Projetos experimentais de ontologias e suas potenciais aplicações

I Oficina do GT Glossário da Cultura

Ministério da Cultura

Brasília, Fevereiro de 2016

Prof. Dalton Martins

dmartins@gmail.com

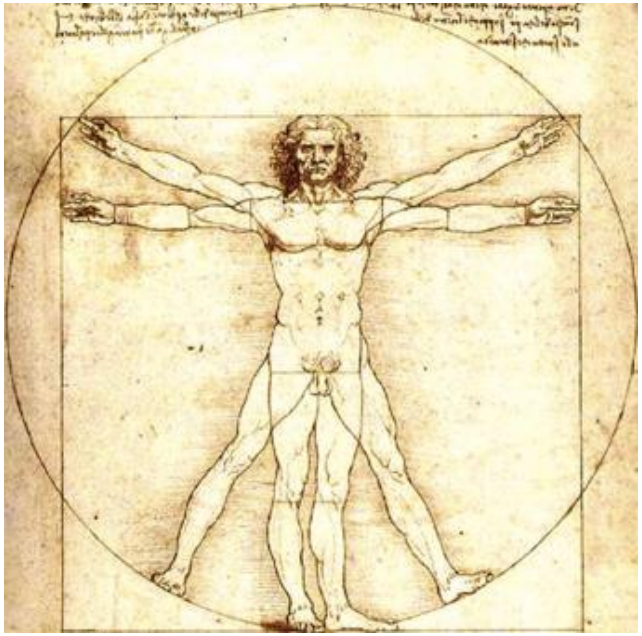
Prof. Marcel Ferrante Silva

marcelf@gmail.com

Laboratório de Políticas Públicas Participativas

Universidade Federal de Goiás

www.l3p.fic.ufg.br



Quais as vantagens de construir uma ontologia?

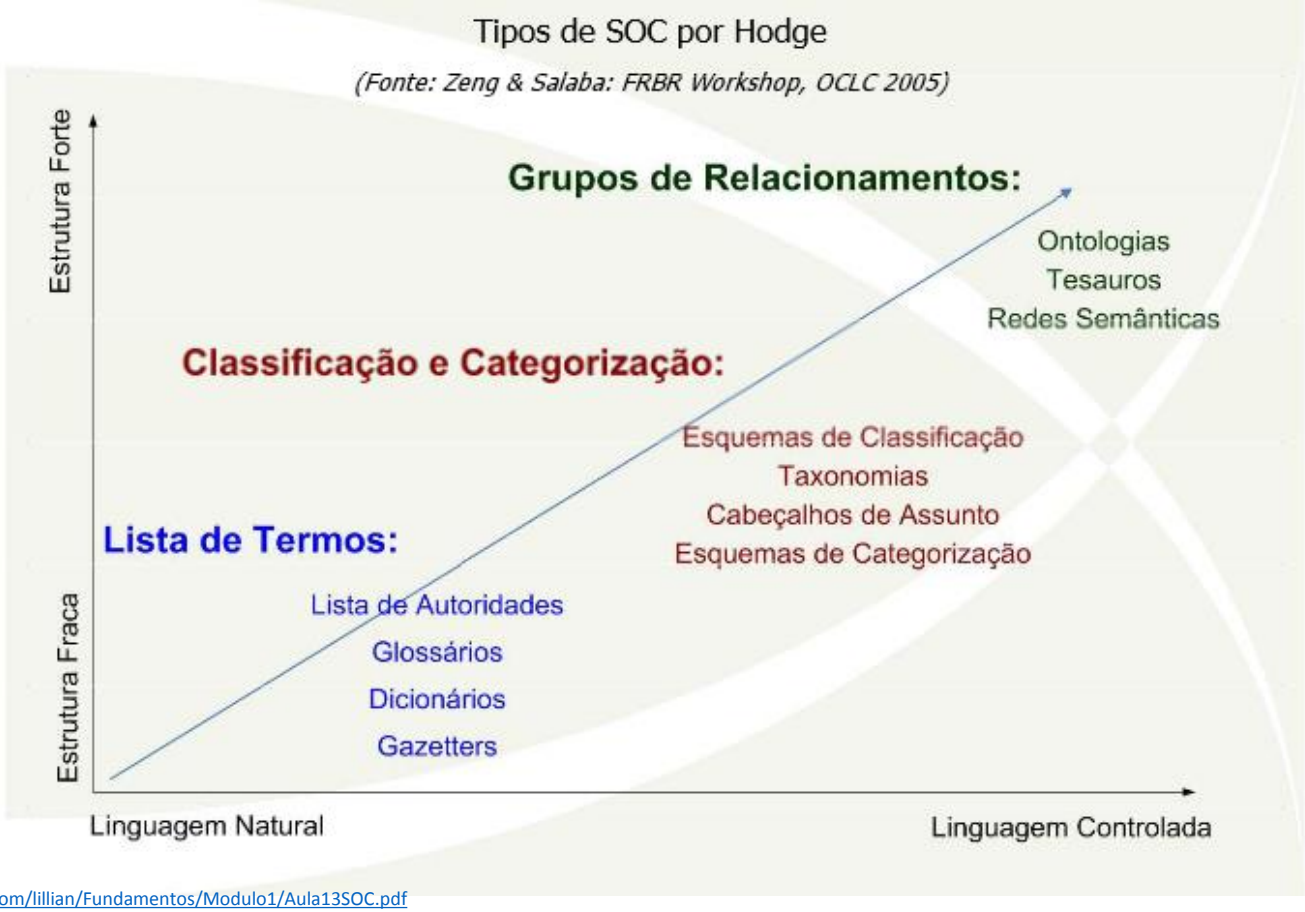
- Criação distribuída e manutenção de informação sobre dados, sobre como é usado, quem usa e como são gerados;
- Padronização de modelos de troca de dados entre instituições e transformação dos dados;
- agregação de dados pelo uso dos padrões RDF/OWL;
- Interpretação dos dados através de semântica precisa e vocabulários controlados, incluindo dados geoespaciais e aspectos temporais;
- Conhecer quem publica o quê e em que formato;
- Garantia de origem e confiabilidade da fonte de dados;
- Correlações e comparações de dados;
- Controle social do processo politico e da produção de políticas públicas;
- Transparência da eficiência e efetividade governamental;
- Participação e conhecimento da sociedade civil sobre as iniciativas governamentais.

