

RNP

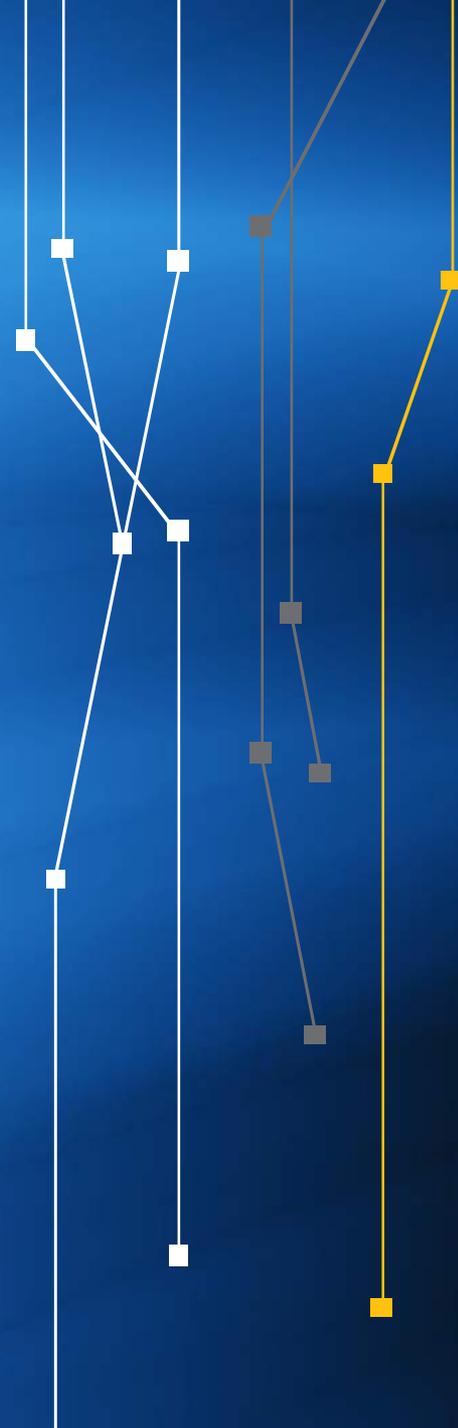
PROMOVENDO O USO INOVADOR DE REDES AVANÇADAS

Alexandre P. Uchoa

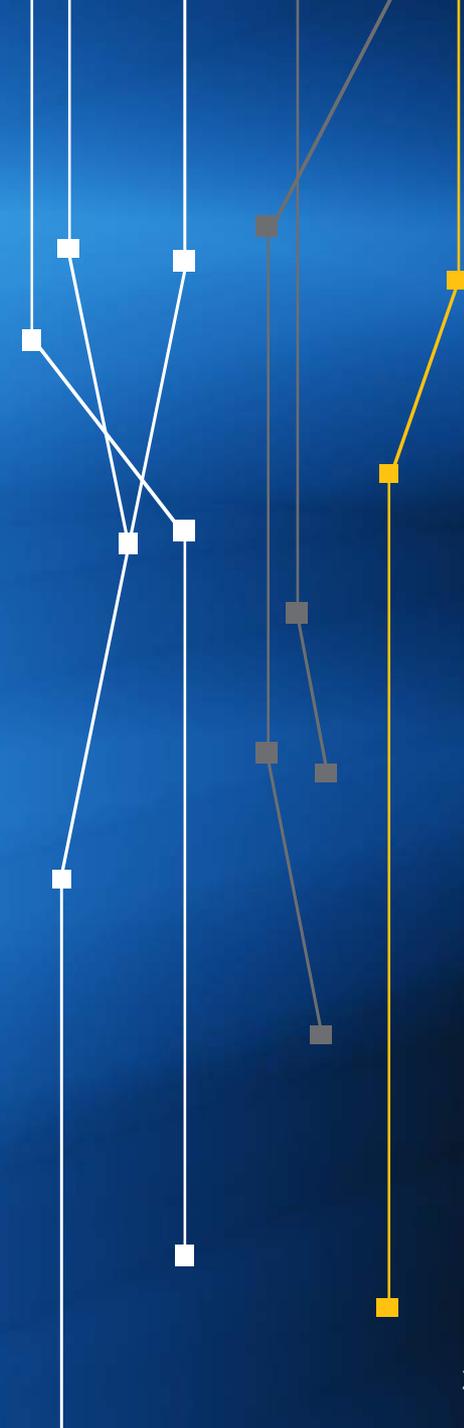
RNP – Rede Nacional de Pesquisa, RJ

+55 21 98153-1143

alexandre.uchoa@rnp.br



Uso da Web Semântica para o Intercâmbio de Dados da Pós-Graduação



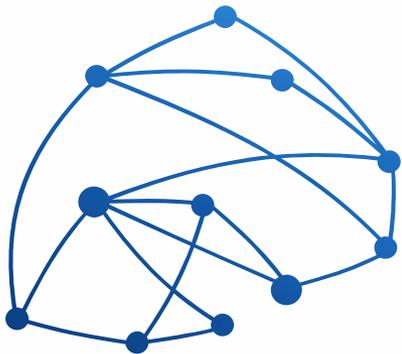
Uso da Web Semântica para a Pós-Graduação

Objetivo

Criar e por em prática uma estratégia de uso da Web Semântica visando facilitar o intercâmbio e a publicação de dados entre organizações, serviços e o Linked Open Data.

Esta estratégia compreende o desenvolvimento de modelos e vocabulários capazes de representar as propriedades e relações dos principais objetos e agentes da pós-graduação nacional.

Tecnologias Semânticas Vs. Web Semântica



RNP

Tecnologias Semânticas

- **Processamento de linguagem natural:** processar o conteúdo de texto não estruturado e buscar extrair e reconhecer os nomes, datas, organizações, eventos, etc., que são mencionados dentro do texto.
- **Mineração de dados:** empregam algoritmos de correspondência de padrões para desvendar tendências e correlações em grandes conjuntos de dados.
- **Inteligência artificial:** utilizam modelos de raciocínio para responder questões complexas. Geralmente incluem algoritmos de aprendizado de máquina.
- **Classificação:** usam heurísticas e regras para agrupar e categorizar dados para apoio à análise de informações.
- **Pesquisa ou busca semântica:** permitem a busca de informações por meio de conceito em vez de palavras-chave ou frases-chave. Pesquisas semântica buscam contextualizar as perguntas para ser capaz de distinguir, por exemplo, o aeroporto John F. Kennedy do presidente.

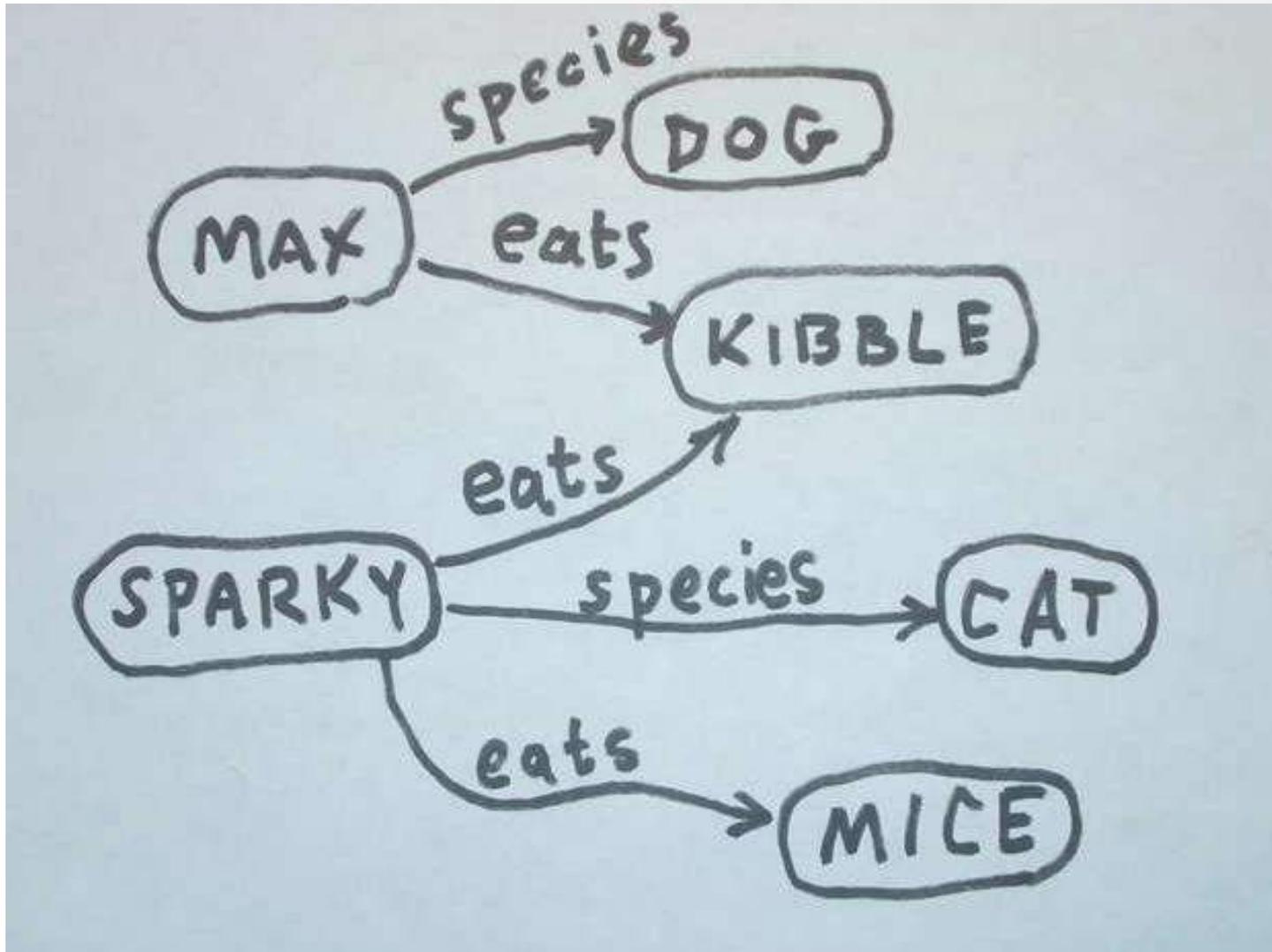
Tecnologias de Web (ou Rede) Semântica

Uma família de **padrões tecnológicos** muito específicos do World Wide Web Consortium (W3C) projetados para descrever e relacionar dados na Web e dentro de empresas.

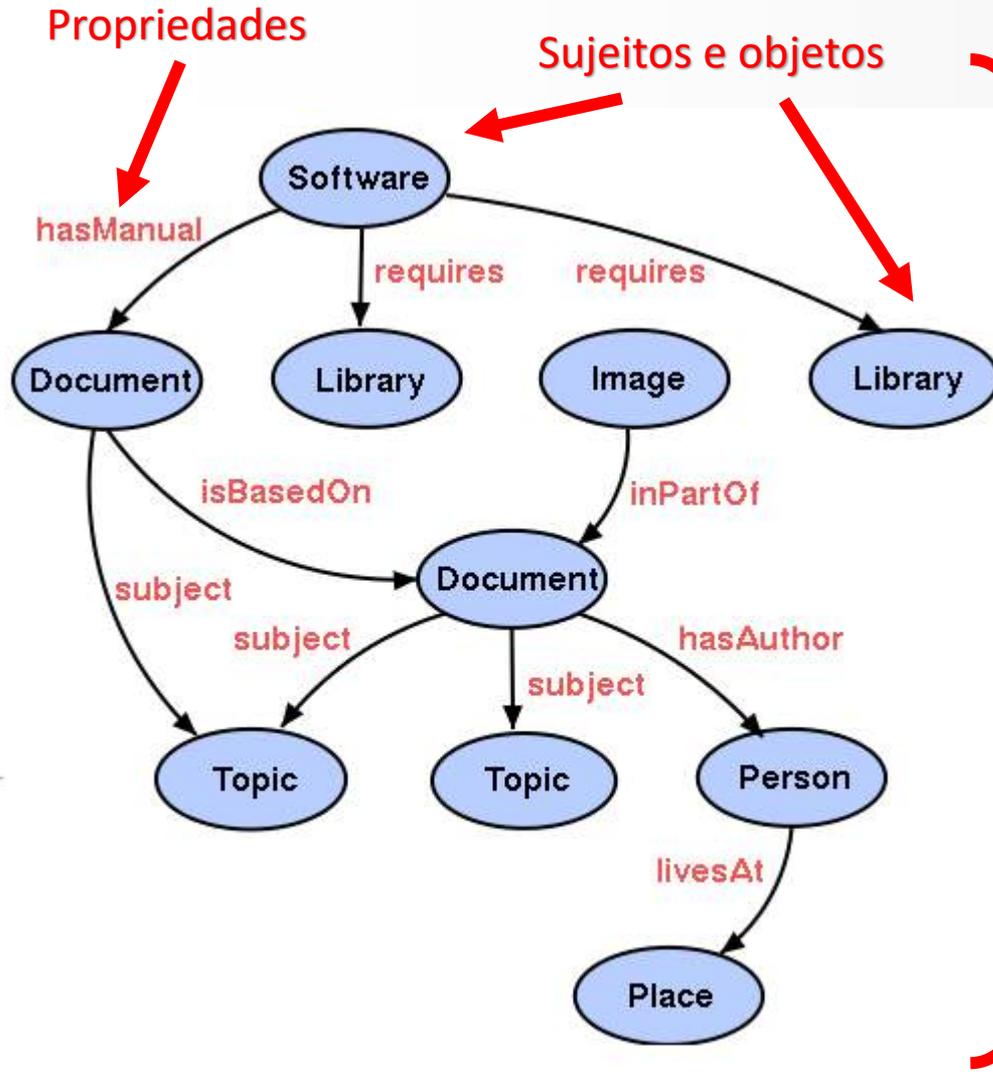
Esses padrões incluem, entre outros:

- **RDF** – linguagem de marcação de dados
- **RDFS, OWL e SHACL** - linguagens para descrever conceitos e seus relacionamentos
- **SPARQL** - linguagem de buscas e consultas na Web Semântica
- **RIF** - linguagem de regras e inferências
- **RDFa** - linguagem para marcar dados dentro de páginas da Web

Representação Semântica do Conhecimento



Representação Semântica do Conhecimento



Modelo Conceitual

Representação Semântica na Web

Vocabulário

Sujeito	Propriedade	Objeto
Software-I	hasManual	Document-1
Software-I	requires	Library-X
Document-1	isBasedOn	Document-13
Document-1	subject	Topic-A
Document-13	subject	Topic-B
Document-13	hasAuthor	John
John	livesAt	SQS-203

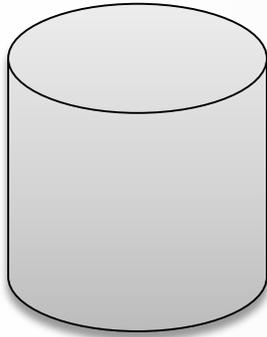
URI

Software-I = <http://example.com/swlib/sw#I>

Tripla	Instancias de classes definidas	Propriedade	Instancias de classes definidas
1	http://example.com/swlib/sw#I	http://capes.gov.br/ws/scheme#:hasManual	http://example.com/swlib/documents#1
2	http://example.com/swlib/sw#I	http://capes.gov.br/ws/scheme#:requires	http://example.com/swlib/libraries#X
3	http://example.com/swlib/documents#1	http://capes.gov.br/ws/scheme#:isBasedOn	http://example.com/swlib/documents#13
4	http://example.com/swlib/documents#1	http://capes.gov.br/ws/scheme#:hasSubject	http://example.com/swlib/topics#A
5	http://example.com/swlib/documents#13	http://capes.gov.br/ws/scheme#:hasSubject	http://example.com/swlib/topics#B
6	http://example.com/swlib/documents#13	http://capes.gov.br/ws/scheme#:hasAuthor	http://example.com/swlib/persons#john
7	http://example.com/swlib/persons#john	http://capes.gov.br/ws/scheme#:livesAt	http://example.com/swlib/locations#sqs203

Dados + Marcações

Semântica dos dados



Dados são acompanhados de anotações que apontam para a fonte de sua descrição semântica

Recursos semânticos descrevem sentido do dado e os relacionam a descrições comuns.



Humano ou Máquina

Torna-os capazes de 'compreender' os dados e de associá-los a outros conteúdos na Web

Nível MÍNIMO de abertura dos dados

★ Disponibilizar seus dados na Web (em qualquer formato) [exemplo ...](#)

Temperature forecast for Galway, Ireland	
Day	Lowest Temperature (°C)
Saturday, 13 November 2010	2
Sunday, 14 November 2010	4
Monday, 15 November 2010	7



Nível MÁXIMO de abertura dos dados

★★★★★ associar dados a outros dados para fornecer contexto [example ...](#)



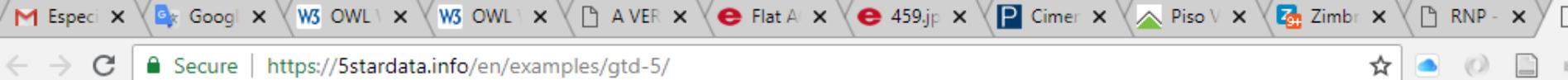
Show embedded data

TEMPERATURE FORECAST FOR GALWAY, IRELAND

Day	Lowest <u>Temperature</u> (°C)
Saturday, 13 November 2010	2
Sunday, 14 November 2010	4
Monday, 15 November 2010	7

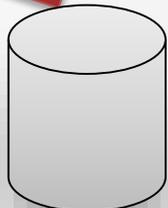
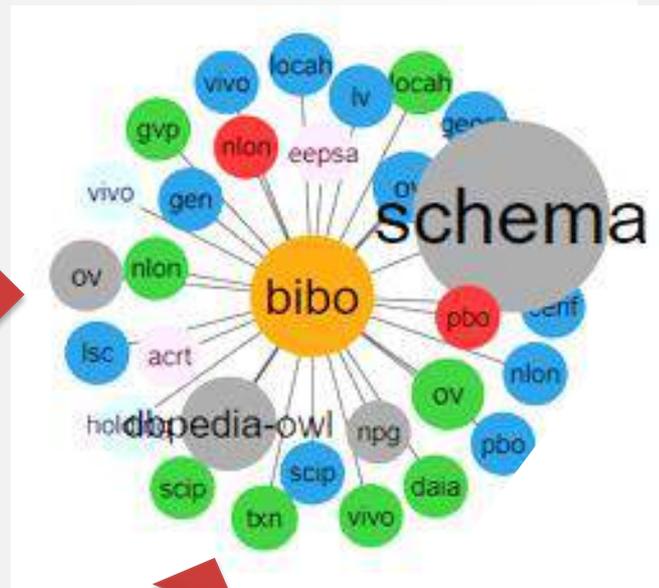
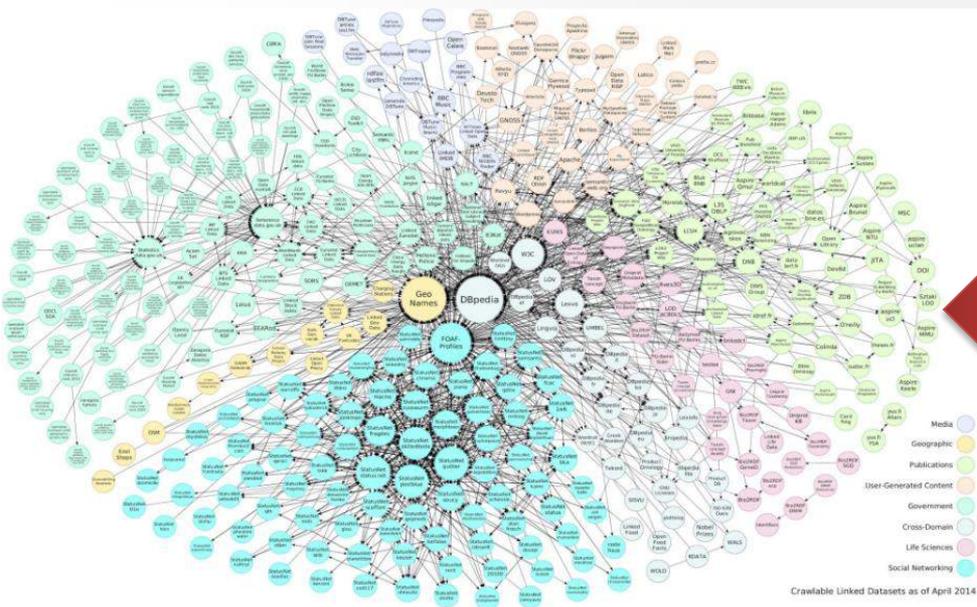
Created: 2012-01-22 by [Michael](#) | Last updated: 2015-08-31 by [James](#) | Code available via [Git-Hub](#)
Unless noted, content on this site is freely available under the [CC0 Public Domain Dedication](#).

Nível MÁXIMO de abertura dos dados



```
<div>
  <button id="show-data">Show embedded data</button>
  <div id="hint" style="display: block;">
    <p>
      To view the embedded data encoded in <a href="http://rdfa.info/">RDFa</a>, hover over the red-dotted a
      You can also <a href="http://graphite.ecs.soton.ac.uk/browser/?uri=http://5stardata.info/en/examples/g
    </p>
    <p>
      In this 5 star example, we used the <a href="http://purl.org/ns/meteo#">Meteo</a> vocabulary to repres
      Each forecast has a URI (for example, '<a href="#forecast20101115">forecast for Monday, 15 November 20
      Additionally, through linking the place (<code>#Galway</code>) to the DBpedia entry <a href="http://db
      we have disambiguated it. Eventually we provided context for the interpretation of the <code>#temp</co
    </p>
  </div>
</div>
<h1 property="dcterms:title">Temperature forecast for Galway, Ireland</h1>
<div id="data" about="#Galway" typeof="meteo:Place" class="highlight">
  <span rel="owl:sameAs" resource="http://dbpedia.org/resource/Galway" class="highlight"></span>
  <table border="1px">
    <tbody><tr>
      <th>Day</th>
      <th>
        <div about="#temp" class="highlight">
          Lowest
          <a rel="rdfs:seeAlso" href="http://en.wikipedia.org/wiki/Temperature" resource="http://dbpedia
          (<span rel="owl:sameAs" resource="http://dbpedia.org/resource/Celsius" class="highlight">°C</s
        </div>
      </th>
    </tr>
    <tr rel="meteo:forecast" resource="#forecast20101113" class="highlight">
      <td>
        <div about="#forecast20101113" class="highlight">
          <span property="meteo:predicted" content="2010-11-13T00:00:00Z" datatype="xsd:dateTime" style=
        </div>
      </td>
    </tr>
  </tbody>
</table>
</div>
```

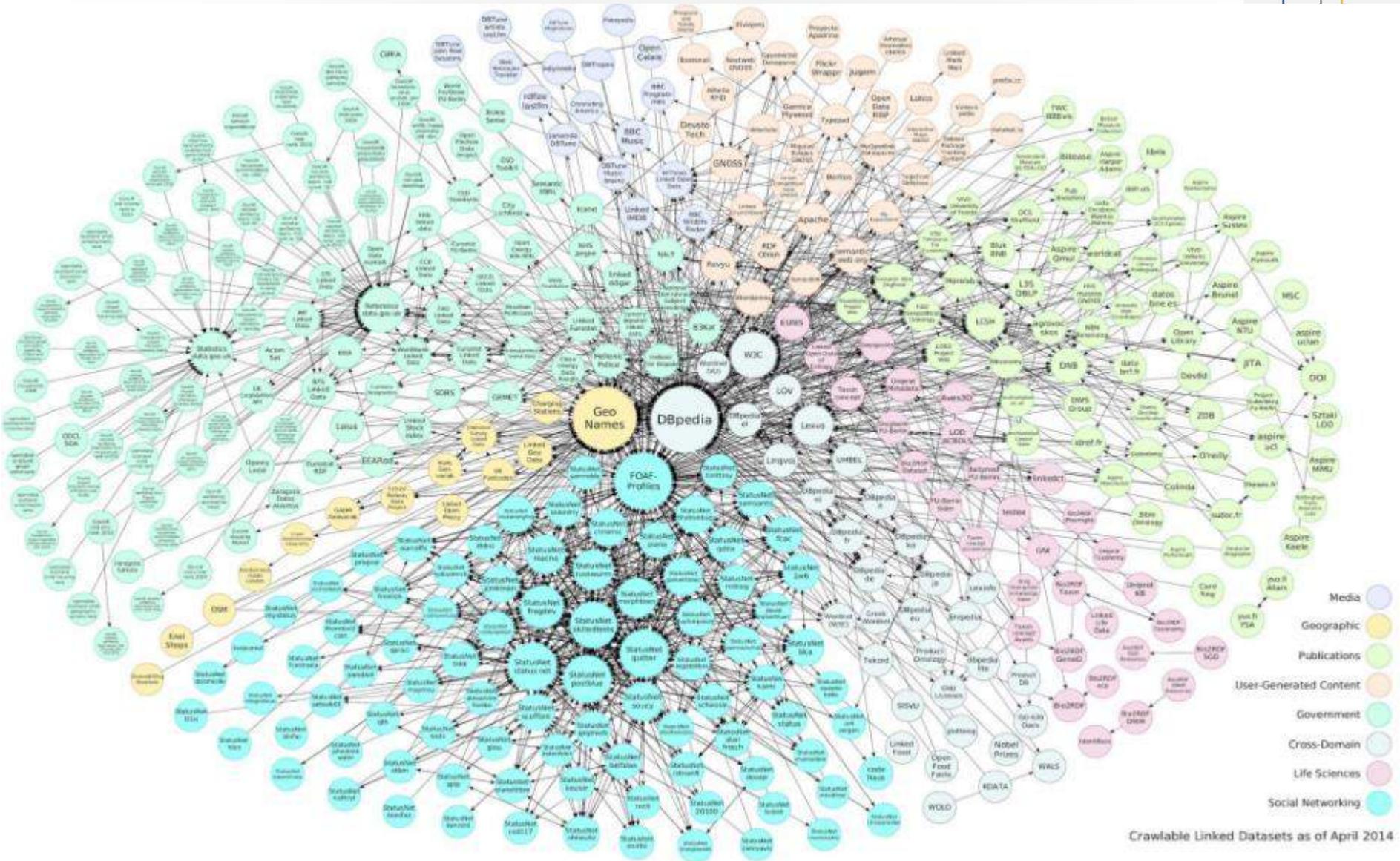
Dados conectados ao mundo



```
graph TD
    subgraph Schema
        bibo((bibo))
        schema((schema))
        vivo1((vivo))
        locah1((locah))
        gen1((gen))
        nlon1((nlon))
        eepsa((eepsa))
        pba1((pba))
        nlon2((nlon))
        pbo1((pbo))
        daia((daia))
        vivo2((vivo))
        scip1((scip))
        bxn((bxn))
        npg((npg))
        dbpedia_owl((dbpedia-owl))
        isc((isc))
        acrt((acrt))
        hold((hold))
    end
    bibo --- schema
    bibo --- vivo1
    bibo --- locah1
    bibo --- gen1
    bibo --- nlon1
    bibo --- eepsa
    bibo --- pba1
    bibo --- nlon2
    bibo --- pbo1
    bibo --- daia
    bibo --- vivo2
    bibo --- scip1
    bibo --- bxn
    bibo --- npg
    bibo --- dbpedia_owl
    bibo --- isc
    bibo --- acrt
    bibo --- hold
```

