

Dados abertos no contexto atual e suas aplicações

Seminário sobre Política de Dados Científicos
Maio/2016



Data



Information



Presentation



Knowledge



Imagem extraída de <http://epicgraphic.com/data-cake/>

EpicGraphic.com

O que são dados?

Simple observações sobre os fatos.

- Facilmente estruturados
- Legível por máquinas
- Frequentemente quantificados
- Intercambiável

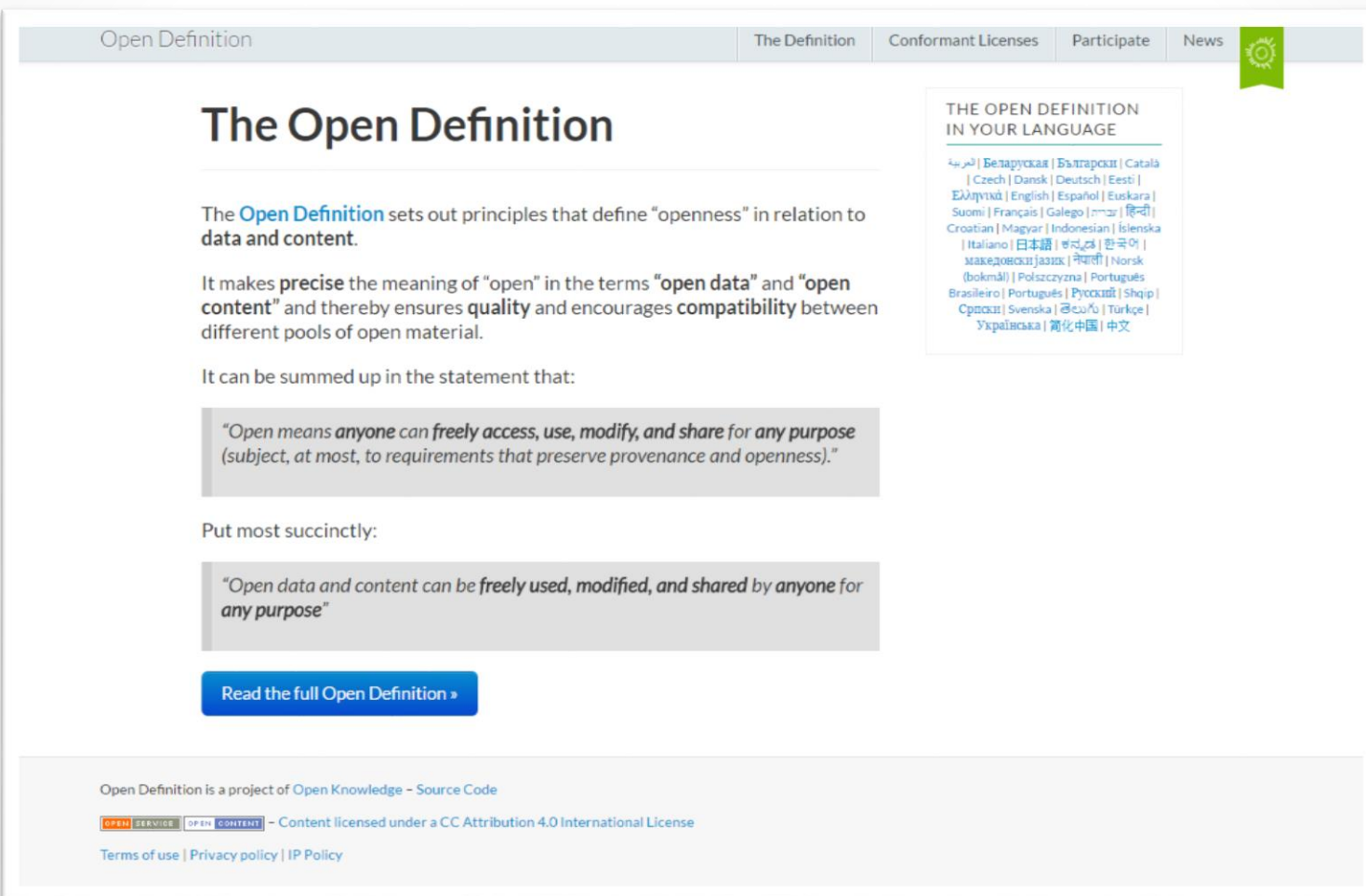


Imagem extraída de <http://opendefinition.org/>

Definição de Dado aberto

Dado aberto é um dado que pode ser livremente utilizado, reutilizado e redistribuído por qualquer um para qualquer propósito.

Os oito princípios (OPENGOVDATA, 2007)

Completos	Primários	Atuais	Acessíveis	Processáveis por máquina	Acesso não discriminatório	Formatos não proprietários	Livres de licença
Todos os dados públicos estão disponíveis	Como coletados na fonte com maior nível de granularidade e sem agregação	Disponibilizados rapidamente para preservar seu valor	Disponibilizados para o público mais amplo possível para qualquer propósito	Razoavelmente estruturados para permitir seu processamento automatizado	Disponível a qualquer um, sem necessidade de identificação ou registro	Disponível em formato que ninguém tenha controle exclusivo	Os dados não estão sujeitos a regulações de direitos autorais, marcas, patentes ou segredo industrial

As três “leis” de Eaves (EAVES, 2009)

1. Se o dado não pode ser encontrado e indexado na Web, ele não existe;
2. Se não estiver aberto e disponível em formato compreensível por máquina, ele não pode ser reaproveitado; e
3. Se algum dispositivo legal não permitir sua replicação, ele não é útil.

Disponibilidade e acesso

- Dados disponíveis
- Internet
- Formato favorável



Dados Abertos

Participação universal

- Sem discriminação
P. ex.: “não comercial”, “fins acadêmicos”, etc.

Reutilização e redistribuição

- Termos de utilização
 - Remixagem
-

As cinco estrelas dos dados abertos (BERNERS-LEE, 2006)