Fonte: http://oegov.org/

O que estamos fazendo?

- É uma taxonomia?
- Um vocabulário controlado?
- Um tesauro?
- Um dicionário?
- (....)

Tipos de Sistemas de Organização do Conhecimento



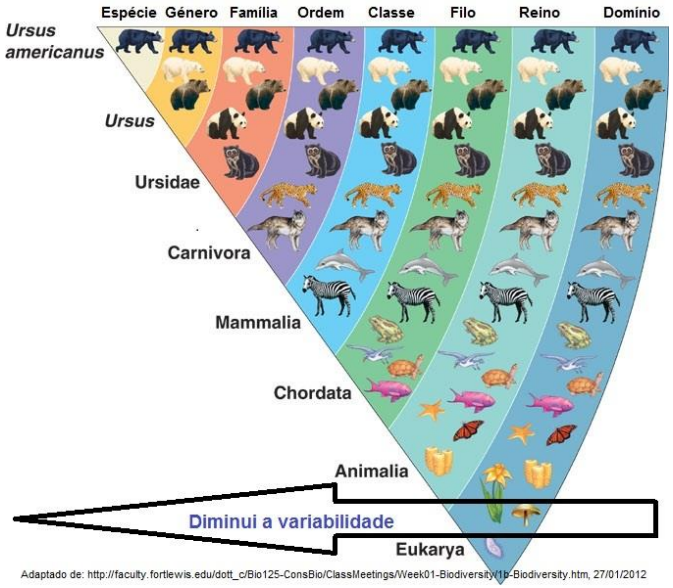
Fonte: http://www.alvarestech.com/lillian/Fundamentos/Modulo1/Aula13SOC.pdf

Listas de termos

- Listas de autoridades:
  - Usada para controlar a variedade entre nomes e entidades;
  - Não inclui uma organização e nem estrutura complexa;
  - A apresentação pode ser alfabética ou um simples esquema de organização
  - Exemplo: nomes de países, nomes de organizações de saúde, etc.
- Glossários:
  - Lista de termos com definições;
  - Os termos podem ser de um assunto específico;
- Dicionários:
  - Lista de palavras em ordem alfabética e suas definições;
  - Escopo é mais geral que os glossários;
  - Podem prover informações sobre a origem das palavras, variações quanto a morfologia e a escrita da palavra e os múltiplos significados nas diferentes áreas.

Classificações

- Classificação, categorização e taxonomia:
  - Organiza em grupos, segundo diferenças ou semelhanças, em certo número de classes metodicamente distribuídos.



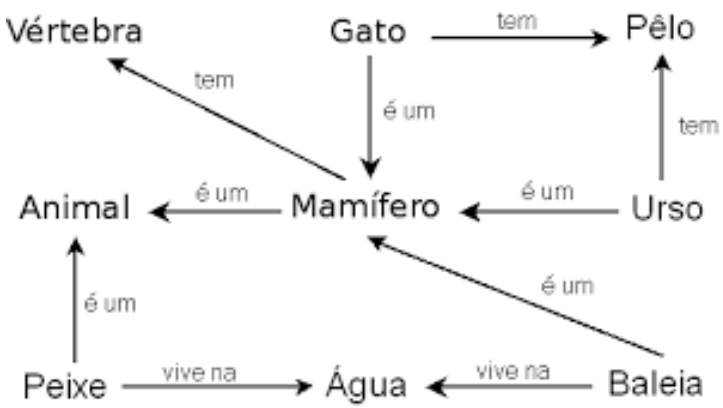
Adaptado de: http://faculty.fortlewis.edu/dott_c/Bio125-ConsBio/ClassMeetings/Week01-Biodiversity/1b-Biodiversity.htm, 27/01/2012

Grupos de relacionamento

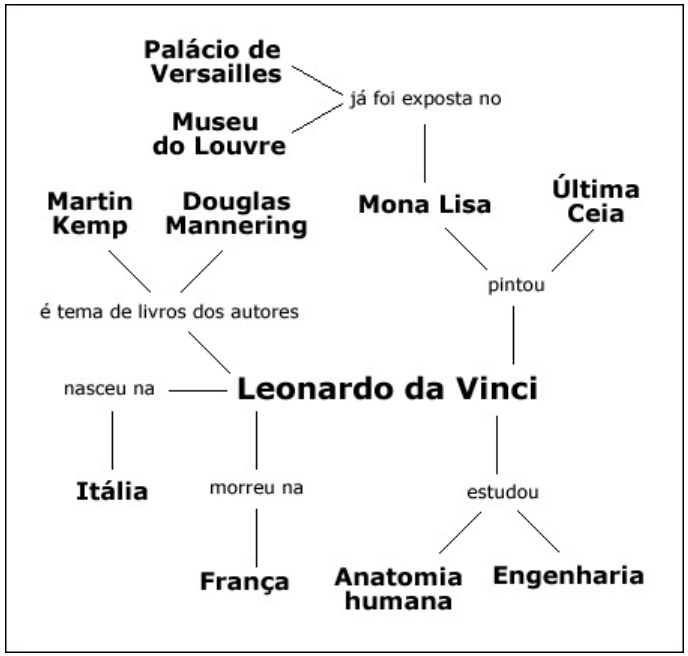
conexão entre termos e conceitos

- Tesauros:
  - Conjunto de termos representando conceitos e as relações hierárquicas, equivalentes e associativas entre eles;
- Redes semânticas:
  - Estrutura de conceitos e termos em **forma de rede**, onde os conceitos são os nós e os relacionamentos expandem-se a partir dos nós.
- Ontologia:
  - Representa **relacionamentos complexos** entre objetos, não cobertos por nenhuma outra forma de organização anterior;
  - Define formalmente as relações entre os termos;
  - É uma especificação formal e explícita de uma conceitualização.