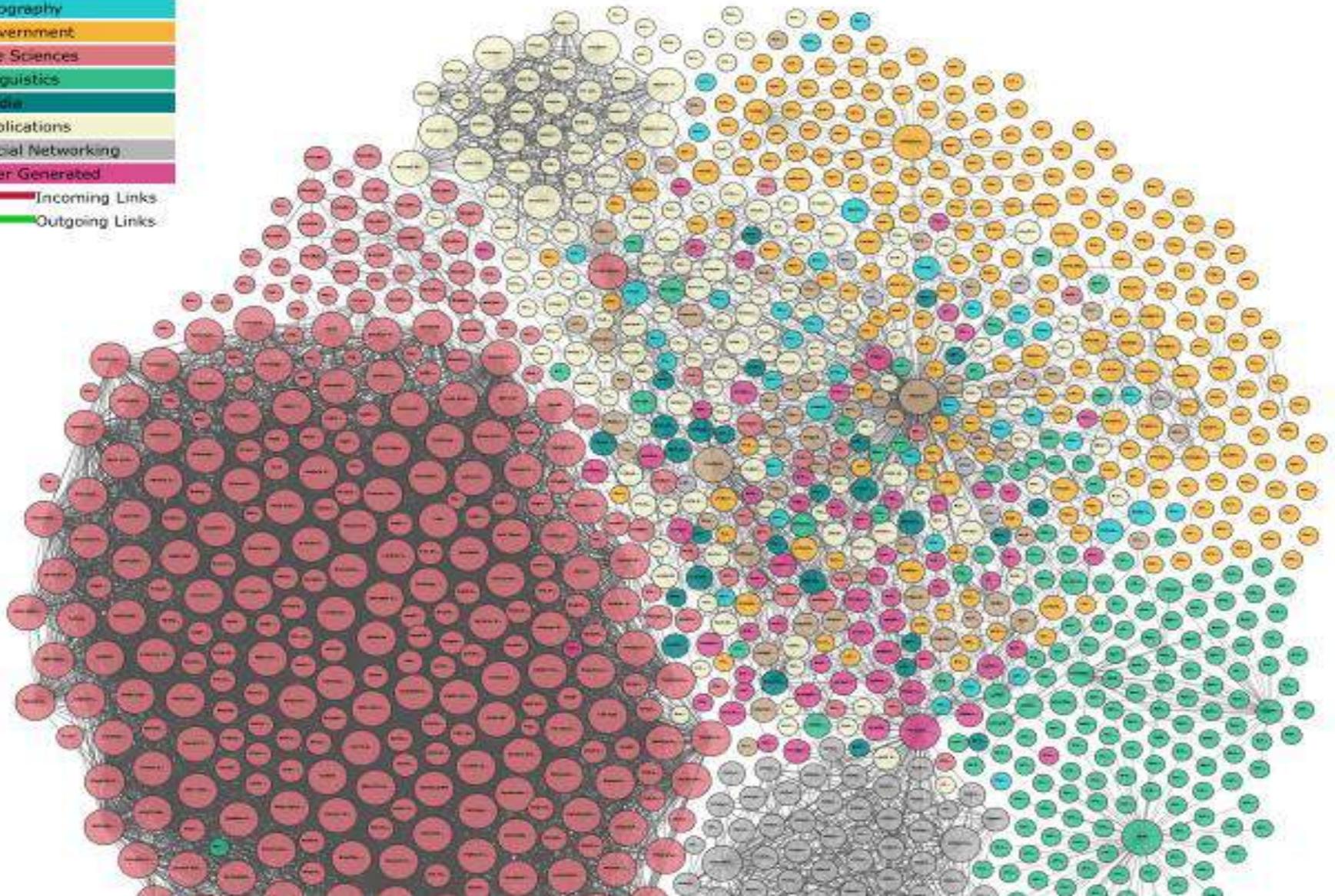


Nuvem da Linked Open Data, 2014

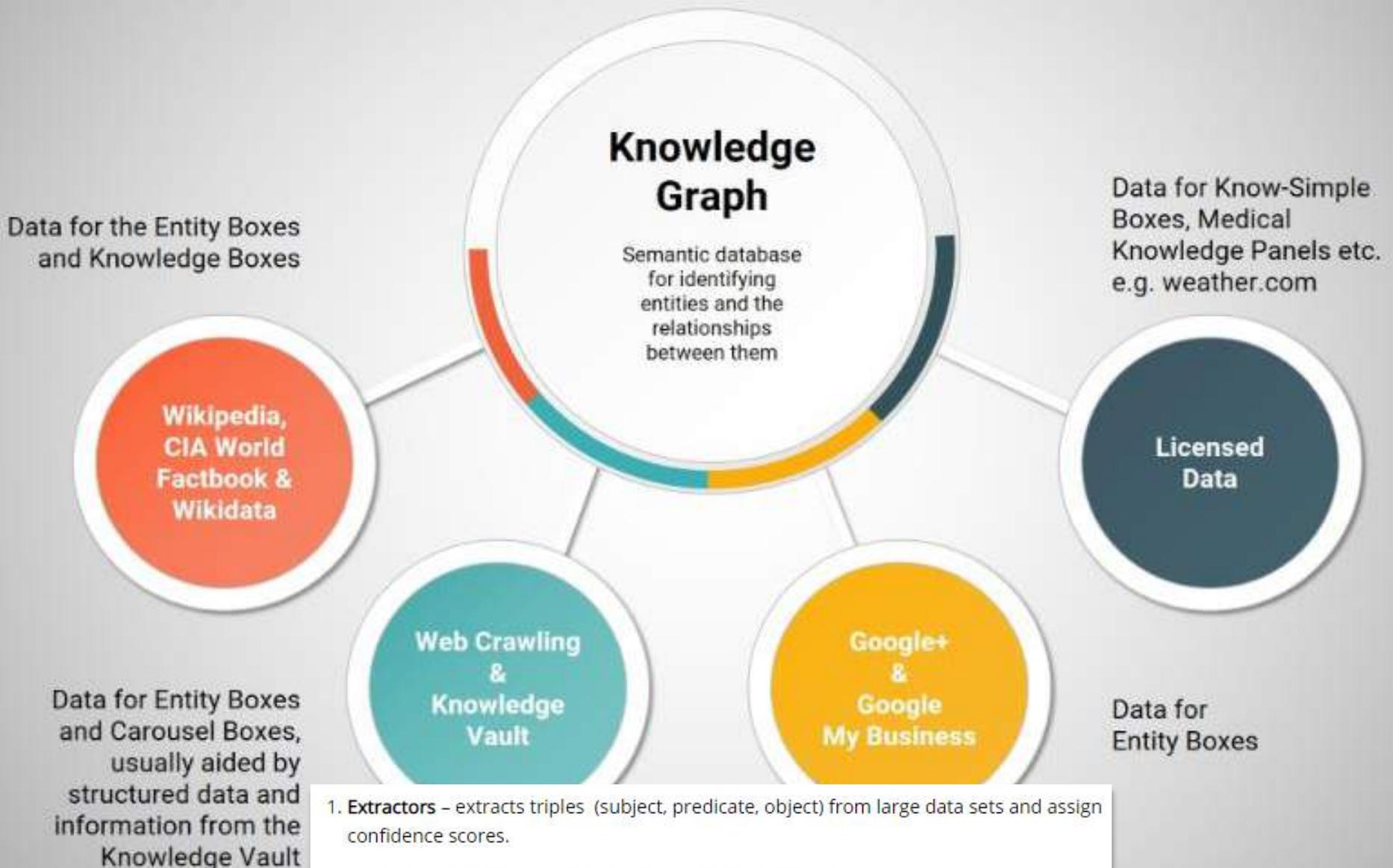


Nuvem da Linked Open Data, 2017

Legend



Data Sources for Google's Knowledge Graph



1. **Extractors** – extracts triples (subject, predicate, object) from large data sets and assign confidence scores.
2. **Graph-based priors** – learning from existing knowledge bases
3. **Knowledge fusion** – final score on factual probability

Tecnologia de Web Semântica em ação

Google  

Todas Notícias Imagens Maps Livros Mais Configurações Ferramentas

Aproximadamente 46.700.000 resultados (0,45 segundos)

CAPES - Coordenação de Aperfeiçoamento de Pessoal de Nível ...

www.capes.gov.br/

CAPES participa de seminário online sobre Plano Nacional de Educação. Governo Aberto ... Acesse a ferramenta de dados estatísticos da CAPES ...

Você já visitou esta página várias vezes. Última visita: 02/07/17

Periódicos CAPES

Portal .periodicos. CAPES oferece acesso aos textos completos de ...

Bolsas e auxílios internacionais

Modalidades de bolsas - Países - Formas de candidatura - ...

Editais Abertos

Editais Abertos. Editais Abertos. 1. Programa CAPES-YALE de ...

Plataforma Freire

Esqueceu a senha? Clique aqui para recuperá-la. Solicitar ...

Bolsas/Estudantes

Avaliação · Bolsas e Auxílios Internacionais · Bolsas ...

Notícias

Abruem. Diretor da CAPES participa de posse de novos ...

[Mais resultados de capes.gov.br »](#)

Coordenação de Aperfeiçoamento de Pessoal de Nível Superior

Agência do governo



Coordenação de Aperfeiçoamento de Pessoal de Nível Superior é uma fundação vinculada ao Ministério da Educação do Brasil que atua na expansão e consolidação da pós-graduação stricto sensu em todos os estados do país. [Wikipédia](#)

Fundação: 11 de julho de 1951, Brasil

Natureza jurídica: Fundação

Criação: 11 de julho de 1951 (67 anos)

Itens também pesquisados

Ver mais 10



Conselho Nacional de Desen...



Ministério da Educação



Fundação de Amparo à Pesquis...



Qualis



Universid... Federal do Rio Gran...

Principais notícias



Capes: Universidades brasileiras vão escolher

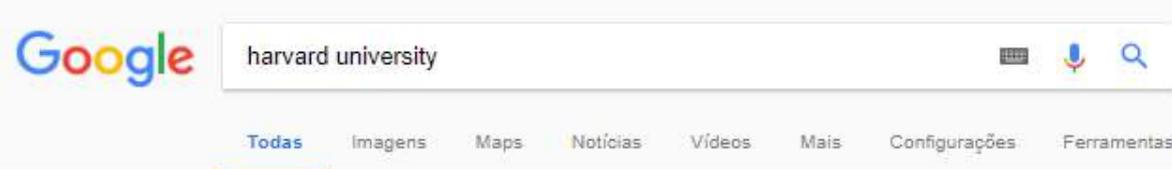


Pagamento de bolsas da Capes não será



Estudantes se mobilizam contra cortes

Tecnologia de Web Semântica em ação



Aproximadamente 225.000.000 resultados (0,85 segundos)

Harvard University

<https://www.harvard.edu/> Traduzir esta página

34 min atrás - Harvard University is devoted to excellence in teaching, learning, and research, and to developing leaders in many disciplines who make a ...

Resultados de harvard.edu

Students

The following links provide University-wide resources for all ...

Admissions

How Aid Works - Visit - Contact Us - Choosing Harvard - Apply - ...

Events

Bees! Permanent Installation, Harvard Museum of Natural ...

Staff

Harvard staff resources are available from Harvard ...

Universidade Harvard – Wikipédia, a enciclopédia livre

https://pt.wikipedia.org/wiki/Universidade_Harvard Traduzir esta página

Universidade de Harvard (em inglês: Harvard University) é uma universidade privada membro da Ivy League, localizada em Cambridge, Massachusetts, ...

Localização: Cambridge, Massachusetts Fundação: 8 de setembro de 1636 (381 anos)

Cores da escola: Carmesim Presidente: Drew Gilpin Faust

História · Campi · Cambridge · Allston

Harvard University - Wikipedia

https://en.wikipedia.org/wiki/Harvard_University Traduzir esta página

Harvard University is a private Ivy League research university in Cambridge, Massachusetts. Established in 1636 and named for its first benefactor clergyman ...

Location: Cambridge, Massachusetts, United ... Students: 22,000

Motto in English: Truth Former names: Harvard College

Harvard University - Página inicial | Facebook

<https://pt-br.facebook.com> Faculdade e universidade Traduzir esta página

Harvard University, Cambridge (Massachusetts). 5.260.870 curtidas · 73.298 falando sobre isso. Devoted to excellence in teaching, learning, and research.



Universidade Harvard

Website Como chegar Salvar

Universidade privada em Cambridge, Massachusetts

Universidade de Harvard é uma universidade privada membro da Ivy League, localizada em Cambridge, Massachusetts, Estados Unidos, e cuja história, influência e riqueza tornam-na uma das mais prestigiadas universidades do mundo. Wikipédia

Endereço: Cambridge, MA 02138, EUA

Custo e taxas de graduação: 44.990 USD (2017)

Taxa de aceitação: 5,4% (2016)

Fundação: 8 de setembro de 1636, Cambridge, Massachusetts, EUA

Subsidiárias: Harvard Business School, MAIS

Sugerir uma edição

Conhece este lugar? Responda perguntas rápidas

Enviar para smartphone

Enviar

Tecnologia de Web Semântica em ação

<https://pt-br.facebook.com> > ... > Faculdade e universidade ▼
Harvard University, Cambridge (Massachusetts). 5.260.870 curtidas · 73.298 falando sobre isso.
Devoted to excellence in teaching, learning, and research,...

