



Tópicos Especiais em Web Semântica e suas aplicações

Dados Abertos

Situando...

- O Consórcio *World Wide Web* (W3C) é um consórcio internacional no qual organizações filiadas, uma equipe em tempo integral e o público trabalham juntos para desenvolver padrões para a web.
- Liderado por Tim Berners-Lee, o W3C tem como missão:
 - Conduzir a Web para que atinja todo seu potencial, desenvolvendo protocolos e diretrizes que garantam o seu crescimento a longo prazo.
- O escritório brasileiro – *W3C Brasil* –, é hospedado pelo Comitê Gestor da Internet (CGI.br) e pelo Núcleo de Informação e Coordenação do Ponto Br (NIC.br).

<http://www.w3c.br/Padroes/WebSemantica>

Introdução

■ Diversos dados são publicados na Web...



Percebe-se que ...

“Saber trabalhar com grandes quantidades de dados procedentes de diversas fontes e com diferentes formatos é uma das habilidades mais desejadas...”

■ O crescimento na Web é exponencial

- Mas, grande parte dos dados **não está disponível** ao público
- Ou **não está estruturada** o suficiente para facilitar sua compreensão e uso ...

Podemos pensar assim

**Publisher ou
Produtor de Dados**

Consumidor de Dados

O Governo produz...

Quem vai consumir ?

Os dados são
verdadeiramente úteis?

The screenshot shows a spreadsheet with columns for 'Ano' (Year) and 'Descrição' (Description). The data lists expenses for each year from 2000 to 2011, categorized by 'Presidência da República' and 'Secretaria de Administração'. The descriptions include 'Despesas Comensais', 'Despesas Correntes', and 'Despesas de Capital'.

Dados sobre Despesas do ano 2011

Como???

Consumidor

App₁

App₂

Um Problema real

■ Gestão de cidades

◦ Serviços públicos

- Diversas informações que precisam ser **compartilhadas, integradas e gerenciadas**

- Serviço de transporte público
- Dados provêm de diversas bases

- Horários, frota, rotas devem estar

Como resolver??? reais aos cidadãos...

- Isso facilitaria a vida do cidadão
- **Análises** poderiam ser feitas...
 - **Aplicativos** com esse fim também

Solução Estudada e Recomendada

■ **Ecosystema de Produção e Consumo de Dados Abertos**

- Com o objetivo de agilizar a **descoberta** de novos dados e conhecimentos
- **Agregar valor** a qualquer informação disponibilizada livremente na Web
- **Facilitar o desenvolvimento de aplicações** que irão consumir esses dados com a agregação da semântica ou do conhecimento

Dados Abertos

Dados abertos são dados que podem ser **livremente utilizados, reutilizados e redistribuídos** por qualquer pessoa – sujeitos, no máximo, à exigência de atribuição à fonte original e ao compartilhamento pelas mesmas licenças em que as informações foram apresentadas.

(Open Definition, 2014)

Dados Abertos

■ Três normas fundamentais:

◦ Disponibilidade e acesso

- Dados disponíveis como um todo e sob custo não maior que um custo razoável de reprodução
 - De forma conveniente e modificável
- Devem ser possíveis de ser **baixados pela Internet**

◦ Reuso e redistribuição

- Os dados devem ser fornecidos sob termos que permitam a reutilização e a redistribuição
 - Inclusive **a combinação** com outros conjuntos de dados

Dados Abertos

■ Três normas fundamentais:

◦ Participação universal

- Todos devem ser capazes de usar, reutilizar e redistribuir
 - Sem discriminação
 - Por exemplo, restrições de uso “não comercial” que impediriam o uso “comercial”, ou restrições de uso para certos fins (ex.: somente educativos) excluem determinados dados do conceito de “abertos”

Assim, é possível que diferentes organizações e sistemas possam trabalhar de forma colaborativa




The screenshot shows the data.gov website. At the top, there is a navigation bar with links: DATA, TOPICS, IMPACT, APPLICATIONS, DEVELOPERS, and CONTACT. Below this is a blue banner with the text "The home of the U.S. Government's open data" and a subtext: "Here you will find data, tools, and resources to conduct research, develop web and mobile applications, design data visualizations, and [more](#)." Below the banner, there is a list of bullet points in blue and green text.