Exemplos de Tesauros, Redes Semânticas e Ontologias



Rede semântica



Ontologia

Quais são os componentes de uma ontologia?

- **Classes**: abstração que representa um conjunto de indivíduos e podem ser organizadas hierarquicamente
- Propriedades de classe: descrevem **atributos** ou **relacionamento** de uma classe
- **Restrições** das propriedades: por exemplo especificam se um propriedade de objeto é inversa ou simétrica
- **Indivíduos**: tem as propriedades definidas pela classe, tem a influência de espaço/tempo

Exemplos dos componentes de uma ontologia

- Classes: **Teatro**
- Atributos: **Data da criação**
- Relacionamento: Teatro **temGestor** Agente
- Restrições: um Teatro **obrigatoriamente** temGestor
- Indivíduos: **Teatro Municipal de Brasilia**

Que problemas resolvem essas soluções de organização?

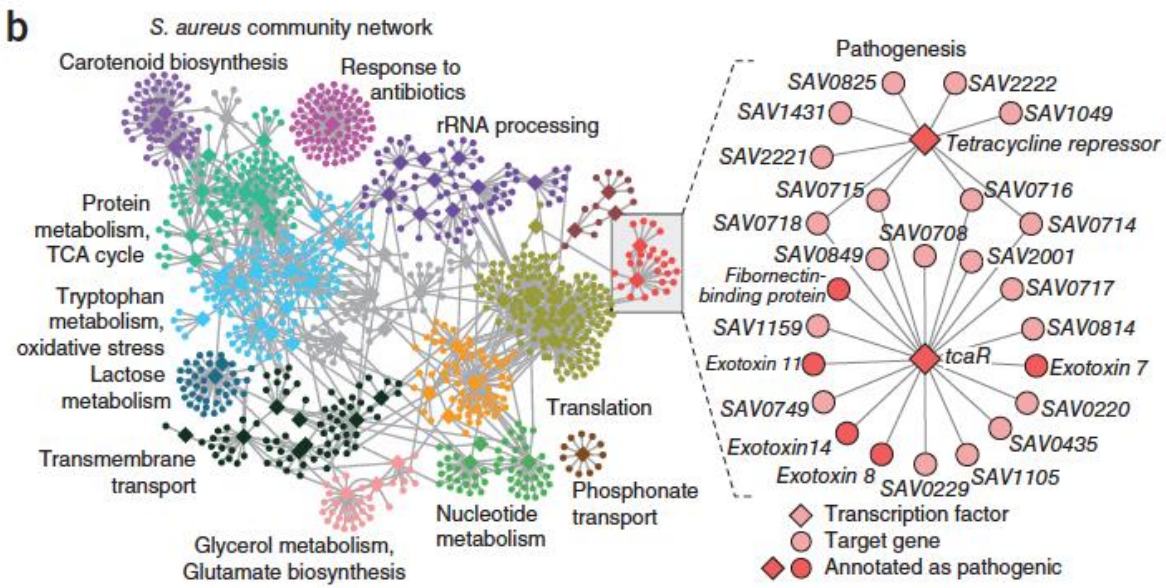
- No geral, são abordagens que apoiam a estruturação, classificação, modelagem e representação de conceitos e relacionamentos.
- Construção de **mapas de conhecimento**:
  - Permite a produção de mapas temáticos, indicadores, análises e relatórios de múltiplas possibilidades de relacionamento da informação
- Navegação para **busca da informação**:
  - Produção de facetas filtragem e navegação;
  - Escolha de termos de busca:
    - Apresentação de termos semanticamente relacionados, facilitando aos usuários selecionarem o que melhor representa o que desejam pesquisar ou mesmo a descoberta de relações inesperadas
  - Busca integrada:
    - Permite agregar campos de busca genéricos e específicos.
- Criação de **novos conhecimentos** a partir do que já existe:
  - Facilita e incentiva a descoberta de novos relacionamentos informacionais antes desconhecidos, amplia o potencial de remixagem da informação.

Exemplos mais concretos...

The Gene Ontology

- Um dos maiores esforços colaborativos na produção de uma ontologia mundial;
- Fornece terminologia para descrição consistente de produtos de pesquisa genética em diferentes bases de dados em termos das associações com processos biológicos, componentes celulares e funções moleculares de diferentes espécies.
- http://geneontology.org/

The Gene Ontology



Observação: os nós em cinza são elementos da ontologia que ainda não foram enriquecidos/comentados pela comunidade.

The Gene Ontology

FlyBase

FB2015_04, released September 3, 2015

Vocabularies Search Page

Home Tools Downloads Links Community Species About Help Archives Jump to Gene Go

Enter a search term

CV Hierarchy: All

Enter text:

Search

Note: Consider using wild cards (*) and/or singular search terms.