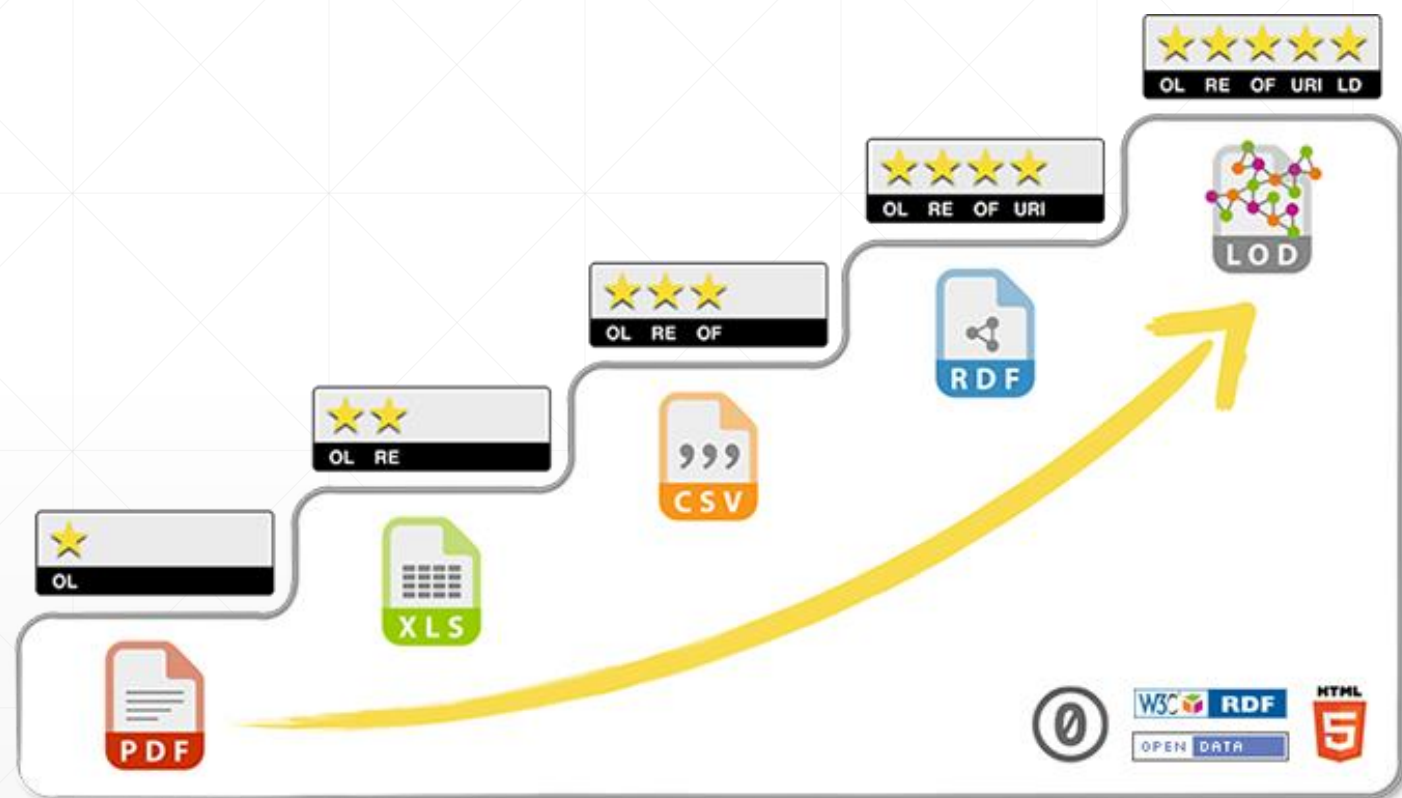


Imagem extraída de <http://5stardata.info/en/>

OL = open license
RE = structured data
OF = non-proprietary format
URI = use of URIs
LD = linked data

Aplicações e proposta de valor

- Dados Abertos Governamentais ou *Open Government Data* (OGD)
 - Melhoria ou criação de produtos e serviços privados
 - Inovação
 - Produção de novos conhecimentos a partir da combinação de várias fontes de dados
-

Artigos científicos

ELSEVIER

Q SEARCH ≡ MENU

Elsevier > About > Open science > Research data > Open data

Open data

Increasing access to research data

Increasing access to research data helps researchers to validate and build upon important discoveries and observations. At Elsevier, we are keen to support researchers to store, share, discover, and use data. We have already been helping authors to post their data in relevant public data repositories and [link back to their published article](#). With our latest Open Data pilot we are adding another option by providing authors with the opportunity to make their supplementary files with raw research data available open access on ScienceDirect.

What are the benefits?

- **Promotes data sharing:** Authors can easily share their raw research data with the public, giving their readers the opportunity to validate, compare and build upon their findings.
- **Greater choice:** Another open access option for

What is the Open Data pilot?

- Elsevier supports the principle that "Raw research data should be made freely available to all researchers" (as expressed in the [STM Brussels Declaration ↗](#)) and authors are free to publicly post their raw research data (see [Author Rights](#)).

Governo – iniciativa EUA

“My Administration is committed to creating an unprecedented level of openness in Government. We will work together to ensure the public trust and establish a system of transparency, public participation, and collaboration. Openness will strengthen our democracy and promote efficiency and effectiveness in Government.”