The Official Website of Harvard University Athletics - Harvard

<https://www.gocrimson.com/> ▼ Traduzir esta página
Contact Us | Directions | News Feeds(RSS) · Harvard Varsity Club · Ivy League · Harvard University ·
NCAA · ECAC Hockey · IMG College · PrestoSports.

Harvard University | Top Universities

<https://www.topuniversities.com/universities/harvard-university> ▼ Traduzir esta página
Get complete details of Harvard University including available Programs, Ranking data, Fees structure, upcoming events.

Harvard University - Profile, Rankings and Data | US News Best ...

<https://www.usnews.com/best-colleges/harvard-university-2155> ▼ Traduzir esta página
Harvard University is a private institution that was founded in 1636. It has a total undergraduate enrollment of 6,710, its setting is urban, and the campus size is ...

Harvard University (@Harvard) | Twitter

<https://twitter.com/harvard> ▼ Traduzir esta página
The latest Tweets from Harvard University (@Harvard). Devoted to excellence in teaching, learning, and research, and to developing leaders who make a ...

Pesquisas relacionadas a harvard university

- | | |
|---|--|
| harvard university cursos | harvard ex-alunos notáveis |
| harvard university como entrar | harvard subsidiárias |
| harvard cursos | quantos cursos harvard oferece |
| harvard business publishing program | numero de cursos oferecidos em harvard |



Enviar para smartphone

Enviar

Ex-alunos notáveis

Ver mais 45



Theodore Kaczynski



Natalie Portman



Mark Zuckerberg



Barack Obama



Matt Damon

Comentários

4.255 comentários no Google

Comentar

Adicionar uma foto

Pesquisas relacionadas

Ver mais 5



Universid... Yale



Instituto de Tecnologia de Massa... Cambridge

Stanfc
Univers

Universid... Stanford



Universid... Columbia Nova Iorque



Universid... de Oxford Oxford

Mais sobre Universidade Harvard

Fontes

Você gerencia a presença on-line de Universidade Harvard?

Feedback

Avança o emprego de Recursos Semânticos

Initiative for Open Citations I4OC

“The present scholarly communication system inadequately exposes the knowledge networks that already exist within our literature. **Citation data are not usually freely available to access, they are often subject to inconsistent, hard-to-parse licenses, and they are usually not machine-readable.**”

“The aim of this initiative is to promote the availability of data on citations that are **structured, separable, and open.**”

Structured = the data representing each publication and each citation instance are expressed in common, machine-readable formats, and that these data can be accessed programmatically.

Separable = the citation instances can be accessed and analyzed without the need to access the source bibliographic products.

Open means the data are freely accessible and reusable.

Re-imagining discovery and access to research: grants, publications, citations, clinical trials and patents in one place



Our Data Universe, Enriched and Connected

Unlike existing tools, Dimensions brings together various research-related data sources (over 138 million pieces thus far) in a venue that is consistent and accessible to the community.

In addition to deep-indexing, the Dimensions team invests in enhancing existing data for increased searchability and

identifying links between related pieces (nearly 4 billion connections so far).

Going far beyond traditional databases, Dimensions provides the community with a data discovery engine with both context and perspective.

Learn more in [A guide to the Dimensions Data Approach... > more](#)

GRID



Global Research Identifier Database

Cataloging the world's research organisations

[DOWNLOAD THE FREE DATABASE](#)

MAKE SENSE OF YOUR INSTITUTIONAL DATA

Capture data accurately



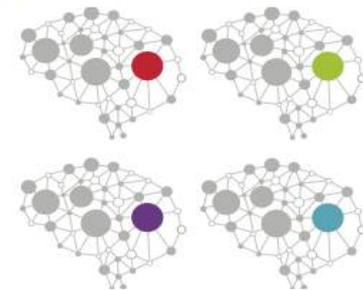
Solve your data capture woes by working with unambiguous institutional information from the start. Our persistent IDs will ensure your data is always perfectly consistent

Ensure robust reporting



Align your data with GRID and open the door to a range of data integration and reporting possibilities. Link to other datasets like ISNI and Eundref seamlessly

Disambiguate your data



No more struggling to find duplicates in your data, or trawling through the history of Paris University to check if XI is Sud or Est. We meticulously clean the data for you

Our Goal

is to revolutionize discovery of scientific knowledge. We are building a visual interface that dramatically increases the visibility of research findings for science and society alike. We are a non-profit organization and we believe that a better way to explore and discover scientific knowledge will benefit us all.



From Smart Content to Smart Data: The Role of SN SciGraph in Open Research

Our topic pages:

Journals

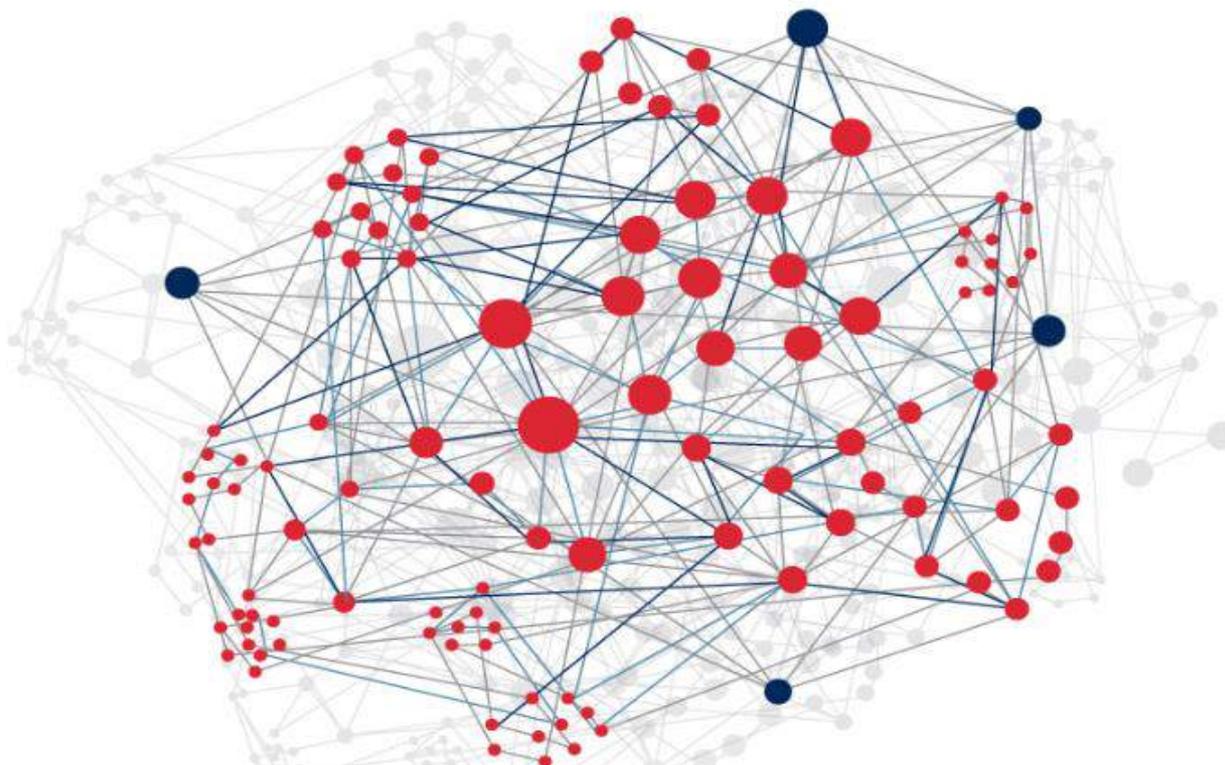
eBooks

Databases

Research & Development

News & Initiatives

System Updates



SN SciGraph

SN SciGraph is the largest Linked Open Data aggregation platform in the scholarly domain. It makes data that is machine-readable, interoperable and re-usable freely available to the global community. Launched in February 2017, SN SciGraph was designed to help researchers, scholars, editors, librarians, funders, and developers overcome the challenge of extracting and merging large datasets from multiple repositories, and in multiple formats.

STM Tech Trends

STM é a principal associação comercial global de editores acadêmicos e profissionais

Participants from 24 organisations, gathering in London, on 1 December 2015,
At Burlington House, Picadilly,
At the Royal Society of Chemistry



Jonathan Morgan	ACS
Richard O Beirne	OUP
Kent Anderson	AAAS
Philip Roberts	CABI
David Smith	IET
Reynold Guida	IEEE
Jeremy Macdonald	RPS
Ann Michael	DeltaThink
Usbrand-J Aalbersberg	Elsevier
Heather Staines	Proquest
Jennifer Robinson	Wolters Kluwer
Sam Bruinsma	Brill
John Sack	HighWire
Richard Kidd	RSC
Will Schweitzer	AAAS
Daniel Schiff	Thieme
James Walker	IOPP
Jasper Simons	APA
Gari Spencer	Wiley
Phill Jones	Digital Science
Paul Groth	Elsevier Labs
Gerry Grenier	IEEE
Thomas Mager	Springer
Eefke Smit	STM
Pam Dixon	World Privacy Forum



2020 STM Tech Trends in summary: The Floodgates of Technology are open!

5 currents interacting: **Current 5 = Publishing**

Publishing: User Centered Publishing delivers Precision Information

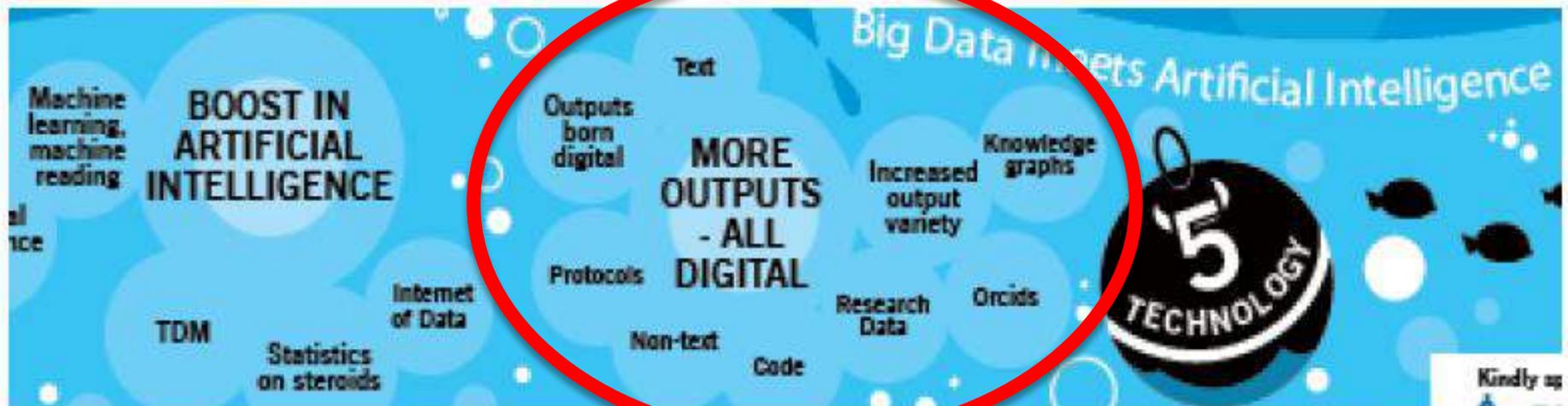
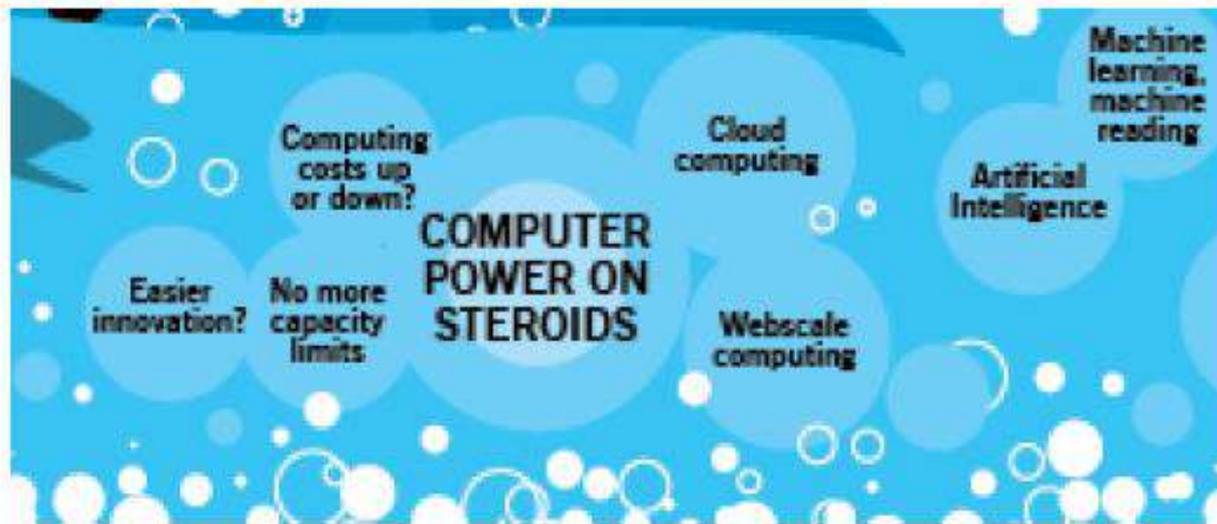
Users: The Machine is the new Reader

Research: Science as a Social Machine

Cyber Security: Data privacy requires a Web of Trust

Technology: Big Data meets Artificial Intelligence

Technology: Big Data Meets Artificial Intelligence



INTEGRITY

VALUE

ScholarX

Editorial services via internet

Open annotation

Open and trustworthy peer review

Lab books & publications connected

MANDATES

Firewalls

CYBER-CRIMINALITY

Hackers

DDOS attacks

Fraudulent IP addresses

Theft

Illegal piracy sites

COLLABORATION AND SHARING

Remote internet labs

Exchange of expert services

Preprint servers

Collaborative writing

Streaming lab data

Transparency on versioning

Social Networks

Portability of manuscripts

Google user data as a service

Lab outputs, streaming

Industry Consolidation

VALUE

SMART SERVICES

Big Data Applications

Automated integrity checks on articles

Checking datasets on validity

BUDGETS

Cyber bullying

INFORMATION WARFARE

Data rights

Misuse of personal data

Non facts

Alternative Facts

Loss of trust

Things as a service (TAAS)

Government Policies

Open linked data

Process changes

AI ASSISTED

Blockchain

Linked open data

Knowledge graphs

Internet of things

RESEARCHERS

Industry Consolidation

Recommended reading services

Automated peer review

Voice for interface

Semantic html

Computer-created articles

Mobile device optimisation

Machine Learning

Text and data mining

VALUE

INNOVATION

Tech Trends 2022

OPEN SCIENCE

- Early sharing
- Includes all research artifacts
- European Open Science Cloud (EOSC)

SOCIAL MEDIA

- Cyber Influencing
- Bots warfare
- Internet Surveillance
- Citizen Science
- How to avoid fake science

EASY ACCESS

- Single Sign-on
- Transparent collaboration
- Complexity of ID Management
- Simple Business Models
- RA21

SHARING PLATFORMS

- A Spotify for Science?
- Find the Napster moment
- How to control governance
- Will it all be Google or Sci-Hub
- Complete platform integration
- Responsible sharing
- Will it all be open?
- CHORUS

BREXIT

Research in Asia

NET NEUTRALITY

METRICS

- New reward systems
- Measuring all outputs

GDPR

- Will it take all development resources?
- May change marketing fundamentally
- Impact on user data analytics
- What does it do to the UX interface?