Or browse the following hierarchy structures:

- Anatomy (FBbt)
- Biological Process (GO)
- Cellular Component (GO)
- Molecular Function (GO)
- Human Disease (DO) NEW
- Molecular Interaction (MI)
- Development (FBdv)
- Allele Class (FBcv)
- Phenotypic Class (FBcv)
- Origin of mutation (FBcv)
- Stock descriptor (FBsv)
- Publication descriptor (FBcv)
- Imaging method (FBbi)
- Sequence ontology (SO)
- Gene Class (SO)
- Chromosome Structure Variation (SO)
- Sequence Variant (SO)

```
Disease
| disease by infectious agent
| disease of anatomical entity
    | cardiovascular system disease
    | endocrine system disease
    | gastrointestinal system disease
    | immune system disease
    | integumentary system disease
    | musculoskeletal system disease
    | nervous system disease
    | [more ...]
| disease of cellular proliferation
    | benign neoplasm
    | cancer
    | pre-malignant neoplasm
| disease of mental health
| disease of metabolism
| genetic disease
| physical disorder
| syndrome
```

FlyBase

FB2015_04, released September 3, 2015

Term Report

Home Tools Downloads Links Community Species About Help Archives Jump to Gene Go

Help Open All Close All

General Information

Term	cancer	ID (Ontology)	DOID:162 (Human Disease)
Definition	"A disease of cellular proliferation that is malignant and primary, characterized by uncontrolled cellular proliferation, local cell invasion and metastasis.		
Also Known As	"malignant neoplasm" ; "malignant tumor " ; "primary cancer"		
Comment			

Annotations

Records which annotation includes this term

Data Class	Field	Records
Alleles (FBal)	HUMAN_DISEASE_INTERACTIONS	30
Alleles (FBal)	HUMAN_DISEASE_MODELS	35

Records which annotation includes this term *OR* any of its *CHILDREN TERMS*

Genes 134 | Human Diseases 1 | Alleles 194

Results list data from **ALL** species. Please use Query Builder to retrieve species specific data.

Exact full annotation statements including this term, and relevant records

Full annotation statements	Relevant records
	Alleles
cancer	59
ameliorates | cancer	21
exacerbates | cancer	6
model of | cancer	36

Google – Mapa do conhecimento

Google

leonardo da vinci

Web Imagens Vídeos Notícias Mapas Mais Ferramentas de pesquisa

Aproximadamente 71.400.000 resultados (0,76 segundos)

ESCOLA LEONARDO DA VINCI DF - ENSINO ...
www.**leonardodavinci**df.com.br/
CENTRO EDUCACIONAL **LEONARDO DA VINCI** DF COLÉGIO PARTICULAR DE BRASÍLIA DF UM DOS MELHORES ENSINOS FUNDAMENTAL E MÉDIO DE ...
Unidade Taguatinga - Unidade Norte - Unidade Sul - Calendário 2015

Leonardo da Vinci – Wikipédia, a enciclopédia livre
https://pt.wikipedia.org/wiki/**Leonardo_da_Vinci**
Leonardo da Vinci é considerado por vários o maior gênio da história, devido a sua multiplicidade de talentos para ciências e artes, sua engenhosidade e ...
Polímata - Mona Lisa - A Última Ceia - Anchiano

Leonardo da Vinci - Biografia - InfoEscola
www.infoescola.com/biografias/**leonardo-da-vinci**/
Artigo sobre o artista e inventor renascentista italiano **Leonardo da Vinci**, como foi a sua vida, principais obras produzidas, entre outras informações.

Leonardo da Vinci - biografia, obras de Da Vinci, frases
www.suapesquisa.com/**leonardo**/
Introdução. **Leonardo da Vinci**, artista renascentista italiano, nasceu em 15/04/1452. Existem algumas dúvidas sobre a cidade de seu nascimento: para alguns ...

Leonardo da Vinci

Cientista

Leonardo di Ser Piero da Vinci, ou simplesmente Leonardo da Vinci, foi um polímata nascido na atual Itália, uma das figuras mais importantes do Alto Renascimento, que se destacou como cientista, ... Wikipédia

Nascimento: 15 de abril de 1452, Vinci, Itália

Falecimento: 2 de maio de 1519, Amboise, França

Sepultamento: Capela de Saint-Hubert, Amboise, França

Google
Mapa do conhecimento

Leonardo da Vinci

Cientista

Leonardo di Ser Piero da Vinci, ou simplesmente Leonardo da Vinci, foi um polímata nascido na atual Itália, uma das figuras mais importantes do Alto Renascimento, que se destacou como cientista, ... Wikipédia

Nascimento: 15 de abril de 1452, Vinci, Itália

Falecimento: 2 de maio de 1519, Amboise, França

Sepultamento: Capela de Saint-Hubert, Amboise, França

Estrutura: Catedral de Milão

Filiação: Piero da Vinci, Caterina

Obras

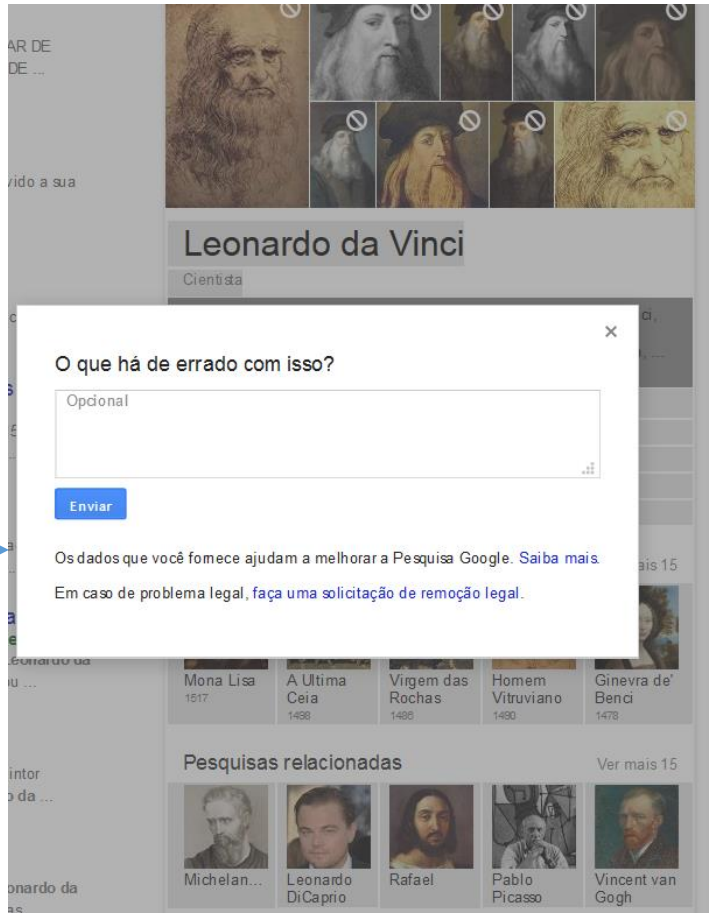
Ver mais 15

Mona Lisa 1517 | A Ultima Ceia 1498 | Virgem das Rochas 1486 | Homem Vitruviano 1490 | Ginevra de' Benci 1478