■ 2007: pensadores e ativistas se reuniram e definiram o conceito de dados abertos governamentais ou públicos

- Pilares: abertura, participação e colaboração

■ 2008: Memorando de Barack Obama sobre Transparência e Dados Governamentais

- Criação do **data.gov**



The screenshot shows the dados.gov.br website. At the top, there is a navigation bar with links: Acesso à informação, Participe, Serviços, Legislação, and Canais. Below this is a green banner with the text "dados.gov.br" and "PORTAL BRASILEIRO DE DADOS ABERTOS". Below the banner, there is a search bar with the text "Pesquisa ..." and a button "PESQUISAR". Below the search bar, there is a table with the title "Publicações mais recentes".

Conjunto de dados	Autor	Quando
Áreas Embargadas	Edson Eyji Sano - CSR	17 Abr

■ Os dados abertos governamentais fazem parte da política de acesso à informação do governo federal (INDA);

■ 2012: Lei de Acesso à Informação

- Necessidade de disponibilizar dados governamentais em **formato aberto**

Lei de Acesso à Informação

LAI: A Lei de Acesso à Informação

A Lei nº 12.527/2011 regulamenta o direito constitucional de acesso às informações públicas. Essa norma entrou em vigor em 16 de maio de 2012 e criou mecanismos que possibilitam, a qualquer pessoa, física ou jurídica, sem necessidade de apresentar motivo, o recebimento de informações públicas dos órgãos e entidades.

A Lei vale para os três Poderes da União, Estados, Distrito Federal e Municípios, inclusive aos Tribunais de Conta e Ministério Público. Entidades privadas sem fins lucrativos também são obrigadas a dar publicidade a informações referentes ao recebimento e à destinação dos recursos públicos por elas recebidos.

Dados Abertos e Dados Abertos Governamentais...



DADOS ABERTOS

Você entregaria parte de seu dinheiro para um administrador que não prestasse contas a você? Os governos são produtores e guardiões de muitos dados que devem estar publicados e acessíveis a todos os cidadãos.

O que são Dados Abertos?

São a publicação e disseminação de informações na internet, compartilhadas em formatos abertos, legíveis por máquinas, e que possam ser livremente reutilizadas de forma automatizada pela sociedade.

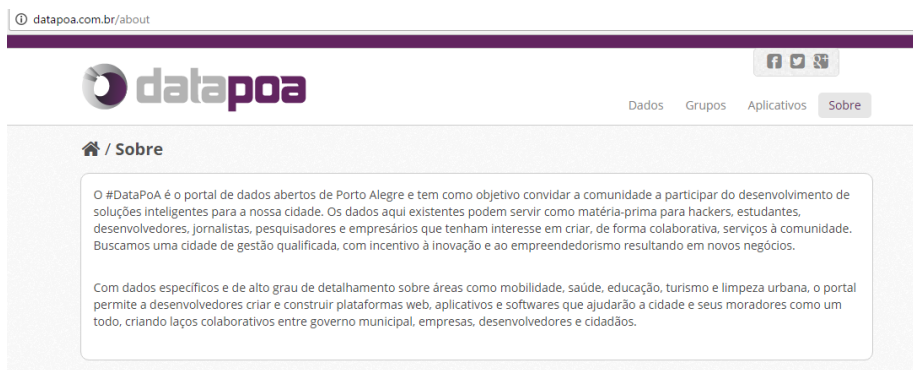
Dados Abertos Governamentais

São os produzidos pelos governos e que devem ser colocados à disposição de qualquer cidadão e para qualquer fim.

Para que servem?

Servem para que os cidadãos e as organizações da sociedade possam reutilizá-los com o intuito de verificar, esclarecer, fiscalizar e acompanhar de acordo com seus interesses. Por exemplo: criar um aplicativo para celular que mostre onde estão as escolas públicas de uma determinada região, como as vagas estão distribuídas e onde há maior demanda por vagas; ou como está sendo gasto o dinheiro público; ou ainda o nível de segurança pública por município ou por bairro.