Publishing to underpin TRUST

- Help avoid crap science
- Quality Assurance of Data

RESEARCH DATA

- Who funds the infrastructure
- Will it bust the pipes?
- Volume is enormous
- What will funders pay for?
- FAIR Data
- Persistent ID's
- Data Management Plans

BLOCKCHAIN

- Can Blockchain help solve the trust issues
- Ensure Authenticity in a Network of Trust
- Is it robust and fast enough?
- Will it all be open?



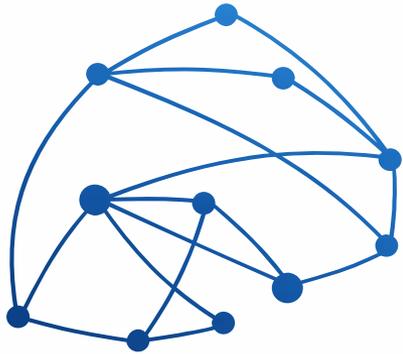
Entering The AI Era
Creative Humans &
Smart Machines



- Open Science
- Research data availability
- Data Management Plans
- All research outputs

- Performance assessments
- Trust and Accountability
- Detect Fraud and Error
- Find research flaws with AI

Por que importa para a Produção Científica Brasileira?



RNP

Como Recursos Semânticos podem ajudar?

- **Relação à fonte correta:** garante que a informação de pessoa ou coisa é da fonte correta pois é inequivocamente identificada (uso de URIs)
- **Resolve múltiplos identificadores:** garante que pessoas e publicações são identificados por números que representam a coisa certa (DOI, passaporte, ORCID)
- **Atributos com significados múltiplos:** atributos de pessoas e coisas têm diferentes significados para diferentes instituições
- **Particularização:** em bases de dados de instituições diferentes, as mesmas informações com a mesma denominação podem ter papéis diferentes
- **Representações múltiplas do mesmo dado:** “nome completo” ou “sobrenome”; “nome do meio”; “primeiro nome“, com ou sem abreviação?
- **Padrão esquemático:** informações passam a poder ser trocadas porque obedecem a um esquema único (ex. XML, JSON, RDF)

Contextos ideais para Recursos Semânticos

Desenvolvimento de novos sistemas

Um vocabulário base pode ser usado como ponto de partida para se projetar os modelos de dados de novos sistemas de informação.

Intercâmbio de informações entre sistemas

Um vocabulário base pode se tornar a base de um esquema de (arquivo) para se trocar dados entre sistemas diferentes.

Integração de dados de múltiplas fontes

O vocabulário permite se corresponder dados provenientes de fontes distintas e criar uma 'tradução', ou correspondência.

Dados abertos e 'descobríveis'

Vocabulários RDF são essenciais para a publicação de dados na Web Semântica.

Por que recorrer à Web Semântica já?

Independência e interoperabilidade de sistemas

A troca de informações se tornará cada vez mais intensa

- Como consequência da **adoção crescente de sistemas computacionais** (e.g. CRIS, RIs, RDMs, serviços web) por agências e instituições acadêmicas, no Brasil e principalmente no mundo.

Muitos conflitos podem advir da falta de interoperabilidade semântica

- E.g. discrepâncias na interpretação de procedimentos e processos, falta de modelos de dados comuns, identificadores e de dados de referência universais.

Maior capacidade de reutilização e descoberta

- A Informações disponíveis em **formato legível por máquina** facilita referência cruzada e a interoperabilidade de dados e serviços e, portanto, aumenta consideravelmente seu poder de **reutilização**.

Por que recorrer à Web Semântica já?

“A capacidade de descobrir e ser descoberto”

Buscadores e robôs (da web) recorrem cada vez mais à semântica

- Para enriquecer conteúdos, contextualizar as informações, identificar relações e compreender significados.

Ser reconhecível por máquina aumenta a visibilidade da informação

- E a possibilidade de descoberta de seus sujeitos e objetos, em ambos os sentidos.

Recursos de web semântica permitem que se busque as informações diretamente da web

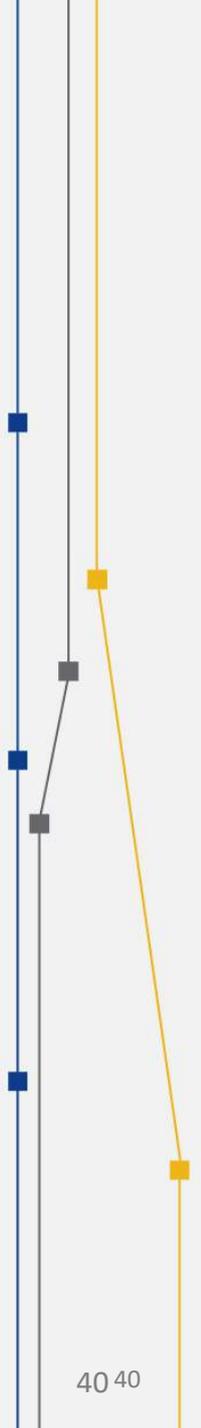


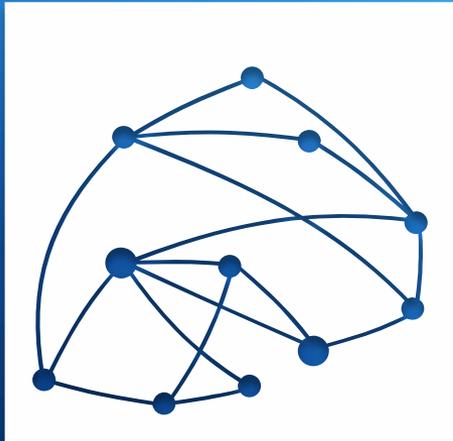
Será que isso é bom para a produção científica nacional?

“Você pode nos contar tudo sobre suas publicações (de novo)?”

VS.

“Achamos que isso é seu. Você pode, por favor, confirmar?”





RNP

Uma Representação para a Pós-Graduação

ISA²

Interoperability solutions for public administrations, businesses and citizens

[European Commission](#) > [ISA²](#) > [Actions](#) > [Improving semantic interoperability in European eGovernment systems](#)[ABOUT](#)[ISA2CONF18](#)[EIF](#)[ACTIONS](#)[SOLUTIONS](#)[S&R AWARDS](#)[DASHBOARD](#)[CONSULTATIONS](#)[LIBRARY](#)[EVENTS](#)[NEWS](#)[BLOG](#)[CONTACT](#)

Improving semantic interoperability in European eGovernment systems



2016.07 SEMIC: Promoting Semantic Interoperability amongst the EU Member States

When is this action of interest to you?

You work in the area of government metadata management and want to check upon and possibly reuse semantic assets (metadata and/or reference data), understand semantic interoperability requirements, approaches, tools, lessons learnt and case studies.

You want to share your own semantic assets.

Which solutions can ISA² offer to you?

The ISA programme has developed a number of semantic specifications and solutions which are available to Public administrations in Europe free of charge.

I. Core Vocabularies

These are simplified, reusable, and extensible data models that capture the fundamental characteristics of an entity in a context-neutral way. Public administrations can use and extend the Core Vocabularies in the following contexts:

PROJECT CORE DATA

Project start/finish: Q2/2016 to Q2/2019

Parties involved: European Commission (DG DIGIT, DG CNECT, DG ESTAT, DG SG), Publications Office

ISA² contribution: EUR 3 329 000

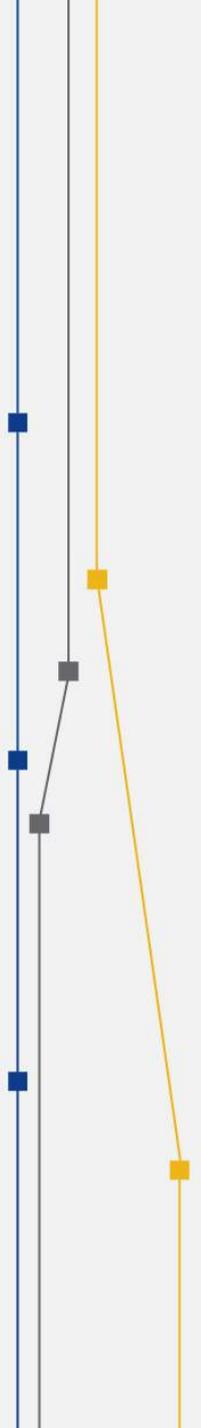
WANT TO KNOW MORE?

Our solutions

- [Core Vocabularies](#)
- [ADMS](#)

Alguns desafios de representação própria

1. Comportar a variedade de relações e papéis desempenhados pelos agentes da pós-graduação
2. Aproveitar recursos abertos já integrados à LOD (e.g. vocabulários e ontologias)
3. Ser independente de implementações técnicas proprietárias (e.g. softwares, serviços)
4. Ser independente de modelos e “padrões” proprietários
5. Requerer o mínimo de manutenção e máxima estabilidade

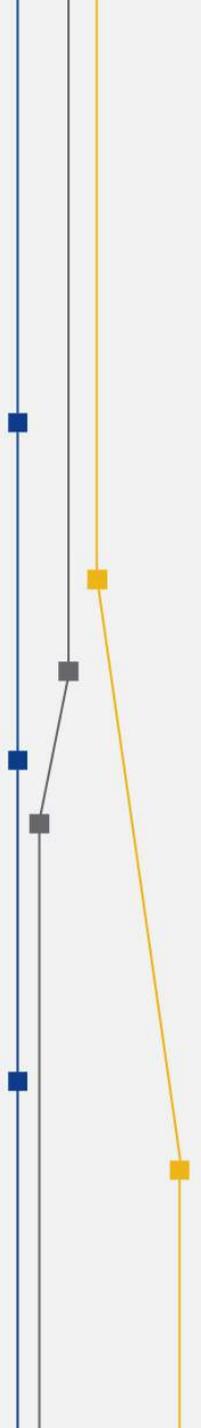


Vocabulários e ontologias em uso na LOD

Usual prefix	Nome	URI
adms	Asset Description Metadata Schema, ADMS	http://www.w3.org/ns/adms#
bibo	The Bibliographic Ontology	http://purl.org/ontology/bibo/
cerif	CERIF Ontology	http://www.eurocris.org/ontologies/cerif/1.3#
dcat	The data catalogue vocabulary, DCAT	http://www.w3.org/ns/dcat#
dcterms	Dublin Core	http://purl.org/dc/terms/
foaf	Friend of a Friend	http://xmlns.com/foaf/0.1/
frbr	Functional Requirements for Bibliographic Records	http://purl.org/vocab/frbr/core#
mo	Music Ontology	http://purl.org/ontology/mo/
org, rov	ORG & RegOrg	http://www.w3.org/ns/org#
owl	Web Ontology Language	http://www.w3.org/2002/07/owl#
qb	The RDF Data Cube Vocabulary	http://purl.org/linked-data/cube#
rdf	RDF	http://www.w3.org/1999/02/22-rdf-syntax-ns#
rdfs	RDFS	http://www.w3.org/2000/01/rdf-schema#
schema	Schema.org	http://schema.org/
skos	Simple Knowledge Organization System	http://www.w3.org/2004/02/skos/core#
umbel	Umbel Ontology	http://umbel.org/umbel#
v	vCard Ontology	http://www.w3.org/2006/vcard/ns#
vann	A vocabulary for annotating descriptions of vocabularies	http://purl.org/vocab/vann/
vivo	VIVO ISF Ontology	http://vivoweb.org/ontology/core#
xsd	XML Schema	http://www.w3.org/2001/XMLSchema#