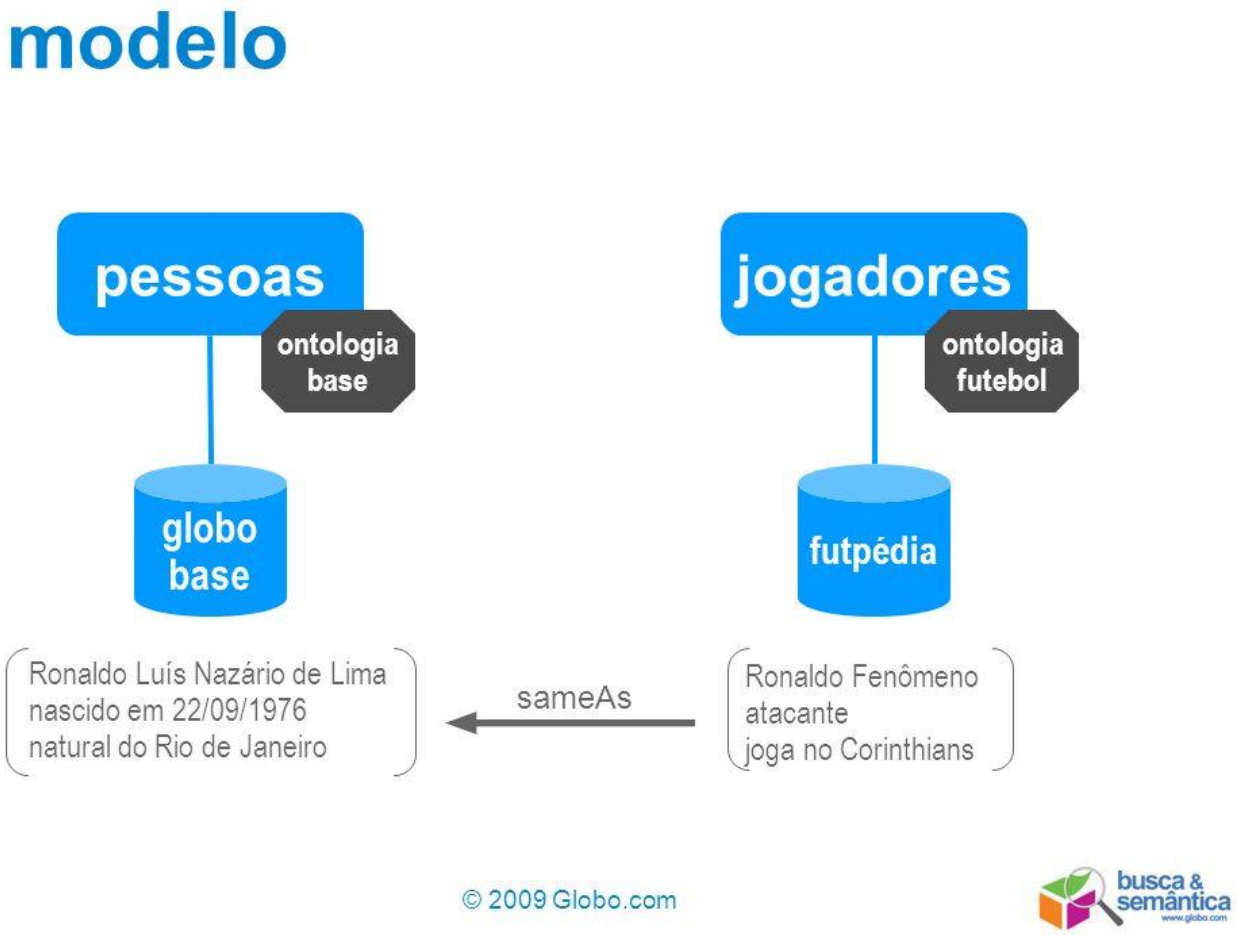
Pesquisas relacionadas

Ver mais 15

Michelan... | Leonardo DiCaprio | Rafael | Pablo Picasso | Vincent van Gogh



Globo.com


modelo
pessoas
ontologia base
globo base
jogadores
ontologia futebol
futpédia
Ronaldo Luís Nazário de Lima
nascido em 22/09/1976
natural do Rio de Janeiro
sameAs
Ronaldo Fenômeno
atacante
joga no Corinthians
© 2009 Globo.com
busca & semântica

Globo.com

solução

- ontologia base
  - modelo central, comum a todo o portal
  - entidades: pessoas, lugares, obras, etc
  - ponto de contato para linked data
- ontologias específicas
  - detalhamento adequado a cada produto
  - modelagem a cargo de cada área
- aplicações comuns
  - widget de edição “aprende” com as ontologias

© 2009 Globo.com

busca & semântica
www.globo.com

Globo.com

globo.com
notícias
esportes
vídeos
globo esporte .com
MESSI
buscar
FOTOS
MESSI
AS ÚLTIMAS
Messi parabeniza Barcelona por título após 'grande ano de trabalho'
'Götze, Müller e Lewandowski podem desbancar Messi', diz técnico alemão
Produtora prepara filme baseado em biografia não autorizada de Messi
DESTAQUES DO TIME
konui

Casos de uso de ontologia no governo

Repositório de ontologia do governo

http://vocab.e.gov.br/

gov.br e-VoG **Repositório de Vocabulários e Ontologias do Governo Eletrônico**

Definição

O e-VoG (Vocabulários e Ontologias do Governo Eletrônico) é um conjunto de padrões, ferramentas e metodologias para possibilitar: o intercâmbio de informações com acordo semântico, de forma a viabilizar o pronto cruzamento de dados de diversas fontes; o uso de metodologias de modelagem conceitual como forma de elicitação do conhecimento tácito das áreas de negócio de governo; o uso de ontologias como ferramenta para explicitar conhecimentos de maneira formal e coerente; o alinhamento conceitual das diversas áreas do conhecimento do governo.

