

CKAN Open Source Data Portal Software

I Ciclo de Oficinas sobre Ferramentas de Gestão de Dados
Grupo de Estudo, Pesquisa e Extensão em Big Data
agosto/2016



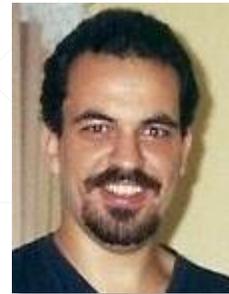
Sobre os autores



Andreiwid Sheffer Corrêa

- Graduado em Ciência da Computação e mestrado em Gestão de Redes e Serviços de Telecomunicações
- Doutorando em Engenharia de Computação pela Poli/USP com pesquisa na área de dados abertos
- 17 anos de experiência no desenvolvimento de software, sendo metade na área de governo
- Professor efetivo do Instituto Federal de São Paulo, no Câmpus Campinas

andreiwid@gmail.com



Ângelo Luís Rocatto Júnior

- Graduado em Engenharia de Computação pela Unicamp
- Experiência de 8 anos como analista de desenvolvimento
- Trabalha com o CKAN há 2 anos
- Atua no desenvolvimento de soluções de software na área de governo
- Tem especialidade na área de dados abertos, geoprocessamento e redes

angelo.roc@gmail.com

Créditos e apoio institucional

- Parte deste material foi baseado no Minicurso de CKAN ministrado por Augusto Herrmann no LOD Brasil.
- Material original sob a licença CC BY disponível em <http://www.inf.ufsc.br/~jose.todesco/LODBrasil/Minicurso/CKAN.pdf>
- Fotos CC BY 2.0 obtidas de <http://photopin.com/>



Conteúdo

- Dados Abertos Governamentais e seus princípios.
 - Apresentando o CKAN.
 - Utilização da ferramenta para gestão de datasets.
 - Hands-on: publicando e acessando dados
 - Caminho das pedras para implementação de uma instância CKAN
-

Dados Abertos Governamentais e seus princípios

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



Definição de Dado aberto

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

Open Definition

The Definition

Conformant Licenses

Participate

News



The Open Definition

The [Open Definition](#) sets out principles that define “openness” in relation to **data and content**.

It makes **precise** the meaning of “open” in the terms “**open data**” and “**open content**” and thereby ensures **quality** and encourages **compatibility** between different pools of open material.

It can be summed up in the statement that:

*“Open means **anyone** can **freely access, use, modify, and share** for any purpose (subject, at most, to requirements that preserve provenance and openness).”*

Put most succinctly:

*“Open data and content can be **freely used, modified, and shared** by anyone for any purpose”*

[Read the full Open Definition »](#)

THE OPEN DEFINITION IN YOUR LANGUAGE

العربية | Беларуская | Български | Catala | Czech | Dansk | Deutsch | Eesti | Ελληνικά | English | Español | Euskara | Suomi | Français | Galego | ગુજરાતી | हिन्दी | Croatian | Magyar | Indonesian | Íslenska | Italiano | 日本語 | ಕನ್ನಡ | 한국어 | Македонски | jазик | नेपाली | Norsk (bokmål) | Poliszczyna | Português Brasileiro | Português | Руски | Shqip | Српски | Svenska | తెలుగు | Türkçe | Українська | 简化中国 | 中文

Open Definition is a project of Open Knowledge - Source Code

[OPEN SERVICE](#) [OPEN CONTENT](#) - Content licensed under a [CC Attribution 4.0 International License](#)

[Terms of use](#) | [Privacy policy](#) | [IP Policy](#)

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

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.