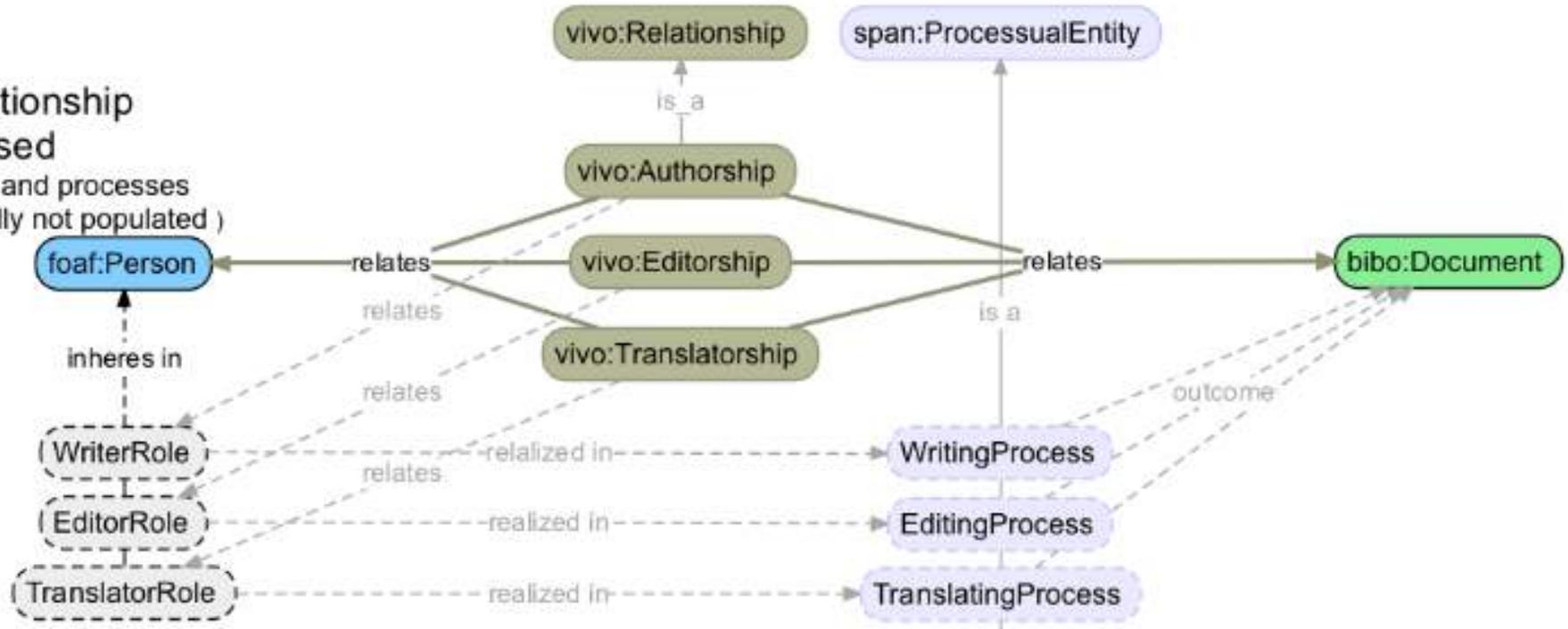
VIVO - Ontologia

- É usada para **representar a especialização de pessoas** envolvidas na criação, transmissão e preservação de conhecimento e trabalhos criativos.
- Representa o conhecimento, descrevendo as atividades e realizações das pessoas **em termos de seus relacionamentos** com artefatos particulares do trabalho, recursos que usam, instituições que os empregam e outros indicadores.
- É independente de conhecimento ou de domínio criativo específico.
- Sustenta a identificação e **avaliação de impacto individual e de grupos de pessoas**, bem como a identificação e reutilização de seus produtos.



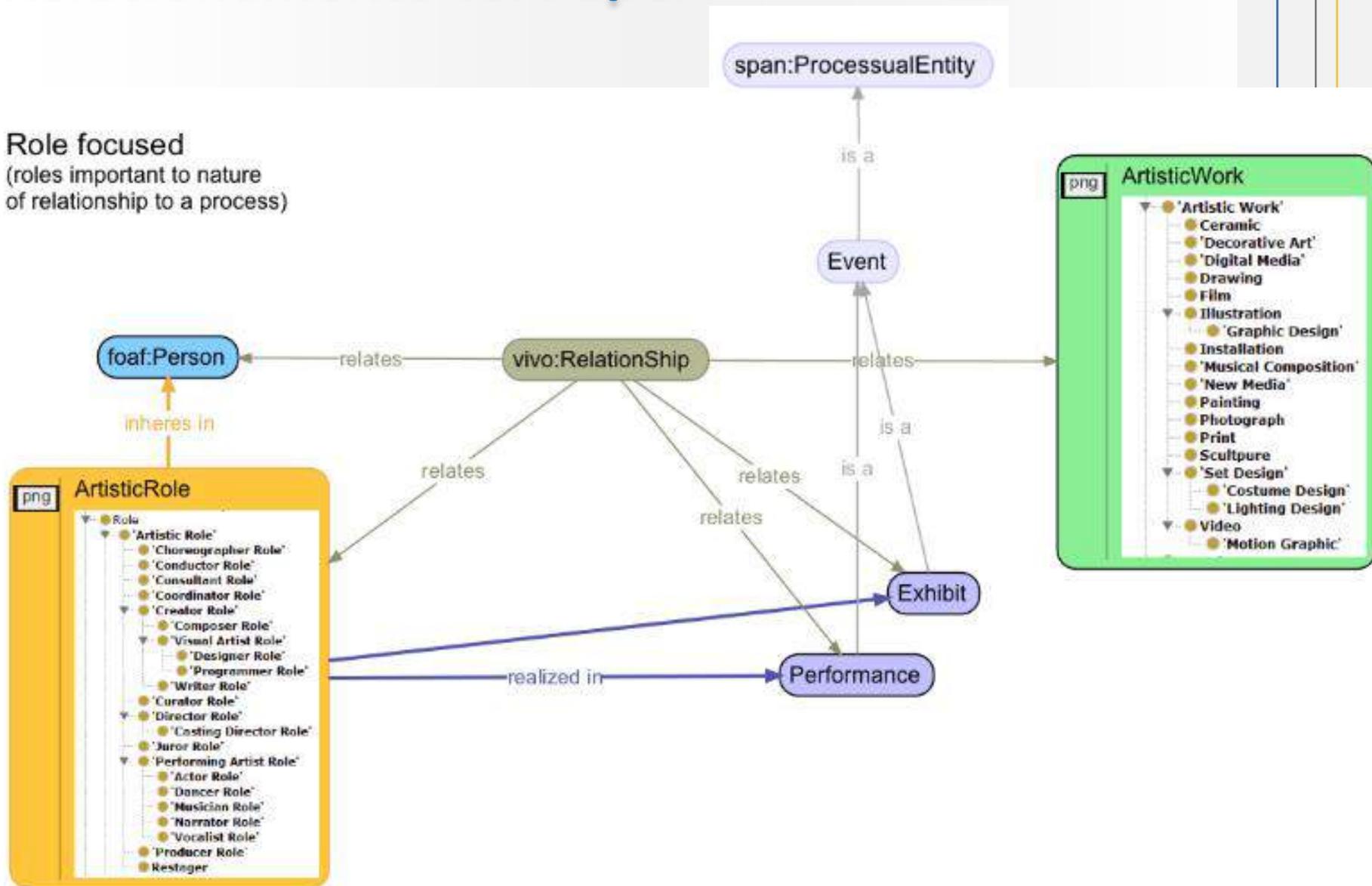
Relacionamento vs. Papel

Relationship
focused
(roles and processes
typically not populated)

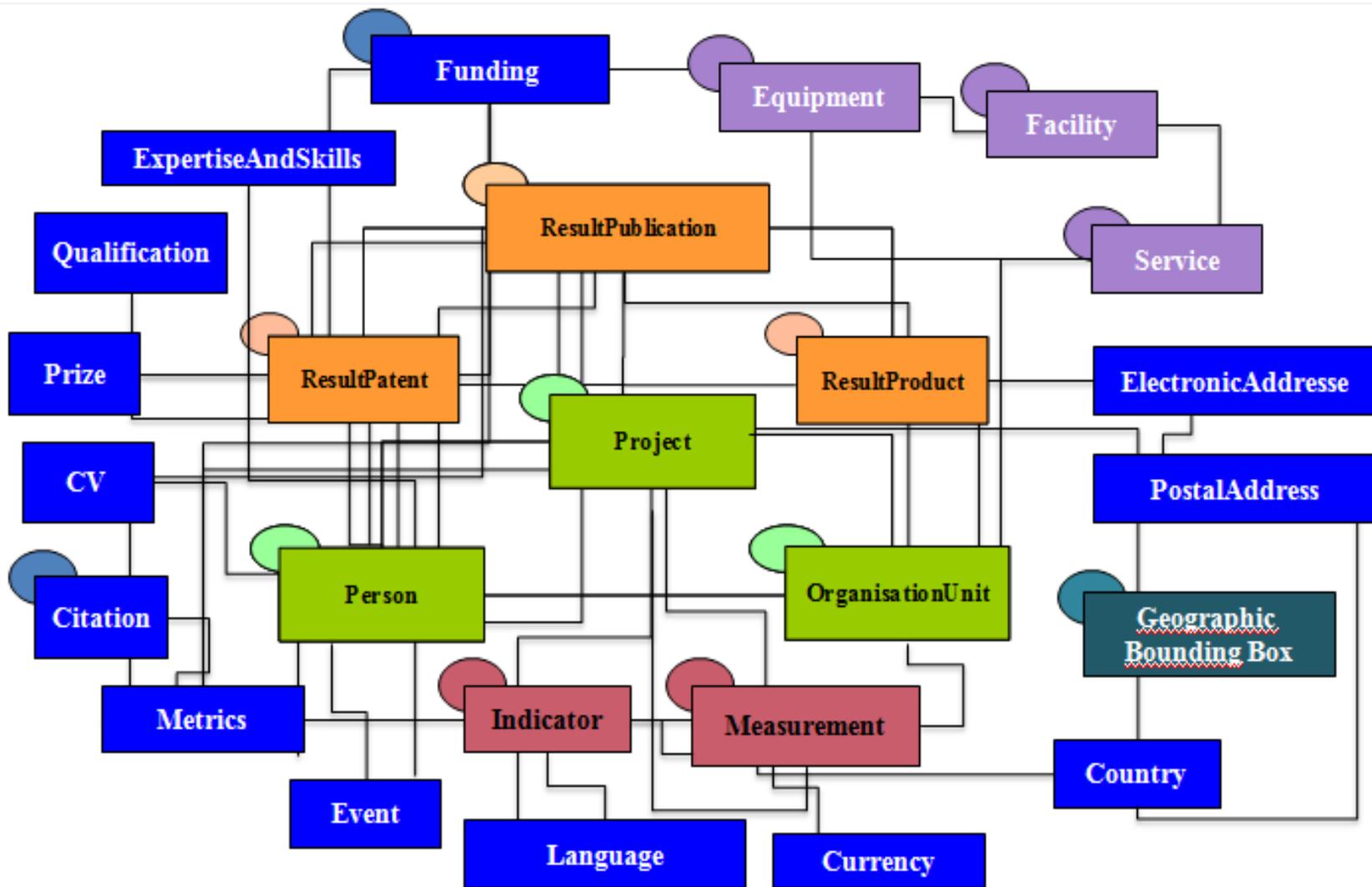


Relacionamento vs. Papel

Role focused
(roles important to nature
of relationship to a process)



CERIF – Modelo de Dados



CERIF – Características do Modelo de Dados

PESSOA

PESSOA_PROJETO

ID Pessoa
ID Projeto
Papel: pesquisador
Início: 01-01-2010
Fim:

PROJETO

PESSOA_PROJETO

ID Projeto
ID Pessoa
papel: gerente
Início: 01-03-2011
Fim: 31-12-2013

CERIF – Características do Modelo de Dados

Esquema A

E.g. Bioquímica,
termos como usados
na Medicina

Classificação no
Esquema A

CLASS _ CLASS

Class_termo A

Class_termo B

Relação:

(e.g. É igual, subtipo,
equivallente, etc...)

Tabela de correspondência
mapeando um termo do
Esquema A para um termo do
Esquema B, incluindo
expressões da natureza ou do
significado da relação entre os
dois termos.

Esquema B

E.g. Bioquímica,
termos como usados
na Biologia

Classificação no
Esquema B

CERIF – Características do Modelo de Dados

Objeto A

Obj-A_Obj-B

ID Obj-A

ID-Obj-B

Relação / Papel

Início:

Fim:

Objeto B

Esquema Class. 1

E.g. Papéis de
uma pessoa num
projeto

Esquema Class. 2

E.g. Papéis de
uma pessoa em
uma produção

Esquema Class. 3

E.g. Tipos de
produção

.....