Dados Abertos - Exemplo



datapoa.com.br/about

datapoa

Dados Grupos Aplicativos Sobre

/ Sobre

O #DataPoA é o portal de dados abertos de Porto Alegre e tem como objetivo convidar a comunidade a participar do desenvolvimento de soluções inteligentes para a nossa cidade. Os dados aqui existentes podem servir como matéria-prima para hackers, estudantes, desenvolvedores, jornalistas, pesquisadores e empresários que tenham interesse em criar, de forma colaborativa, serviços à comunidade. Buscamos uma cidade de gestão qualificada, com incentivo à inovação e ao empreendedorismo resultando em novos negócios.

Com dados específicos e de alto grau de detalhamento sobre áreas como mobilidade, saúde, educação, turismo e limpeza urbana, o portal permite a desenvolvedores criar e construir plataformas web, aplicativos e softwares que ajudarão a cidade e seus moradores como um todo, criando laços colaborativos entre governo municipal, empresas, desenvolvedores e cidadãos.

<http://datapoa.com.br/>

Dados Abertos – notícia recente



UFRN UNIVERSIDADE FEDERAL DO RIO GRANDE DO NORTE

Acesso à Informação BRASIL

Institucional Ensino Pesquisa Extensão Serviços Documentos Sala de Imprensa Eventos Fale Conosco

Início » Notícias da UFRN » 20400039

UFRN é primeira universidade a aprovar Plano de Dados Abertos 30/5at/2016 às 17:14

A Universidade Federal do Rio Grande do Norte (UFRN) é pioneira na aprovação do Plano de Dados Abertos (PDA), que orienta as ações de implementação e promoção de abertura de dados da instituição. O documento foi aprovado pelo Conselho de Administração (Consad) da UFRN em sessão extraordinária na última quinta-feira, 29, e será colocado em prática para disponibilizar dados acessíveis ao público em um Portal de Dados Abertos próprio da universidade.

Durante apresentação ao Consad, o diretor de sistemas da Superintendência de Informática (Sinfo) e coordenador da comissão elaboradora do PDA, Gibeon Soares de Aquino Júnior, expôs os próximos passos do trabalho. Em outubro está prevista a criação do Portal de Dados Abertos da UFRN, que será o primeiro do Brasil entre as universidades. Até o final de 2016 espera-se que seja publicado o primeiro conjunto de dados abertos da instituição.

"Esta iniciativa é de extrema importância por permitir ao cidadão acompanhar o que acontece nas instituições públicas", citou o vice-reitor da UFRN, José Daniel Diniz Melo. O PDA atende à Política de Dados Abertos do Poder Executivo Federal, instituída pelo Decreto

ENCONTRE NA UFRN

Palavra-chave

OK

ESCOLHA SUA ÁREA

Público externo

Alunos

Docentes

Técnicos Adm.

<http://dados.ufrn.br/>

Dados Abertos Governamentais - leis

- O especialista em políticas públicas e ativista dos dados abertos David Eaves propôs as seguintes “leis”:
 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.
- As leis foram propostas para os Dados Abertos Governamentais, mas se aplicam aos Dados Abertos de forma geral

Dados Abertos Governamentais - Princípios

1. **Completos**
 - Todos os dados públicos são disponibilizados. Dados são informações eletronicamente gravadas, incluindo, mas não se limitando a, documentos, bancos de dados, transcrições e gravações audiovisuais.
2. **Primários**
 - Os dados são publicados na forma coletada na fonte, com a mais fina granularidade possível, e não de forma agregada ou transformada.
3. **Atuais**
 - Os dados são disponibilizados o quão rapidamente seja necessário para preservar o seu valor.
4. **Acessíveis**
 - Os dados são disponibilizados para o público mais amplo possível e para os propósitos mais variados possíveis.

Dados Abertos Governamentais - Princípios

5. Processáveis por máquina

- Os dados são razoavelmente estruturados para possibilitar o seu processamento automatizado.

6. Acesso não discriminatório

- Os dados estão disponíveis a todos, sem que seja necessária identificação ou registro.

7. Formatos não proprietários

- Os dados estão disponíveis em um formato sobre o qual nenhum ente tenha controle exclusivo.