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

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

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



Participação universal

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

Reutilização e redistribuição

- Termos de utilização
 - Remixagem
-

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

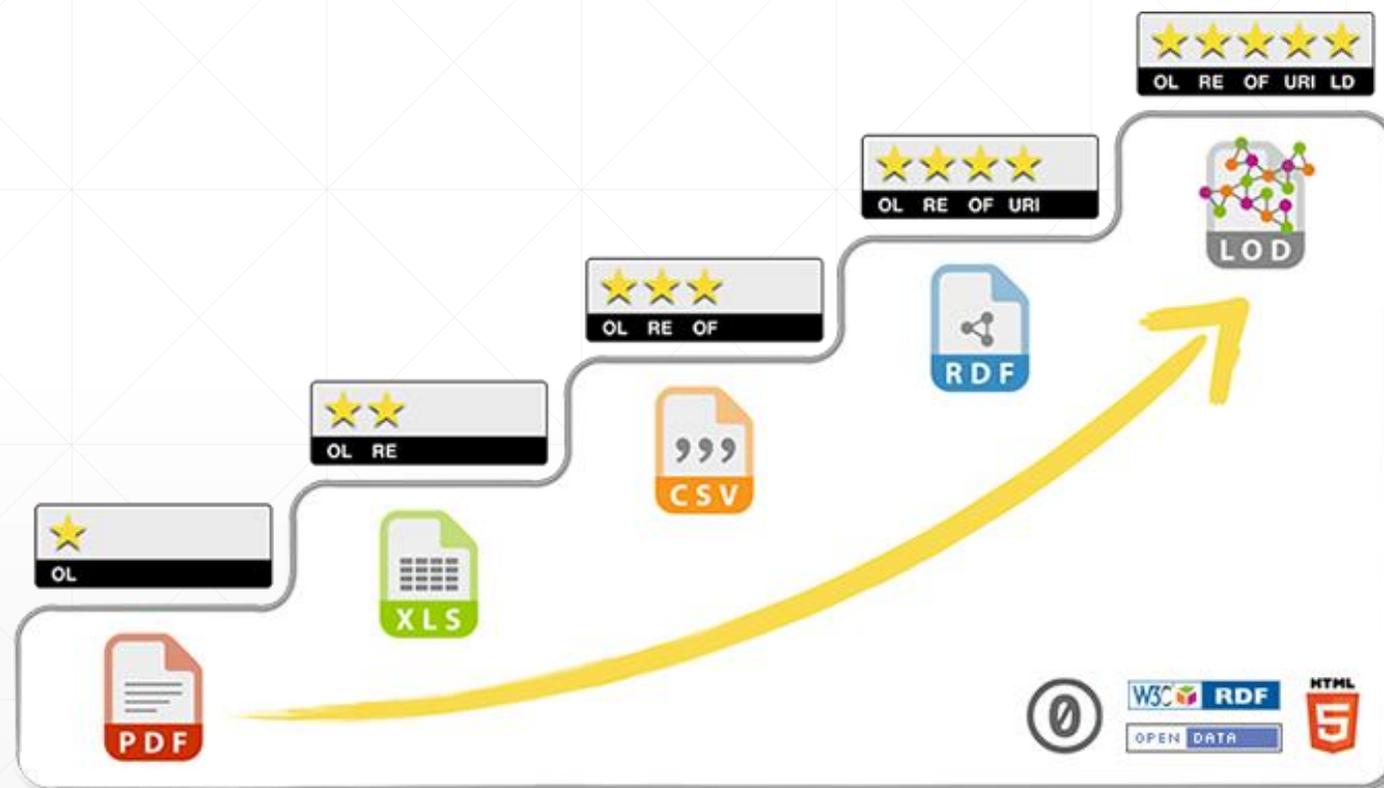


Imagem extraída de <http://5stardata.info/en/>

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

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

- Processos de trabalho
- Formato

Apresentação

- Busca e descoberta
- Metadados
- Múltiplos formatos

Disseminação

- API
- Visualização avançada

Catálogo

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"
  }
}
```

Apresentando o CKAN

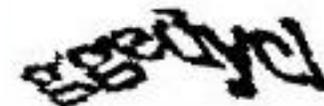
O que é

- CKAN = Comprehensive Knowledge Archive Network
 - Software livre para catalogação de dados abertos
 - Desenvolvido pela Open Knowledge (<http://okfn.org>)
 - Open-source (<https://github.com/ckan/ckan>)
 - Em desenvolvimento desde 2007
 - Disponível em mais de 50 idiomas (inclusive PT-BR)
-

Para que serve

- Guardar dados e metadados
 - Pode também apenas linkar dados diretamente
- Permitir a busca fácil de dados
- Não há captcha
- Possibilita implementar os princípios de OGD
- Trabalha com formatos abertos
 - CSV, JSON, XML, RDF
- Mas também pode incluir outros recursos (PDF, HTML, etc)

Please prove you're not a robot



Type the two words:

106 is#4life



Algumas instâncias ao redor do mundo



Data.gov



data.gov.au



Data.gov.uk



datos.gob.mx



Data.Rio



Dados Recife



The Brazilian Senate

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

Utilização da ferramenta para gestão de datasets

Principais funcionalidades

- Catalogação pela interface web, pela API ou por ferramentas de importação
 - Pesquisa em todos os metadados
 - Busca avançada
 - organização, etiqueta, formato, licença
 - Organização baseada em *datasets* (conjuntos de dados) e recursos
 - Recursos relacionados ou similares
 - Mesmos dados em diferentes formatos, períodos distintos, etc.
-

Pré-visualização dos dados

- Pré-visualização do recurso como tabela, gráfico, mapa, etc.
- Ajustável - permite reordenação por atributo, escolha de colunas para eixos do gráfico, etc.
- Usa a biblioteca `recline.js`

Contratos em CSV

Baixar

URL: <http://compras.dados.gov.br/contratos/v1/contratos.csv>

Fornece uma lista das informações relacionadas aos registros de Contratos.

URL: <http://compras.dados.gov.br/contratos/v1/contratos.csv>

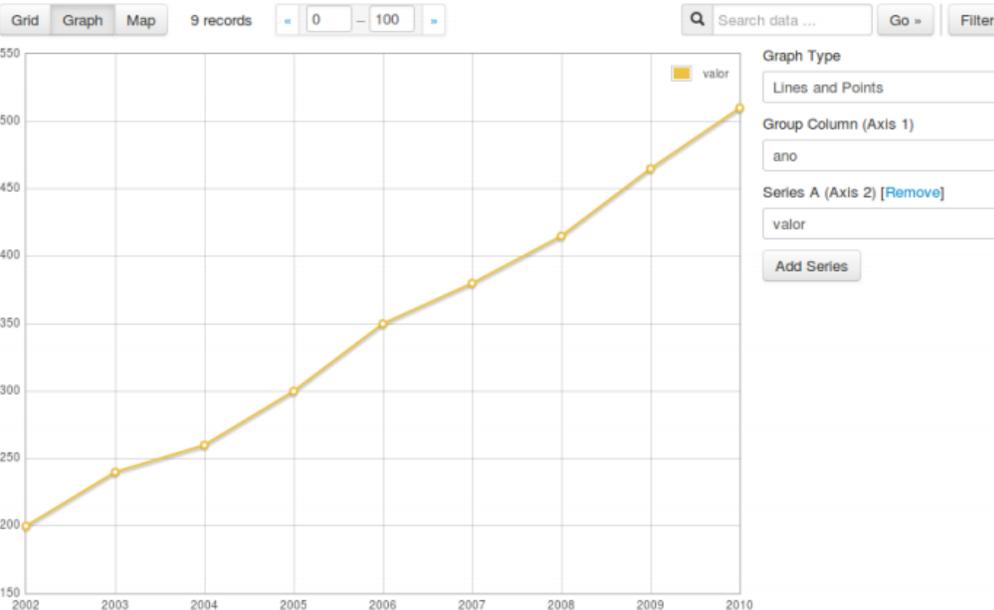
Exemplo: <http://compras.dados.gov.br/contratos/v1/contratos.csv?uasg=153229>

Grid Graph Map 500 records 0 100 Search data ... Go Filters

Identifica...	UASG	Modalida...	Número ...	Tipo de ...	Licitação...	Origem d...	Número	Objeto	Número ...	Número ...
38093150...	380931: ...	05: PRE...	23/2013	50: CONT...	PREGÃO...	SISPP	1/1014	Prestaçã...	0	46201006...