Esq. Class. N

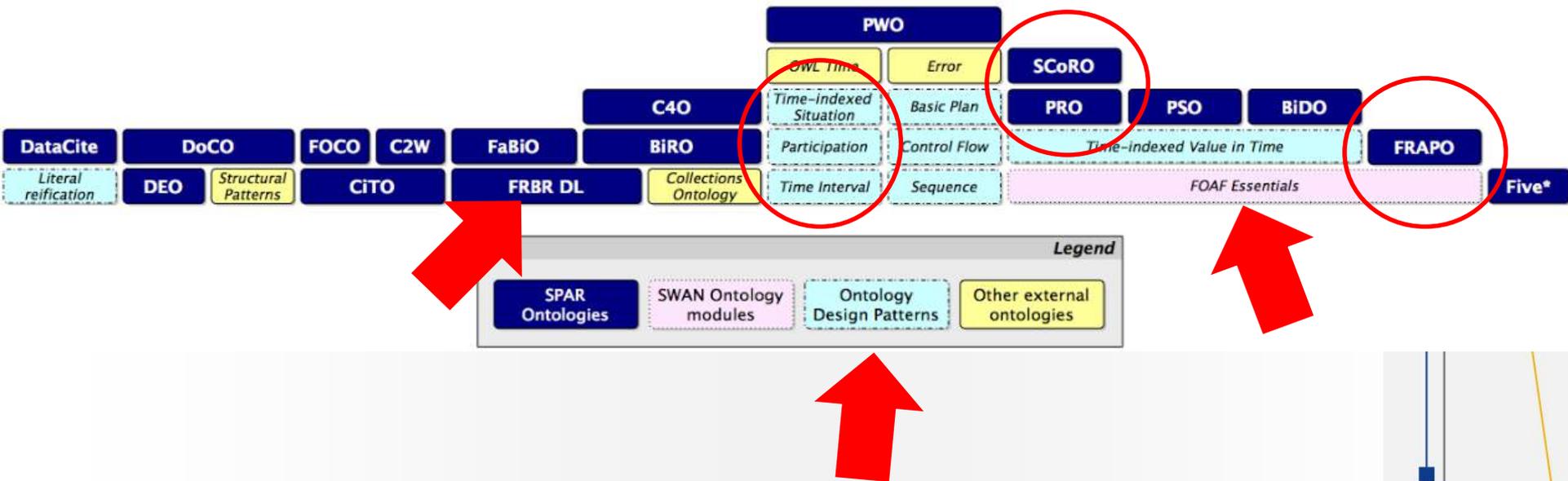
E.g. Tipos de
vínculo.

Vocabulário controlado para expressar o “significado” da relação.

Fonte: **euroCRIS**

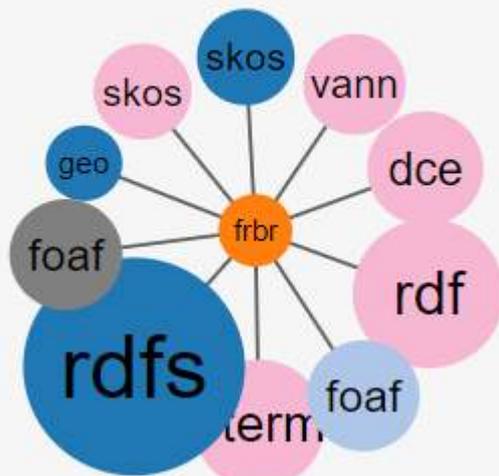
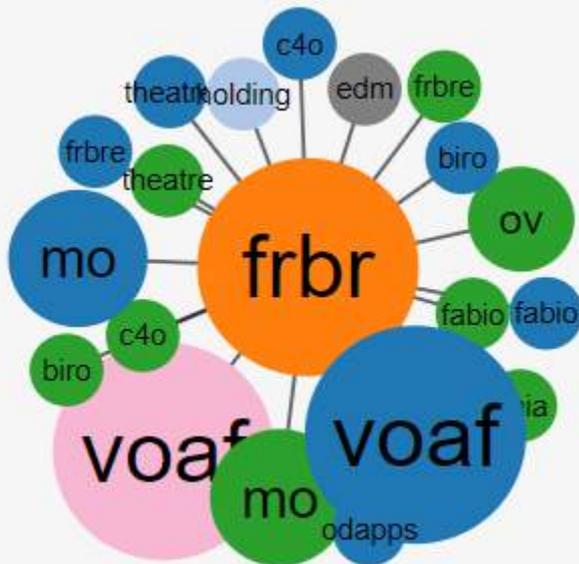
Família de Ontologias SPAR

Semantic Publishing and Referencing Ontologies (SPAR)

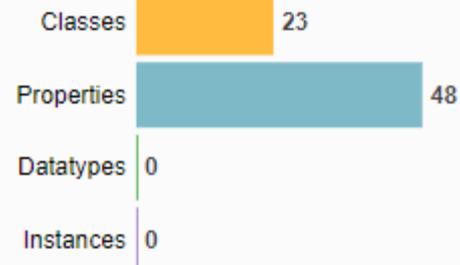


Functional Requirements for Bibliographic Record (FRBR)

- É um modelo proposto pela **Federação Internacional de Associações de Bibliotecas (IFLA)** para descrever documentos e sua evolução.
- Funciona tanto para recursos físicos como digitais e provou ser muito flexível e poderoso.
- Um dos aspectos mais importantes do FRBR é o fato de que ele **não está associado a nenhum esquema ou implementação de metadados específico**.

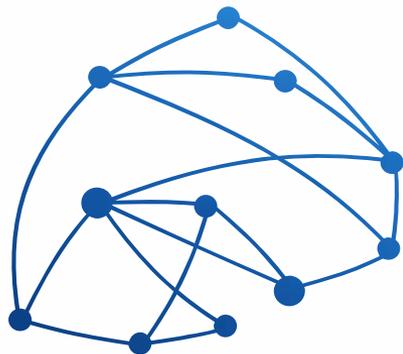


Statistics



LOD

Vocabulary used in **12 datasets**



RNP

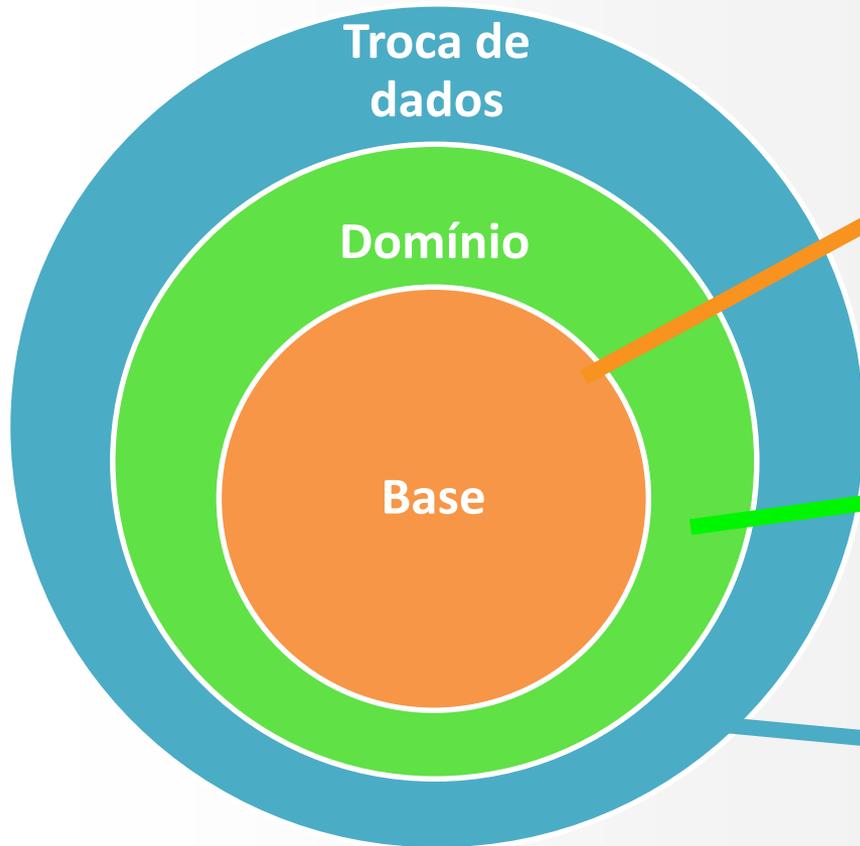
Uma Representação Semântica para a Pós-Graduação

Opção por um Vocabulário Base

Um vocabulário (base) é um modelo de dados simplificado, reutilizável e extensível que captura as características fundamentais das entidades de dados de maneira neutra em relação ao contexto.

Vocabulários são a base da construção da representação semântica de entidades e sua inserção no Linked Open Data.

Opção por um Vocabulário Base



1. Vocabulário base: um modelo de dados neutro em relação ao contexto que captura as características fundamentais de entidades-chave

2. Vocabulário de domínio: uma visão conceitual de um domínio específico

3. Vocabulário para troca de dados: define e descreve a estrutura e o conteúdo para troca de dados em um contexto específico

Opção por um Vocabulário Base

Esquema XML, APIs específicas

Esquemas de metadados, vocabulário especializado

Vocabulário RDF/RDFa publicado na Web



Vocabulário da Pós-Graduação (VPG)

Estratégia

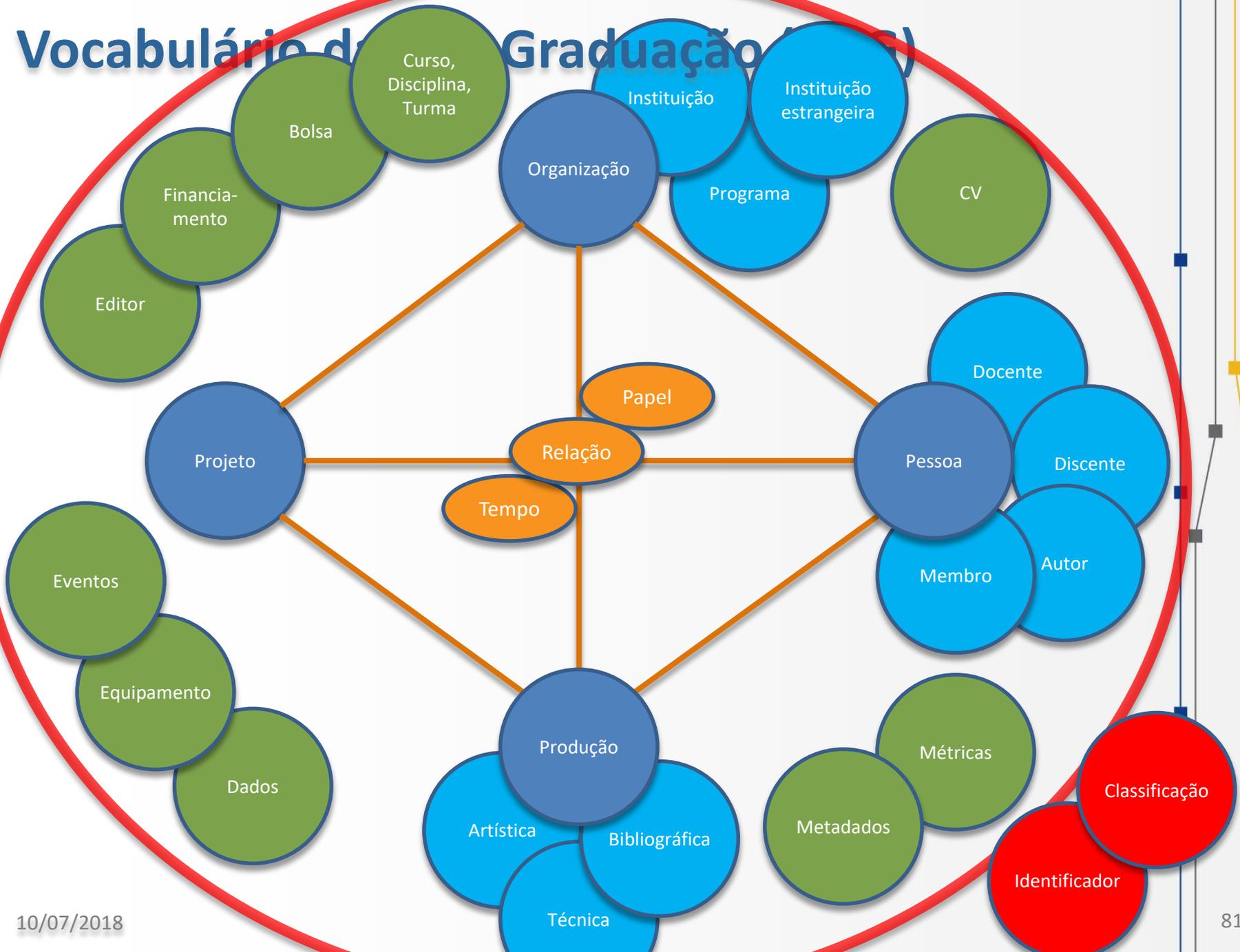
1. (Re)aproveitar esquemas, padrões de metadados, vocabulários e ontologias existentes e consagrados
2. Comportar a variedade de relações e papéis existente entre sujeitos e objetos da pós-graduação
3. Representar a temporalidade de tais relações e papéis
4. Possibilitar a acomodação de taxonomias e classificações de domínios específicos
5. Permitir futuras expansões e particularizações de domínios e casos de uso específicos

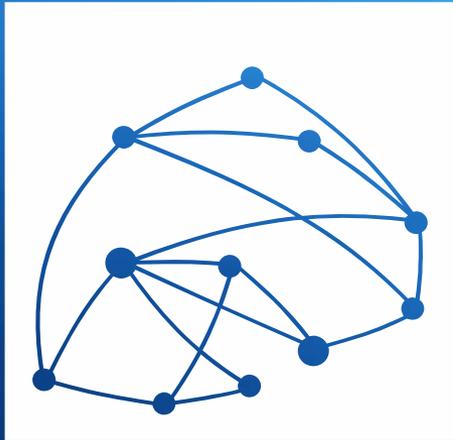
Critério para aproveitamento

Ser aberto e integrado significa “falar a mesma língua que a Web”. Nesse sentido, vêm sendo privilegiados vocabulários e padrões que:

1. Sejam **estáveis**, mantidos por consórcios ou órgãos colegiados;
2. Já **amplamente adotados** no LOD e outras entidades e consórcios;
3. **Independentes** de modelos de dados ou padrões de metadados específico;
4. Cubram ao máximo a **diversidade de classificações**, taxonomias e categorias exigidas pelo sistema brasileiro;
5. Sejam **flexíveis quanto à inclusão** de novos termos e propriedades para abarcar novos casos de uso (e.g. bolsas, financiamentos, gestão acadêmica etc.);
6. Possibilitem **fácil integração** ou correspondências com outros vocabulários;
7. Sejam o **mais genérico possíveis** visando acomodar mais facilmente modelos e vocabulários proprietários já adotados pela comunidade local.