Um dos produtos do e-VoG é o Repositório de Vocabulários e Ontologias de Governo Eletrônico, local para acesso a todas as referências ontológicas do Governo Eletrônico Federal.

Além do repositório persistente de ontologias de referência na ePING, também estão previstos os seguintes produtos para o e-VoG:

- Capacitação em ontologias;
- Conjunto de vocabulários e ontologias de domínio;
- Manual de boas práticas e processo de engenharia de ontologias;
- Política de URIs para Publicação de Dados no Governo.

Lista de vocabulários disponíveis:

Estáveis

EDGV

Autores: Comissão Nacional de Cartografia (CONCAR) do Governo Brasileiro

Descrição: Estruturação de Dados Geoespaciais Vetoriais.

Namespace URI: http://vocab.e.gov.br/edgv/2.1

VCGE 2.0.3

Autores: ePING

Descrição: Vocabulário Controlado de Governo Eletrônico.

Formatos: VCGE em formato JSON, VCGE em formato SKOS (RDF/N3), VCGE em formato CSV, VCGE em formato SQL, VCGE 2.0.3 PDF, Descrição dos Arquivos de Exportação do VCGE PDF, VCGE 2.0.3 Detalhado PDF

Namespace URI: http://vocab.e.gov.br/id/governo

Orçamento Federal Brasileiro

Autor: SOF/MP

Descrição: Modelo ontológico da classificação da despesa do orçamento federal brasileiro.

Formatos: Orçamento federal Brasileiro no formato OWL/XML, jOWL

Namespace URI: http://vocab.e.gov.br/2013/09/loa

Repositório de ontologia do governo

Em desenvolvimento

OPS

Autores: Comunidade no Corais
Descrição: Ontologia de Participação Social.
Formatos: Ontologia de participação Social no formato OWL/XML
Namespace URI: http://vocab.e.gov.br/2013/12/ops

SIAFI

Autores: ePING
Descrição: Sistema Integrado de Administracao Financeira.
Namespace URI: http://vocab.e.gov.br/siafi

SICAF

Autores: ePING
Descrição: Sistema de Cadastrameto Unificado de Fornecedores.
Namespace URI: http://vocab.e.gov.br/sicaf

Estruturas Organizacionais — SIORG

Autores: ePING
Descrição: Sistema de Orgãos.
Namespace URI: http://vocab.e.gov.br/siorg

Versões antigas

VCGE - Versão de março de 2011

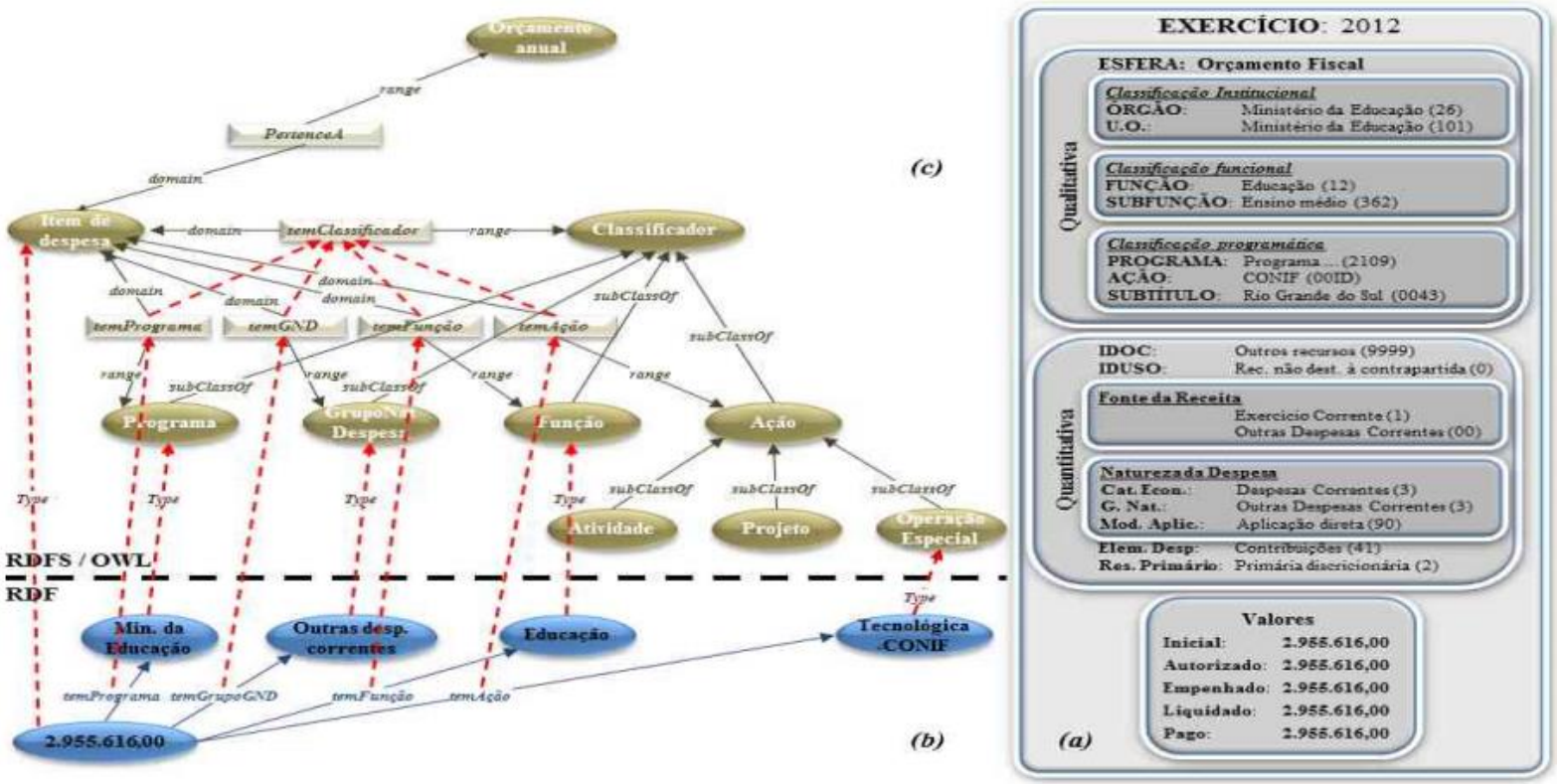
Autores: ePING
Descrição: Vocabulário Controlado de Governo Eletrônico.
Formatos: JSON, SKOS (RDF/N3), SKOS (RDF/XML), PDF
Namespace URI: http://vocab.e.gov.br/2011/03/vcge