15303254...	153032: ...	03: CON...	1/1978	54: CON...	CONCOR...	SISPP	1/1978	Concessã...	3	0001/78
11460650...	114606: ...	06: DISP...	1/1999	50: CONT...	DISPENS...	SISPP	1/1985	Contrataç...	1	PA/03615...
15404753...	154047: ...	06: DISP...	1/1985	53: ARRE...	DISPENS...	SISPP	1/1985	locacao d...	0	23110000...
15404753...	154047: ...	06: DISP...	1/1986	53: ARRE...	DISPENS...	SISPP	1/1986	locacao d...	0	00000000...
38004450...	380044: ...	06: DISP...	1/1987	50: CONT...	DISPENS...	SISPP	1/1987	Contrato ...	0	35183001...
17011650...	170116: ...	07: INEXI...	1/1988	50: CONT...	INEXIGIB...	SISPP	1/1988	Contrato ...	1	10720000...
51208450...	512084: ...	06: DISP...	1/1990	50: CONT...	DISPENS...	SISPP	1/1990	LOCAÇA...	0	35314/71...
51211150...	512111: ...	06: DISP...	1/1990	50: CONT...	DISPENS...	SISPP	1/1990	Locação ...	1	35328022...
17011950...	170119: ...	06: DISP...	1/1988	50: CONT...	DISPENS...	SISPP	1/1990	Locação ...	2	10720000...
51206050...	512060: ...	07: INEXI...	1/1991	50: CONT...	INEXIGIB...	SISPP	1/1991	Renovaçã...	0	35301018...
51209950...	512099: ...	06: DISP...	13/1991	50: CONT...	DISPENS...	SISPP	1/1991	Locação ...	0	35301016...
17011950...	170119: ...	06: DISP...	1/1991	50: CONT...	DISPENS...	SISPP	1/1991	Locação ...	3	13009000...

URL: <http://pgi.gov.br/pgiadmin/csv/serie/1567.csv>

Valor nominal do salário mínimo no ano



Dados geoespaciais

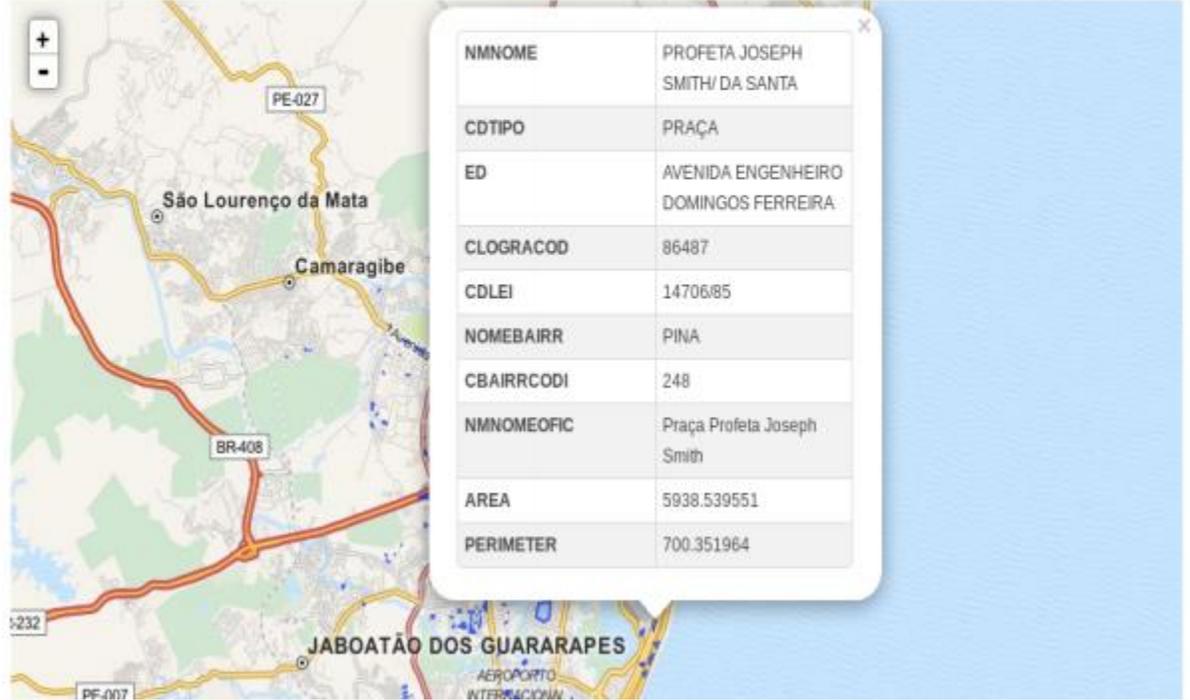
- Por meio da extensão ckanext-spatial
- Visualizar o conteúdo dentro do mapa (p. ex.: delimitações de áreas de praças e parques)