Transparency and Open Government Memorandum for the Heads of Executive Departments and Agencies (OBAMA, 2009)



Governo – iniciativa Brasil

- Cofundador e partícipe na *Open Government Partnership* (OGP) desde 2011
 - Lei de Acesso à Informação (LAI) de aplicabilidade federal (união, estados e municípios) em todos os órgãos e entidades
 - Estabelece o direito de acesso à qualquer tipo informação, sem a necessidade de justificar o pedido
 - LAI também estabelece:
 - Itens obrigatórios de divulgação
 - Requisitos técnicos
 - Entre outros
-

Lei de Acesso à Informação Brasileira (LAI) – Lei nº. 12.527/2011

Aplicável aos níveis nacional, estaduais, municipais e do Distrito Federal

Transparência passiva



Cidadãos pedem
pela informação



O governo responde as
informações solicitadas



A LAI define:

- Tempo de resposta
- Formas de recurso, em caso de negativa
- Uso da Internet como principal ferramenta

Transparência ativa



Governo divulga informações
de forma regular



Cidadãos consomem
as informações



A LAI define:

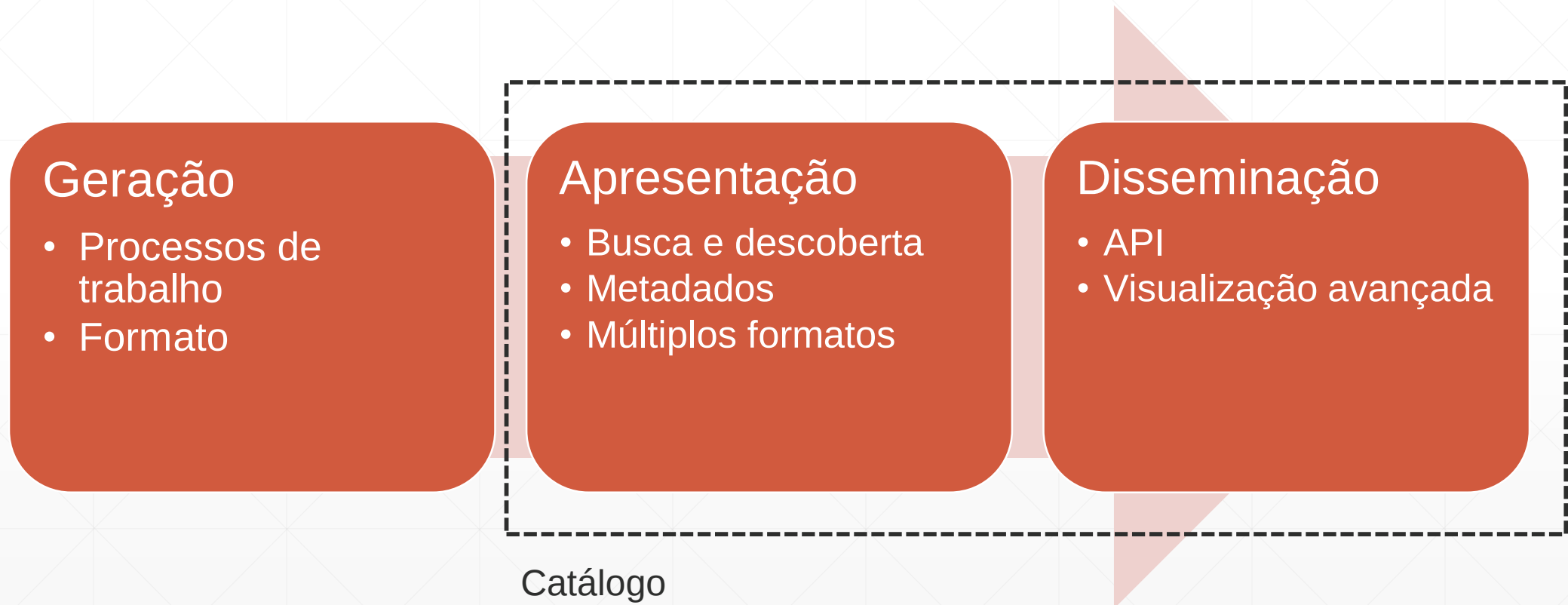
- Uso da Internet para divulgação ampla
- *Itens obrigatórios de divulgação*
- *Requisitos técnicos*