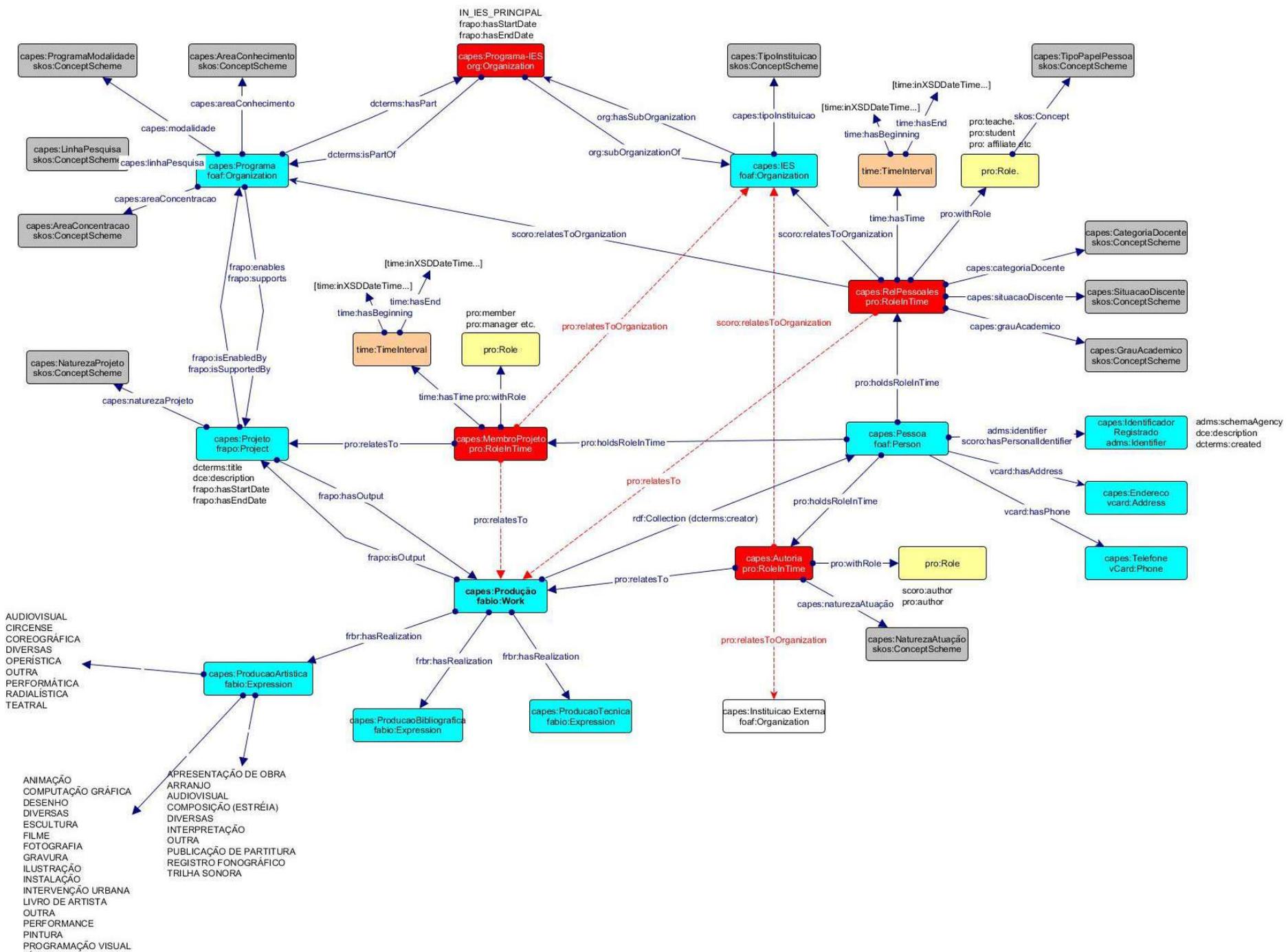
Vocabulário de Graduação (V.G.)





RNP

Status do Projeto



Classes Principais – Versão de estudo

Entidade legada

AREA_CONCENTRACAO
 AREA_CONHECIMENTO
 AUTOR
 CATEGORIA_DOCENTE
 DISCENTE
 DOCENTE
 ENDERECO
 GRAU_ACADEMICO
 IDENTIFICADOR
 IES
 LINHA_PESQUISA
 MEMBRO_PROJETO
 NATUREZA_ATUACAO
 PAIS
 PAPEL
 PARTICIPANTE
 PESSOA
 PRODUCAO

 PPG-IES
 PROGRAMA_MODALIDADE
 PPG
 PROJETO
 RELACIONAMENTO
 TELEFONE
 FINALIDADE_ENDEREÇO

Base vocab:Class

skos:ConceptScheme
 skos:ConceptScheme
 pro:RoleInTime
 skos:ConceptScheme
 pro:RoleInTime
 pro:RoleInTime
 vcard:Address
 skos:ConceptScheme
 dcterms:Identifier
 foaf:Organization
 skos:ConceptScheme
 pro:RoleInTime
 skos:ConceptScheme
 vcard:Country
 pro:Role; scoro:Role
 pro:RoleInTime
 foaf:Person
 frbr:Work
 foaf:Organization;
 schema:Organization
 skos:ConceptScheme
 foaf:Organization
 foaf:Project
 pro:RoleInTime
 vcard:Phone
 skos:ConceptScheme

Class label

capes:AreaConcentração
 capes:AreaConhecimento
 capes:Autor(ia)
 capes:CategoriaDocente
 capes:RelPessoales
 capes:RelPessoales
 capes:EnderecoFisico
 capes:GrauAcademico
 capes:IdentificadorRegistrado
 capes:InstituicaoEnsino
 capes:LinhaPesquisa
 capes:MembroProjeto
 capes:NaturezaAtuacao

 capes:PapelPessoa
 capes:RelPessoales
 capes:Pessoa
 capes:Producao

 capes:Programales
 capes:ProgramaModalidade
 capes:Programa
 capes:Projeto
 capes:RelPessoales
 capes:Telefone
 capes:TipoFinalidadeEndereco

TIPO_IDENTIFICADOR_REGISTRADO
 TIPO_INSTITUICAO
 TIPO_PAPEL
 TIPO_PRODUCAO
 TIPO_RELACIONAMENTO
 TIPO_SITUACAO_DISCENTE

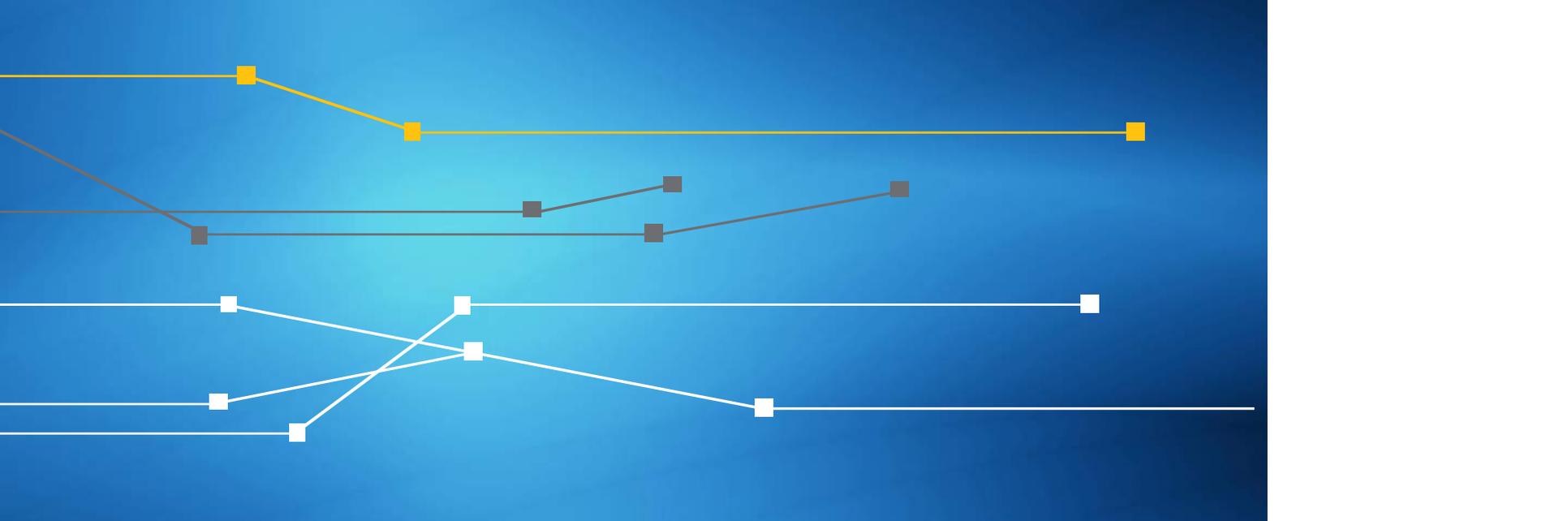
capes:TipoIdentificador
 capes:TipoInstituicao
 capes:TipoPapel
 capes:TipoProducao
 capes:TipoRelacionamento
 capes:TipoSituacaoDiscente

Propriedades Sugeridas – Versão de estudo

CLASSE PRINCIPAL	Atributo (Sucupira)	vocab:Property
PPG; PRODUCAO	NM_AREA_CONCENTRACAO	skos:prefLabel
AUTOR; PPG	CD_AREA_CONHECIMENTO	skos:altLabel
AUTOR; PPG	NM_AREA_CONHECIMENTO	skos:prefLabel
AUTOR	NM_AUTOR	vocab:authorString; vcard:givenName
AUTOR	NM_ABNT_AUTOR	vcard:preferredName
DOCENTE	NM_CATEGORIA_DOCENTE	skos:prefLabel
DOCENTE	DS_CATEGORIA_DOCENTE	dce:description
EMAIL; PESSOA; PPG-IES	DS_CORREIO_ELETRONICO	vcard:hasEmail
DOCENTE	DH_CADASTRO	time:hasBeginning
DOCENTE	DH_EXCLUSAO	time:hasEnd
DOCENTE	DT_FIM	frapo:hasEndDate
DOCENTE	DT_INICIO	frapo:hasStartDate
ENDERECO; PESSOA	DS_COMPLEMENTO	vcard:streetAddress
ENDERECO; PESSOA	DS_LOGRADOURO	vcard:streetAddress
ENDERECO; PESSOA	DS_NUMERO_ENDERECO	vcard:streetAddress
ENDERECO; PESSOA	NM_CIDADE_EXTERIOR	vcard:locality
ENDERECO; PESSOA	NR_CEP	vcard:postalCode
PESSOA; URL	DS_URL	foaf:homepage
DISCENTE	CD_GRAU_ACADEMICO	skos:altLabel
DISCENTE	NM_GRAU_ACADEMICO	skos:prefLabel
IDENTIFICADOR; PESSOA	DS_IDENTIFICADOR_REGISTRADO	dce:description
IDENTIFICADOR; PESSOA	NM_ORGAO_EXPEDIDOR	adms:schemaAgency
IDENTIFICADOR; PESSOA	DT_INICIO_VALIDADE	dcterms:valid
IDENTIFICADOR; PESSOA	DT_FIM_VALIDADE	dcterms:valid
IDENTIFICADOR; PESSOA	DT_EXPEDICAO	dcterms:issued
IES; DISCENTE; PARTICIPANTE	CD_INSTITUICAO_CAPES	dcterms:identifier
IES; DISCENTE; PARTICIPANTE	CD_INSTITUICAO_ENSINO	dcterms:identifier
IES; DISCENTE; PARTICIPANTE	SG_INSTITUICAO_ENSINO	frapo:hasAcronym
PPG-IES; PPG; PROJETO; PRODUCAO	DS_LINHA_PESQUISA	dce:description
PPG-IES; PPG; PROJETO; PRODUCAO	NM_LINHA_PESQUISA	skos:prefLabel
MEMBRO_PROJETO	DT_FIM	frapo:hasEndDate
MEMBRO_PROJETO	DT_INICIO	frapo:hasStartDate
MEMBRO_PROJETO	TP_RESPONSABILIDADE_MEMBRO	dce:description
AUTOR	NM_NATUREZA_ATUACAO	skos:prefLabel; dcterms:name; rdf:label
AUTOR; PESSOA; ENDERECO	CD_PAIS_GEONAME	gn:countryCode
AUTOR; PESSOA; ENDERECO	CD_PAIS_ISO	gn:isoCountryCode
AUTOR; PESSOA; ENDERECO	NM_PAIS	vcard:countryName

Status do Projeto

1. Levantamento dos dados, processos e modelos de dados existentes
2. Identificação das principais entidades, relações e taxonomias em prática
3. Seleção e remodelagem dos dados base para o vocabulário
4. Levantar vocabulários, esquemas e padrões de metadados abertos que possam ser aproveitados
- 5. Definição dos recursos abertos que serão aproveitados e os que serão criados**
- 6. Confecção do vocabulário em suas diferentes formas de aplicação:**
 - Linguagem OWL/RDF para publicação na Web
 - Esquema XML para troca de dados por meio de arquivos
- 7. Elaboração de documentação e difusão**
- 8. Capacitação e suporte ao uso em beta**
- 9. Publicação na Linked Open Data**



Obrigado!

Alexandre P. Uchoa

RNP – Rede Nacional de Pesquisa, RJ

+55 21 98153-1143

alexandre.uchoa@rnp.br



MINISTÉRIO DA
DEFESA

MINISTÉRIO DA
CULTURA

MINISTÉRIO DA
SAÚDE

MINISTÉRIO DA
EDUCAÇÃO

MINISTÉRIO DA
**CIÊNCIA, TECNOLOGIA,
INOVAÇÕES E COMUNICAÇÕES**