8. Livres de licenças

- Os dados não estão sujeitos a regulações de direitos autorais, marcas, patentes ou segredo industrial. Restrições razoáveis de privacidade, segurança e controle de acesso podem ser permitidas na forma regulada por estatutos.

Dados Abertos Governamentais



Efeitos dos dados abertos governamentais sobre as políticas públicas

Inclusão. Fornecer dados em formatos padronizados abertos e acessíveis permite que qualquer usuário utilize qualquer ferramenta de software para adaptá-los às suas necessidades.

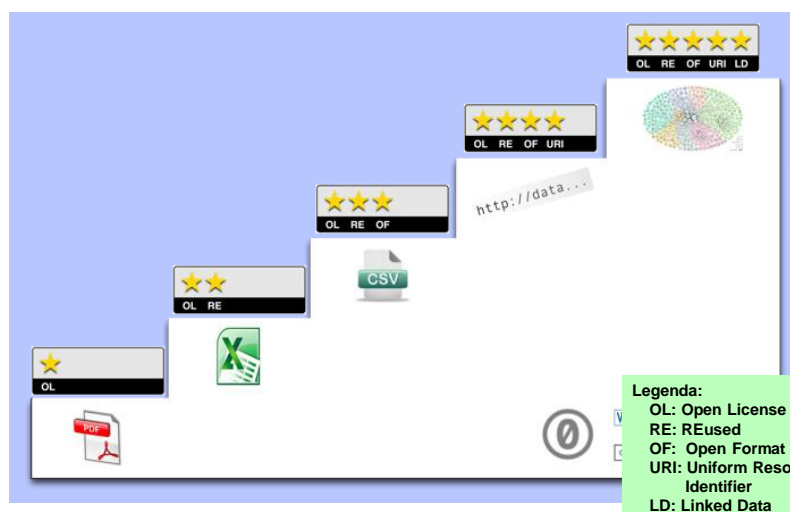
Transparência. Informações do setor público abertas e acessíveis melhoram a transparência, pois as partes interessadas podem usá-las da maneira mais adequada ao seu propósito, obtendo uma ideia melhor do governo.

Responsabilidade. Os conjuntos apropriados de dados abertos, devidamente associados, podem oferecer vários pontos de vista sobre o desempenho do governo no atingimento de suas metas em políticas públicas.

Portais de Transparência e Dados Abertos



Dados Abertos - Níveis



Fonte: (Berners-Lee, 2008) <http://5stardata.info/en/>

Esquema de 5 ★

- Cada nova estrela alcançada torna os dados progressivamente mais úteis, mais fáceis pras pessoas e aplicações utilizarem
 - ★☆☆☆☆ Os dados estão disponíveis na Web, independente de formato, **sob uma licença aberta**
 - Ex: um documento PDF sob uma licença aberta
 - ★★☆☆☆ A condição anterior + disponível como **dados estruturados legíveis por máquina**
 - **Ex:** Arquivo Excel ao invés de uma imagem escaneada de uma tabela
 - ★★★☆☆ Todas as anteriores + utilizar um **formato não proprietário**
 - Ex: um arquivo CSV ao invés de um Excel
 - ★★★★☆ Todas as anteriores + **utilizar URIs bem desenhadas** para identificar os recursos, então as pessoas podem referenciá-los
 - ★★★★★ Todas as anteriores + **ligar os dados com dados de outros**

Ciclo de Vida dos Dados



Figura 1.1 – Ciclo de Vida da Informação. Fonte: adaptado de Floridi (2010).

Dados Abertos – Como?

■ Possuem formatos variados

- Semiestruturados, es...
estruturados...

Metadados descrevendo
os conjuntos de dados
podem ajudar...

■ Nem sempre seguem padrões...

■ Nem sempre se consegue entender sua semântica...

Semântica proveniente
de vocabulários pode
ajudar...

Recomendação do
W3C!

Recomendação do
W3C!