Extraída e traduzida de Corrêa, Corrêa e da Silva (2014)

Plano para abertura de dados

Ciclo de vida

Pontos importantes a considerar



Geração

- Entender e considerar dados abertos desde a geração do dado
 - O formato deve se encaixar nos princípios de abertura de dados
 - Não há um formato mais recomendado que o outro
-

CSV

Character-separated values

Continente;País;Capital
África;Angola;Luanda
América do Norte;Estados Unidos;Washington DC
América Central;México;Cidade do México
América do Sul;Brasil;Brasília
Europa;Espanha;Madri
Europa;Alemanha;Berlim
Oceania;Austrália;Camberra
Ásia;Japão;Tóquio



Continente	País	Capital
África	Angola	Luanda
América do Norte	Estados Unidos	Washington DC
América Central	México	Cidade do México
América do Sul	Brasil	Brasília
Europa	Espanha	Madri
Europa	Alemanha	Berlim
Oceania	Austrália	Camberra
Ásia	Japão	Tóquio

XML

```
<?xml version="1.0" encoding="UTF-8"?>
<Exemplo>
  <Localidade número="1">
    <Continente>África</Continente>
    <País>Angola</País>
    <Capital>Luanda</Capital>
  </Localidade>
  <Localidade número="2">
    <Continente>América do Norte</Continente>
    <País>Estados Unidos</País>
    <Capital>Washington DC</Capital>
  </Localidade>
  <Localidade número="3">
    <Continente>América Central</Continente>
    <País>México</País>
    <Capital>Cidade do México</Capital>
  </Localidade>
  <Localidade número="4">
    <Continente>América do Sul</Continente>
    <País>Brasil</País>
    <Capital>Brasília</Capital>
  </Localidade>
  <Localidade número="5">
    <Continente>Europa</Continente>
    <País>Espanha</País>
    <Capital>Madri</Capital>
  </Localidade>
  <Localidade número="6">
    <Continente>Europa</Continente>
    <País>Alemanha</País>
    <Capital>Berlim</Capital>
  </Localidade>
  <Localidade número="7">
    <Continente>Oceania</Continente>
    <País>Austrália</País>
    <Capital>Camberra</Capital>
  </Localidade>
  <Localidade número="8">
    <Continente>Ásia</Continente>
    <País>Japão</País>
    <Capital>Tóquio</Capital>
  </Localidade>
</Exemplo>
```

JSON

```
{
  "localidade 1": {
    "Continente": "África",
    "País": "Angola",
    "Capital": "Luanda"
  },
  "localidade 2": {
    "Continente": "América do Norte",
    "País": "Estados Unidos",
    "Capital": "Washington DC"
  },
  "localidade 3": {
    "Continente": "América Central",
    "País": "México",
    "Capital": "Cidade do México"
  },
  "localidade 4": {
    "Continente": "América do Sul",
    "País": "Brasil",
    "Capital": "Brasília"
  },
  "localidade 5": {
    "Continente": "Europa",
    "País": "Espanha",
    "Capital": "Madri"
  },
  "localidade 6": {
    "Continente": "Europa",
    "País": "Alemanha",
    "Capital": "Berlim"
  },
  "localidade 7": {
    "Continente": "Oceania",
    "País": "Austrália",
    "Capital": "Camberra"
  },
  "localidade 8": {
    "Continente": "Ásia",
    "País": "Japão",
    "Capital": "Tóquio"
  }
}
```

