase: exemplo

Find topics of type /film/director

Try it out:

Select an item from the list:

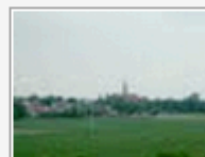
Steven Spielberg	Film director
Roland Emmerich	Film director
Götz Spielmann	Film director
Brad Armstrong	Pornographic actor
Godfrey Ho	Film director
Robert Beaucage	Scriptwriter
Wolfgang Staudte	Film director

Freebase: exemplo

Bozza

Select an item from the list:

Eugène Bozza	Composer
Anthony Bozza	Writer
Tullio Bozza	Olympic athlete
Bozzai	City/Town/Village
Bozza	Canonical Version
Bozza	Organization
Piero Bozza	Film editor



Bozzai

/m/02rjf4l

Time zone(s): **Central European Time Zone**

Area: **4.43**

Bozzai is a village in Vas County, Hungary.

City/Town/Village

Oportunidades

- Horizonte 2020: este é um dos eixos principais de financiamento na área das TIC;
- Criação de novos produtos: site da BBC, Retrievo, Archeevo, ...
- Se ligarmos à LOD o nosso repositório estamos a maximizar a sua visibilidade e a aumentar potencialmente o número de visitas;
- Os motores de busca tipo google já tiram partido da informação semântica se esta estiver disponível.

Dificuldades

- Se descrever arquivisticamente consome recursos, descrever conteúdos consome ainda mais;
- Voluntariado (folksonomias na prática)
 - é preciso preparar o terreno;
 - controlar/definir os vocabulários usados;
 - criar política de contribuições.

Mas já podemos ir fazendo...

The screenshot shows the archeevo website interface. At the top, there is a dark red header with the archeevo logo on the left, a home icon, and the text 'SIMPLE SEARCH'. Below the header, the main content area is divided into two columns. The left column is titled 'CLASSIFICATION SCHEME' and contains a blue folder icon followed by 'JUD/TCBRG Tribunal da Comarca de Braga' and a red document icon followed by '00033 Inventário por Óbito de Alzira Jesus Rodrigues 1'. The right column is titled 'INVENTÁRIO POR ÓBITO DE ALZIRA JESUS RODRIGUES' and contains a yellow box with the text 'Record not reviewed.', followed by 'DESCRIPTION LEVEL' with a red document icon and 'File', 'REFERENCE CODE' with 'PT/UM-ADB/JUD/TCBRG/00033', and 'DATE RANGE' with '1995' and a green checkmark. Three blue arrows point from labels at the bottom to specific parts of the record: 'Toponímia' points to 'Tribunal da Comarca de Braga', 'Cronologia' points to '1995', and 'Antroponímia' points to 'ALZIRA JESUS RODRIGUES'.

archeevo | SIMPLE SEARCH

CLASSIFICATION SCHEME

JUD/TCBRG Tribunal da Comarca de Braga

00033 Inventário por Óbito de Alzira Jesus Rodrigues 1

INVENTÁRIO POR ÓBITO DE ALZIRA JESUS RODRIGUES

Record not reviewed.

DESCRIPTION LEVEL

File

REFERENCE CODE

PT/UM-ADB/JUD/TCBRG/00033

DATE RANGE

1995 ✓

Toponímia Cronologia Antroponímia

Schema.org

schema.org

Home Schemas Documentati

Home Schemas Docume

Thing > Place

A person (alive, de

Entities that have a somewhat fixed, physical extension.

Property	Expected Type	Description
Properties from Place		
address	PostalAddress	Physical address of the item.
aggregateRating	AggregateRating	The overall rating, based on a collection of reviews or ratings, of the item.
containedIn	Place	The basic containment relation between places.
event	Event	Upcoming or past event associated with this place or organization. Supercedes events .
faxNumber	Text	The fax number.
geo	GeoCoordinates or GeoShape	The geo coordinates of the place.

Exemplo: a minha homepage



José Carlos Ramalho

Prof. Auxiliar

Department of Informatics // CCTC Research Center
University of Minho

Campus de Gualtar 4700-057 Braga - Portugal

Phone: +351 253 604479 | FAX: +351 253 604471 | Ext.: 604479

jcr @ di.uminho.pt | <http://www.di.uminho.pt/~jcr/>

<http://www.di.uminho.pt/~jcr>

O que foi adicionado...

```
<h1><span property="name">José Carlos Ramalho</span></h1>
<p><i property="jobTitle">Prof. Auxiliar</i> <br/>
<span property="worksFor" resource="http://
www.di.uminho.pt">Department of Informatics // CCTC
Research Center<br/> University of Minho<br/>
<span property="workLocation" resource="http://
www.geonames.org/2742032/braga.html"><a href="http://
www.geonames.org/2742032/braga.html">Campus de
Gualtar 4700-057 Braga - Portugal</a></span><br/>
<span property="telephone">Phone: +351 253 604479</
span> | FAX: +351 253 604471 | Ext.: 604479<br/>
    jcr @ di.uminho.pt | <a href="http://www.di.uminho.pt/
%7Ejcr/">http://www.di.uminho.pt/~jcr/</a> </p>
```

Algumas Ferramentas

- [Google structured data testing tool](#)
- [Yandex Structured Data validator](#)
- [RDFa Play](#)
- [Structured data linter](#)

Algumas questões...

- O que vamos anotar?
- Como? Com que vocabulários?
- Que ligações queremos fazer?

Exemplo: www.retrievo.pt



SCIENCE
IN THE CLOUD

6.310.273 registos de 181 bases de dados

PROCURAR CIÊNCIA

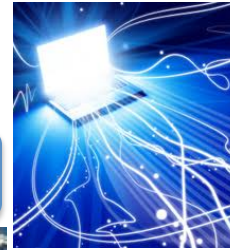
 **RETRIEVO, agregador de conteúdos e pesquisa federada**

O Retrievo é um portal que permite a pesquisa e recuperação de conteúdos armazenados em diferentes fontes de informação. Trata-se de uma solução que disponibiliza um ponto de acesso único e privilegiado a todos os ativos de informação disponíveis na sua organização ou oriundos de fontes de informação externas (e.g. bases de dados científicas).

Pensamentos “soltos”



Informação



Estamos algures entre a Informação e o Conhecimento...



Linguagem



Conhecimento





José Carlos Ramalho
KEEPS/Universidade do Minho
jcr@keep.pt / jcr@di.uminho.pt

Questões?

KEEPSOLUTIONS
University of Minho (SPIn-CIF)

<http://www.keep.pt>