- Pesquisar por dados que se situem dentro de um perímetro delimitado na pesquisa

Organizações / Secretaria de Infraestrutura ... / Parques e praças / Parques e Praças

Parques e Praças

URL: <http://dados.recife.pe.gov.br/storage/f/2014-05-23T19%3A41%3A50.416Z/parquespracas.geojson>

Layer (nível) de geoprocessamento contendo a localização dos parques e praças do Recife.



The image shows a web interface for a geospatial data layer. It features a map of Recife, Brazil, with a popup window displaying details for a specific square. The popup window contains a table with the following data:

NM NOME	PROFETA JOSEPH SMITH/ DA SANTA
CD TIPO	PRAÇA
ED	AVENIDA ENGENHEIRO DOMINGOS FERREIRA
CLOGRACOD	86487
CD LEI	14706/85
NOME BAIRR	PINA
CB AIRRCODI	248
NM NOME OFIC	Praça Profeta Joseph Smith
AREA	5938.539551
PERIMETER	700.351964

Histórico de atualização

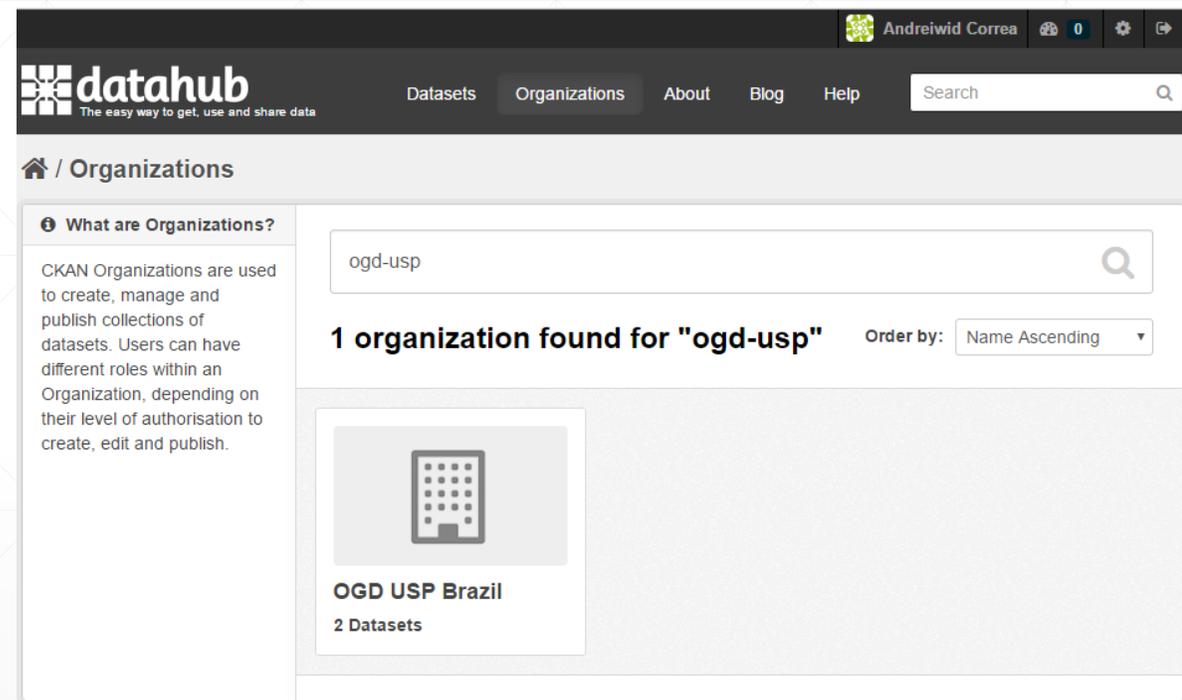
- É possível saber quando e quem atualizou



The screenshot displays a web interface for a data portal. The main heading is 'Orçamento Federal'. Below it, there are navigation tabs: 'Conjunto de dados', 'Grupos', 'Fluxo de Atividades', and 'Relacionado'. The 'Fluxo de Atividades' tab is selected, showing a vertical timeline of updates. Each update entry includes a user profile icon, a green gear icon, the text 'luis.araujo atualizou o recurso', the specific year of the budget (e.g., 'Orçamento Federal 2014'), and the text 'do conjunto de dados Orçamento Federal há 1 mês'. On the left side, there is a sidebar with sections for 'Seguidores' (0), 'Social' (with links for Google+, Twitter, and Facebook), and 'Licença' (Licença Aberta para Bases de Dados (ODbL) do Open Data Commons).