KML

```
<?xml version="1.0" encoding="UTF-8"?>
<kml xmlns="http://www.opengis.net/kml/2.2">
<Document>
  <Placemark>
    <name>São Paulo</name>
    <description>Cidade de São Paulo</description>
    <Point>
      <coordinates>-23.5476258,-46.6360159</coordinates>
    </Point>
  </Placemark>
</Document>
</kml>
```

geoJSON

```
{
  "type": "Feature",
  "geometry": {
    "type": "Point",
    "coordinates": [125.6, 10.1]
  },
  "properties": {
    "name": "Dinagat Islands"
  }
}
```

Apresentação/Disseminação CKAN

- Software livre para catalogação de dados abertos
 - Desenvolvido pela Open Knowledge (<http://okfn.org>)
 - Open-source
 - Disponível em mais de 50 idiomas (inclusive PT-BR)
 - Disponibiliza uma *sandbox* em <http://demo.ckan.org>.
-

 / Datasets

 Organizations

Herefordshire (1)

 Tags

trees (1)

nature conservation (1)

arboretum (1)

 Formats

KML (1)

GeoJSON (1)

CSV (1)

 Licenses

 Add Dataset

Specimen Trees in Queenswood Country Park in Herefordshire



1 dataset found for "Specimen Trees in Queenswood Country Park in Herefordshire"

Order by:

Relevance



Specimen Trees in Queenswood Country Park in Herefordshire

The Queenswood Country Park in Herefordshire has an arboretum with many specialist trees. There is a web page that allows you to search for specimen trees This file contains the...

[GeoJSON](#) [CSV](#) [KML](#)

You can also access this registry using the [API](#) (see [API Docs](#)).