Uma Ontologia das Classificações da Despesa do Orçamento Federal

Este artigo apresenta, em OWL, a Ontologia da Classificação da Despesa do Orçamento Federal, que tem como propósito possibilitar à comunidade de **desenvolvedores e técnicos em finanças** públicas, o **acesso completo, irrestrito e automático aos dados do orçamento federal** brasileiro.

http://ceur-ws.org/Vol-938/ontobras-most2012_paper30.pdf

www.siop.planejamento.gov.br



Dados abertos

Manual de Referência

Carta de Serviços

Ontologia do Orçamento Federal Brasileiro em OWL

PLOA 2015	LOA 2011	LOA 2007	LOA 2003
LOA 2014	LOA 2010	LOA 2006	LOA 2002
LOA 2013	LOA 2009	LOA 2005	LOA 2001
LOA 2012	LOA 2008	LOA 2004	LOA 2000

Atendimento SIOP

Dias úteis, das 8h às 18h

(61) 2020.2121

suportesiop@planejamento.gov.br

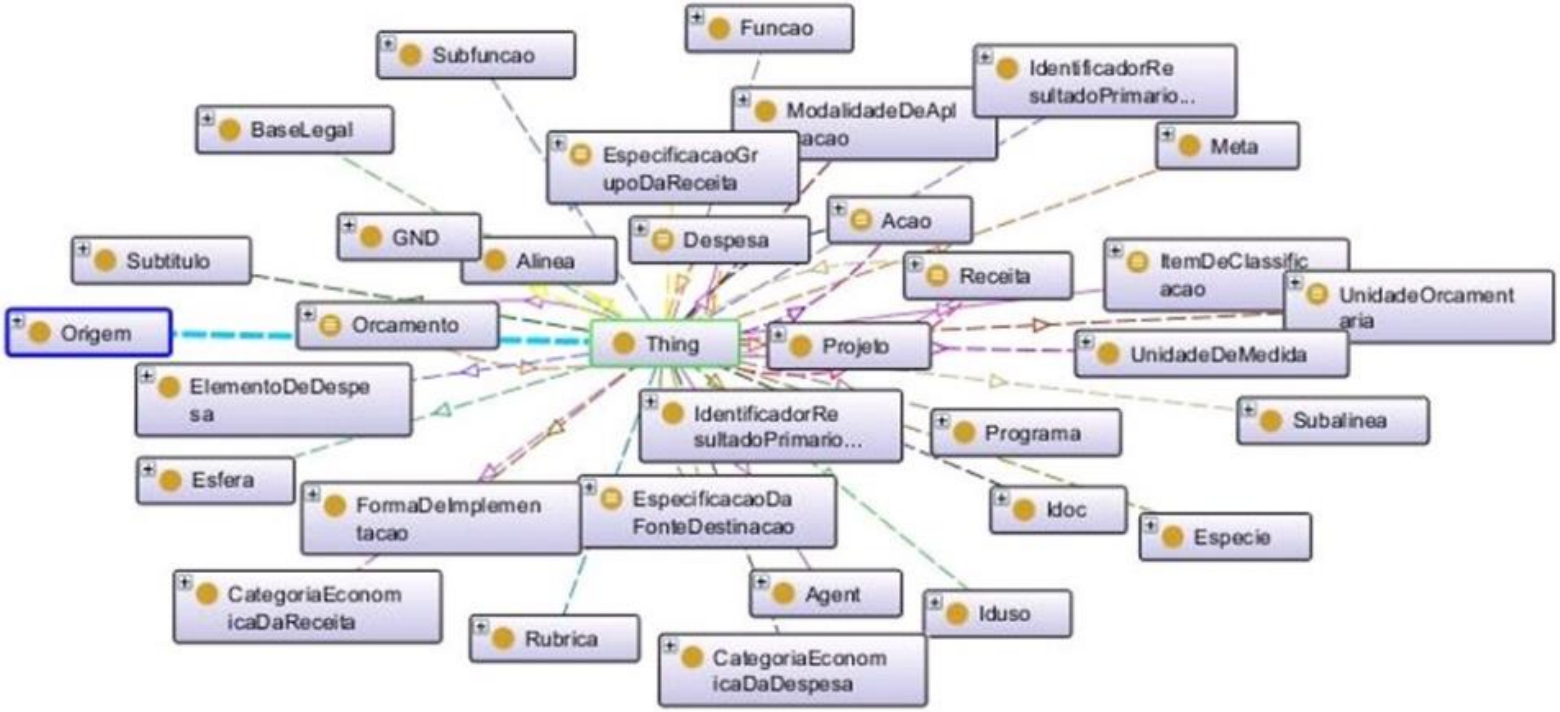
vocab.e.gov.br/2013/09/loa.owl


```
    ///////////////////////////////////////////////////////////////////////////////////////
    //
    // Classes
    //
    ///////////////////////////////////////////////////////////////////////////////////////

  -->
  <!--  http://vocab.e.gov.br/2013/09/loa#Acao  -->
▼<owl:Class rdf:about="http://vocab.e.gov.br/2013/09/loa#Acao">
   <rdfs:subClassOf rdf:resource="http://vocab.e.gov.br/2013/09/loa#Classificador"/>
  </owl:Class>
  <!--  http://vocab.e.gov.br/2013/09/loa#Atividade  -->
▼<owl:Class rdf:about="http://vocab.e.gov.br/2013/09/loa#Atividade">
   <rdfs:subClassOf rdf:resource="http://vocab.e.gov.br/2013/09/loa#Acao"/>
  </owl:Class>
▼<!--
    http://vocab.e.gov.br/2013/09/loa#CategoriaEconomica
  -->
▼<owl:Class rdf:about="http://vocab.e.gov.br/2013/09/loa#CategoriaEconomica">
   <rdfs:subClassOf rdf:resource="http://vocab.e.gov.br/2013/09/loa#Classificador"/>
  </owl:Class>
  <!--  http://vocab.e.gov.br/2013/09/loa#Classificador  -->
▼<owl:Class rdf:about="http://vocab.e.gov.br/2013/09/loa#Classificador">
  ▼<rdfs:subClassOf>
   ▼<owl:Class>
    ▼<owl:intersectionOf rdf:parseType="Collection">
     ▼<owl:Restriction>
        <owl:onProperty rdf:resource="http://vocab.e.gov.br/2013/09/loa#codigo"/>
        <owl:someValuesFrom rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
      </owl:Restriction>
     ▼<owl:Restriction>
        <owl:onProperty rdf:resource="http://www.w3.org/1999/02/22-rdf-syntax-ns#label"/>
        <owl:someValuesFrom rdf:resource="http://www.w3.org/2001/XMLSchema#string"/>
      </owl:Restriction>
    </owl:intersectionOf>
   </owl:Class>
  </rdfs:subClassOf>
  </owl:Class>
  <!--  http://vocab.e.gov.br/2013/09/loa#ElementoDespesa  -->
▼<owl:Class rdf:about="http://vocab.e.gov.br/2013/09/loa#ElementoDespesa">
   <rdfs:subClassOf rdf:resource="http://vocab.e.gov.br/2013/09/loa#Classificador"/>
  </owl:Class>
  <!--  http://vocab.e.gov.br/2013/09/loa#Esfera  -->
▼<owl:Class rdf:about="http://vocab.e.gov.br/2013/09/loa#Esfera">
   <rdfs:subClassOf rdf:resource="http://vocab.e.gov.br/2013/09/loa#Classificador"/>
  </owl:Class>
```