Organizações no CKAN

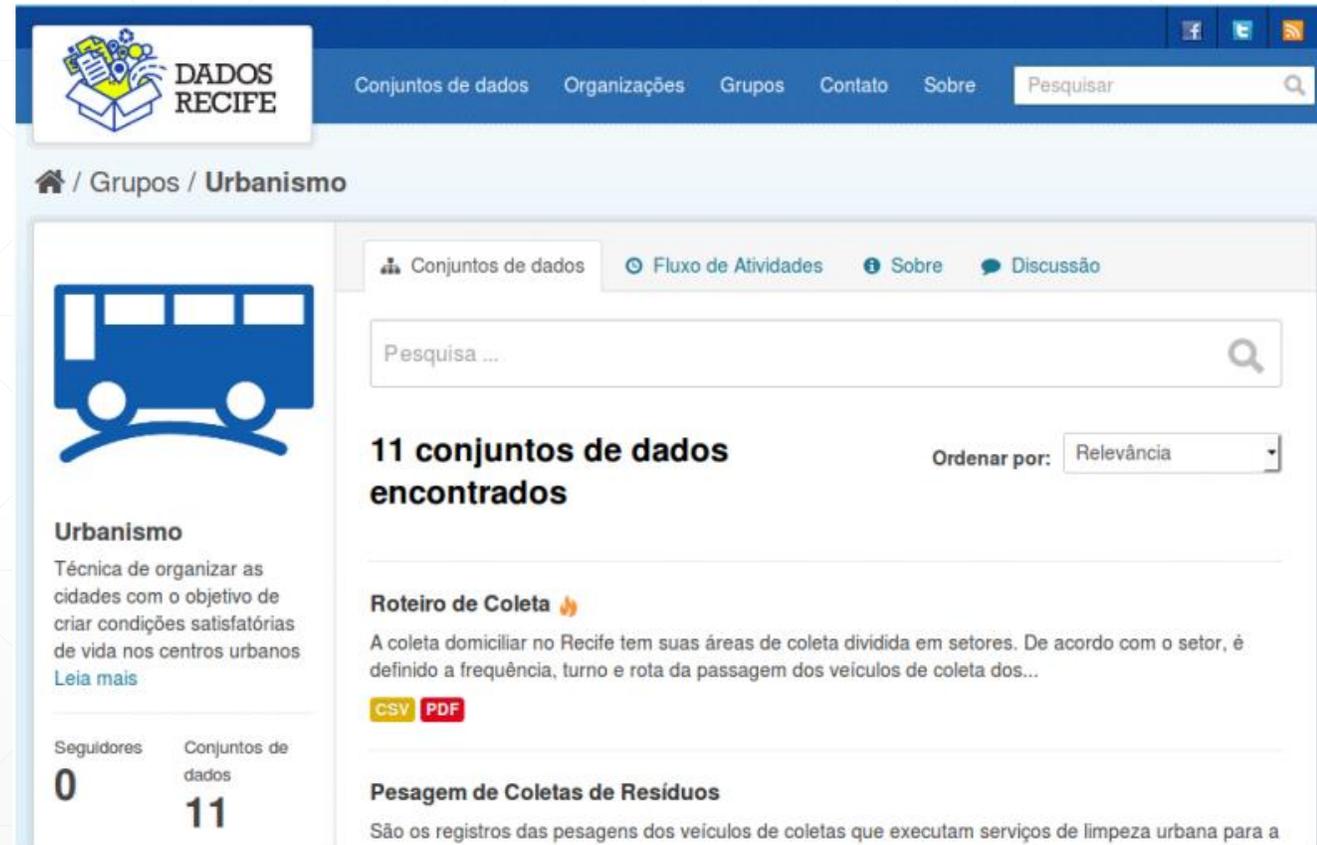
- Possibilita que cada instituição tenha gestão sobre os seus próprios dados no catálogo



The screenshot shows the DataHub website interface. At the top, there is a navigation bar with the DataHub logo and the tagline "The easy way to get, use and share data". The navigation menu includes "Datasets", "Organizations", "About", "Blog", and "Help". A search bar is located on the right side of the navigation bar. Below the navigation bar, the page title is "Organizations". On the left side, there is a sidebar with a section titled "What are Organizations?" containing a brief explanation of CKAN Organizations. The main content area shows search results for "ogd-usp". A search bar at the top of the results area contains the text "ogd-usp". Below the search bar, it says "1 organization found for 'ogd-usp'" and "Order by: Name Ascending". The search results list one organization: "OGD USP Brazil" with "2 Datasets".

Grupos no CKAN

- Possibilita classificações temáticas dos dados



The screenshot displays the CKAN website interface for the 'DADOS RECIFE' organization. The main navigation bar includes links for 'Conjuntos de dados', 'Organizações', 'Grupos', 'Contato', and 'Sobre', along with a search bar. The current page is titled 'Urbanismo' and features a blue bus icon. The page