Exemplo – Conjunto de Dados sobre Paradas de ônibus (stops-2015-05-05)

:stops-2015-05-05

```
a dcat:Dataset ;
dct:title "Bus stops of MyCity" ;
dcat:keyword "transport","mobility","bus" ;
dct:issued "2015-05-05"^^xsd:date ;
dcat:contactPoint
<http://data.mycity.example.com/transport/contact> ;
dct:temporal <http://reference.data.gov.uk/id/year/2015> ;
dct:publisher :transport-agency-mycity ;
dcat:theme :mobility ;
dcat:distribution :stops-2015-05-05.csv ;
```

Passos para a Abertura dos Dados

1. Escolha os conjuntos de dados que planeja abrir.
2. Disponibilize os dados estruturados em um formato aberto reutilizável.
 - Aplique uma licença livre
 - Deve-se também considerar métodos alternativos de disponibilização como uma API (*application programming interface*)
3. Torne-os mais fáceis de encontrar
 - Publique os dados na rede ou organize um repositório/catálogo central para listar todos os conjuntos de dados.

Catálogos/Portais de Dados

- Data.Gov (<http://data.gov/>)
- Data.Gov.Uk (<http://data.gov.uk/>)
- <http://dados.recife.pe.gov.br/>
- Fornecem links para feeds ou APIs e permitem que os usuários indiquem os conjuntos de dados mais relevantes para as comunidades, num processo de gestão colaborativa da informação pública.

Exemplos de aplicativos

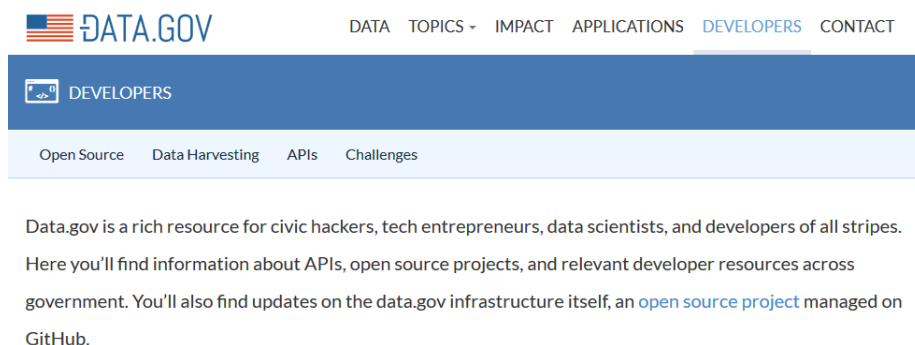
■ <http://alagamentos.topical.com.br/>

■ <http://cmsp.topical.com.br/>

“Para criar esses serviços, desenvolvedores tiveram de “**raspar**” informações de sites públicos. **É como se a sociedade estivesse fazendo todo o trabalho sozinha.** “

“Quando os dados governamentais estiverem abertos, convidando as pessoas a os reutilizarem, será possível criar mais.”

O que temos a ver com isso???



The screenshot shows the Data.gov website. At the top, there is a navigation bar with links: DATA, TOPICS, IMPACT, APPLICATIONS, DEVELOPERS (highlighted), and CONTACT. Below this is a blue header with the Data.gov logo and a sub-header 'DEVELOPERS'. Underneath, there are four tabs: Open Source, Data Harvesting, APIs, and Challenges. The main content area contains the following text:

Data.gov is a rich resource for civic hackers, tech entrepreneurs, data scientists, and developers of all stripes. Here you'll find information about APIs, open source projects, and relevant developer resources across government. You'll also find updates on the data.gov infrastructure itself, an [open source project](#) managed on GitHub.

Significado de Hackathon

Compartilhar 37 G+

O que é Hackathon:

Hackathon significa maratona de programação. O termo resulta de uma combinação das palavras inglesas "hack" (programar de forma excepcional) e "marathon" (maratona).

■ Competições em todo o mundo

- Uma série de conjuntos de dados e problemas para os quais os programadores têm um período de tempo para desenvolver aplicativos.
- O melhor aplicativo construído com base nos dados é, então, premiado.
- Vários governos e organizações de diversos países (como Reino Unido, Estados Unidos, Noruega, Finlândia, etc) vêm realizando

■ Show Us a Better Way (<http://showusabetterway.co.uk>)

Mais links

- <http://br.okfn.org/dados-abertos/>
- http://opendatahandbook.org/guide/pt_BR/