Proposta de uma Ontologia para o Orçamento Público Federal Brasileiro

http://gpnti.marilia.unesp.br/dti/2013/apresentacoes/GiseledaSilvaCraveiro.pdf



Vocabulário Controlado do Governo Eletrônico

O VCGE é um esquema para ser utilizado no elemento assunto.categoria (subject.category) do Padrão de Metadados do Governo Eletrônico (e-PMG)[1]. Termos retirados do VCGE tornarão mais direta, para os gerentes de sítios e portais governamentais, a apresentação dos serviços disponibilizados em uma estrutura de diretório baseada nos indexadores do VCGE.

O VCGE ajuda os cidadãos a encontrar informações mesmo sem o conhecimento de qual órgão o assunto é responsabilidade.

O VCGE está organizado em uma poli-hierarquia. Isto é, um dado termo pode ter um ou mais de um Termos Gerais (TG).

- VCGE
  - Administração
  - Agropecuária
  - Comércio e Serviços
  - Comunicações
  - Cultura
  - Defesa Nacional
  - Economia e Finanças
  - Educação
  - Energia
  - Esporte e Lazer
  - Habitação
  - Indústria
  - Meio ambiente
  - Pesquisa e Desenvolvimento
  - Previdência Social
  - Proteção Social
  - Relações Internacionais
  - Saneamento
  - Saúde
  - Segurança e Ordem Pública
  - Trabalho
  - Transportes
  - Urbanismo

Informações | Exemplos | Sugestões

Termo Autorizado (TA):

VCGE

URI:

http://vocab.e.gov.br/id/governo#esquema

Termos Específicos (TE):

Administração
Agropecuária
Comércio e Serviços
Comunicações
Cultura
Defesa Nacional
Economia e Finanças
Educação
Energia
Esporte e Lazer
Habitação
Indústria
Meio ambiente

Como faremos isso?

- Etapas da metodologia de construção de ontologias:
  1. Definir escopo e domínio da ontologia;
     - Que domínio a ontologia deve atender?
     - Para quê vamos usar a ontologia?
     - Que tipos de questões informacionais a ontologia deve responder?
     - Quem vai usar e manter a ontologia?
  2. Avaliar ontologias existentes que podem ser reutilizadas;
  3. Identificar e escolher os termos da ontologia;
  4. Definir as classes e hierarquia de classes;
  5. Definir as propriedades das classes;
  6. Definir as facetas das propriedades das classes: possibilidades de respostas das propriedades;
  7. Criar instâncias

Como faremos isso?

- Até dezembro de 2016;
- Encontros mensais para realização das 07 etapas de trabalho;
- Etapas virtuais de trabalho:
  - Possibilidade de ampliar a inclusão de mais pessoas;
  - Maior dinâmica a discussões e deliberações durante todo o processo.