content includes a search bar, a list of 11 data sets, and a dropdown menu for sorting by 'Relevância'. The first data set listed is 'Roteiro de Coleta', which is described as a collection of data for household waste collection in Recife, divided into sectors. The second data set is 'Pesagem de Coletas de Resíduos', which records the weight of waste collection vehicles performing urban cleaning services.

DADOS RECIFE

Conjuntos de dados Organizações Grupos Contato Sobre Pesquisar

/ Grupos / Urbanismo

Conjuntos de dados Fluxo de Atividades Sobre Discussão

Pesquisa ...

11 conjuntos de dados encontrados Ordenar por: Relevância

Urbanismo
Técnica de organizar as cidades com o objetivo de criar condições satisfatórias de vida nos centros urbanos
[Leia mais](#)

Seguidores: 0 Conjuntos de dados: 11

Roteiro de Coleta 🔥
A coleta domiciliar no Recife tem suas áreas de coleta dividida em setores. De acordo com o setor, é definido a frequência, turno e rota da passagem dos veículos de coleta dos...
CSV PDF

Pesagem de Coletas de Resíduos
São os registros das pesagens dos veículos de coletas que executam serviços de limpeza urbana para a

Acesso via API

- Comunica-se por requisições http (estilo RPC)
 - Processa e retorna metadados no formato JSON
 - Possibilita fazer tudo o que se faria pela interface gráfica (ex.: pesquisar, consultar, etc.)
 - Utiliza chave de acesso do usuário para acesso de escrita para criar, atualizar ou excluir conjuntos de dados