Algumas instâncias ao redor do mundo



Data.gov



data.gov.au



Data.gov.uk



datos.gob.mx

Mais em
<http://ckan.org/instances>



Data.Rio



Dados Recife



The Brazilian Senate

Datahub.io

- Instância CKAN em produção
- Patrocinada pela comunidade Open Knowledge

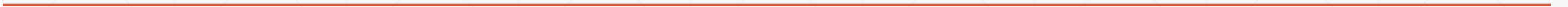


Outros softwares de catálogo/repositório

- Socrata (<http://ckan.org/>)
 - Junar (<http://www.junar.com/>)
-

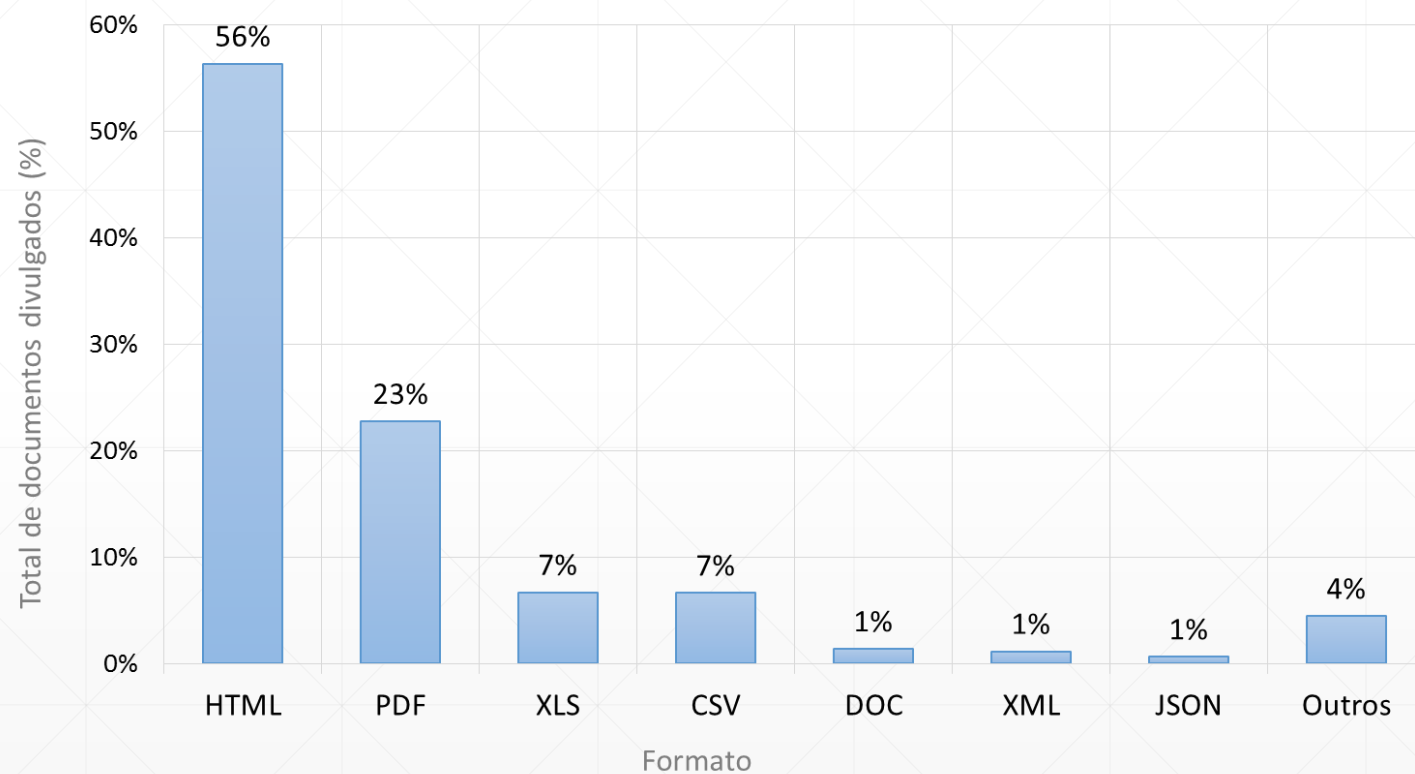
OpenSpending

- Visualização avançada
- Integração com datasets em instâncias CKAN



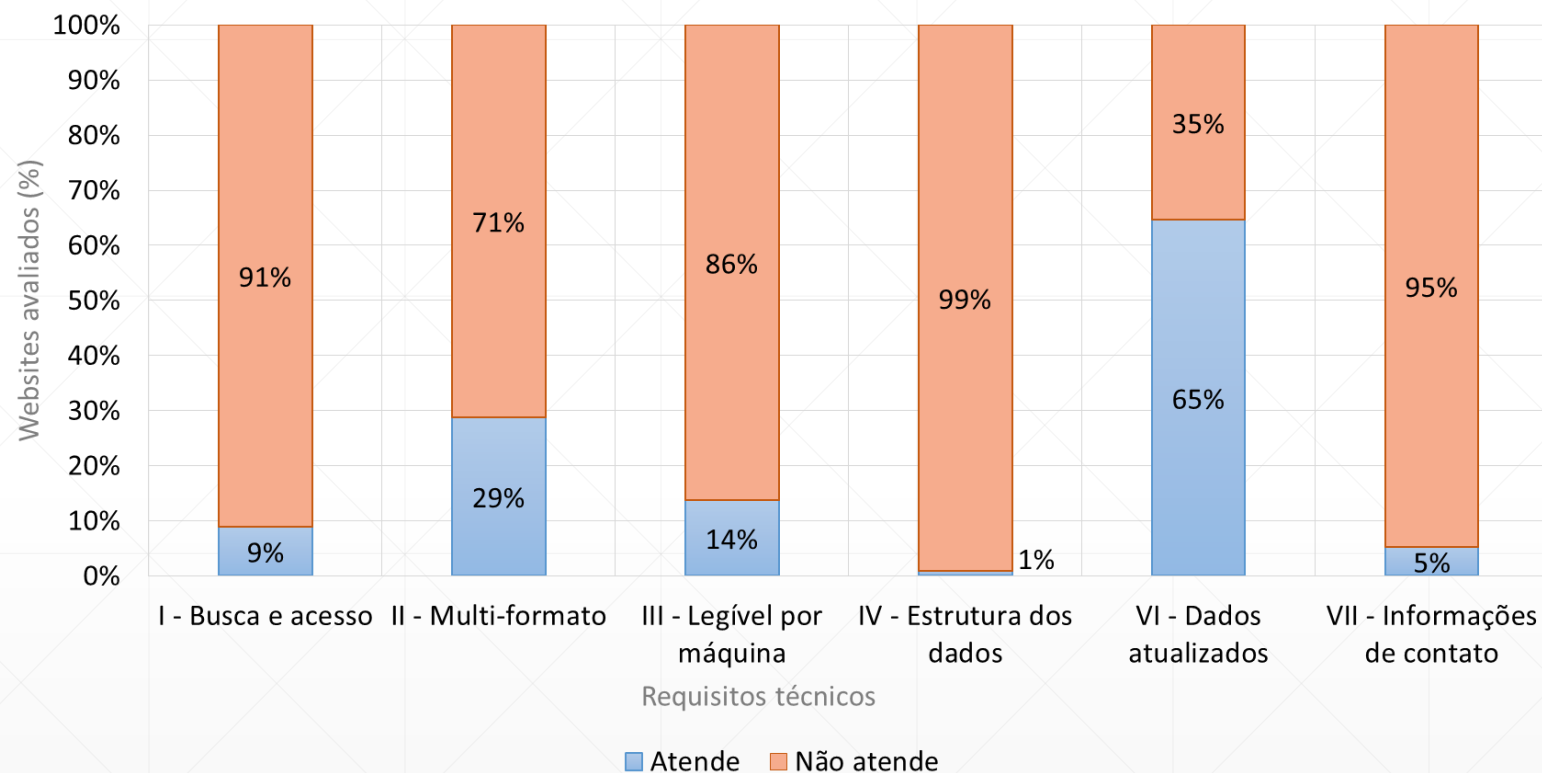
Retrato dos dados abertos no Brasil

Formatos preferenciais



Formatos preferenciais. Fonte: autores.

Requisitos técnicos da LAI



Atendimento dos requisitos técnicos que os websites precisam atender, conforme § 3º do Art. 8º da LAI

Fonte: autores.

Referências

BERNERS-LEE, T. **Linked Data. W3C Design Issues**. Disponível em: <<http://www.w3.org/DesignIssues/LinkedData.html>>. Acesso em: 13 dez. 2013.

CORRÊA, A. S.; CORRÊA, P. L. P.; DA SILVA, F. S. C. **Transparency Portals Versus Open Government Data: An Assessment of Openness in Brazilian Municipalities**. Proceedings of the 15th Annual International Conference on Digital Government Research. **Anais...**: dg.o '14. New York, NY, USA: ACM, 2014. Disponível em: <<http://doi.acm.org/10.1145/2612733.2612760>>. Acesso em: 10 out. 2014

DAVENPORT, T. H.; PRUSAK, L. **Working Knowledge: How Organizations Manage What They Know**. [s.l.] Harvard Business Press, 2013.

EAVES, D. **The Three Laws of Open Government Data**. Disponível em: <<http://eaves.ca/2009/09/30/three-law-of-open-government-data/>>. Acesso em: 27 nov. 2014.

OBAMA, B. **Transparency and Open Government | The White House**. Disponível em: <http://www.whitehouse.gov/the_press_office/TransparencyandOpenGovernment/>. Acesso em: 13 dez. 2013.

OPENGOVDATA. **Eight Principles of Open Government Data - OpenGovData.org**. Disponível em: <<http://opengovdata.org/>>. Acesso em: 27 nov. 2014.

The Open Definition - Open Definition - Defining Open in Open Data, Open Content and Open Knowledge. Disponível em: <<http://opendefinition.org/>>. Acesso em: 23 maio. 2016.

Obrigado!

andreiwid@usp.br