 - Útil para processar grandes volumes de dados (ex.: catalogação em lote de novos dados, correções em massa, etc.)
-

Mais funcionalidades

- Temas personalizáveis
 - FileStore (upload de arquivos)
 - Federável/Harvesting (extensão ckanexharvest)
 - Feedback dos usuários
-

Comece a testar

- demo.ckan.org
 - Livre para experimentar, catalogar dados e conhecer o CKAN
 - Conteúdo é apagado periodicamente
 - [Datahub.io](https://datahub.io)
 - Instância CKAN em produção
 - Patrocinada pela comunidade Open Knowledge
 - Existe um processo para criar organizações
-

Outros softwares de catálogo/repositório

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

OpenSpending

- Visualização avançada
 - Integração com *datasets* em instâncias CKAN
-

Hands-on: publicando e acessando dados

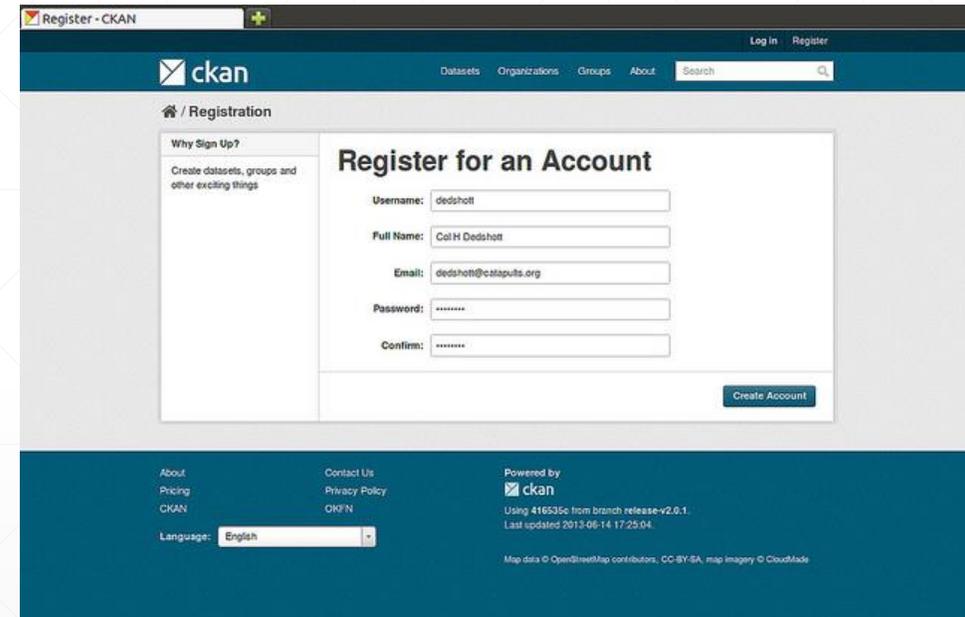
Preparativos para o hands-on

- Baixe os arquivos de apoio
 - Link <https://1drv.ms/f/s!AnpqtVDedGRjhd9QaHJWycVQ2MpVmw>
 - Arquivos DEINFO_ED_CENSOESCOLAR_2015 (CSV e XLS)
 - Amostra do censo escolar da cidade de SP para 2015
 - Fonte: <http://geosampa.prefeitura.sp.gov.br/>
 - Arquivo cotaparlamentar01-05_2016_resumo (CSV e XLS)
 - Gastos mensais com cotas parlamentares dos deputados federais nos meses de 01 a 08 (parcial) do ano de 2016
 - Fonte: <http://www2.camara.leg.br/transparencia/dados-abertos>
-

Registro e acesso

Executar os seguintes passos:

- Acesse demo.ckan.org
- Clique em “Register”
- Preencha as informações solicitadas para registro (tela ao lado)
- Agora localize o link “Log in” e faça o acesso com seu usuário

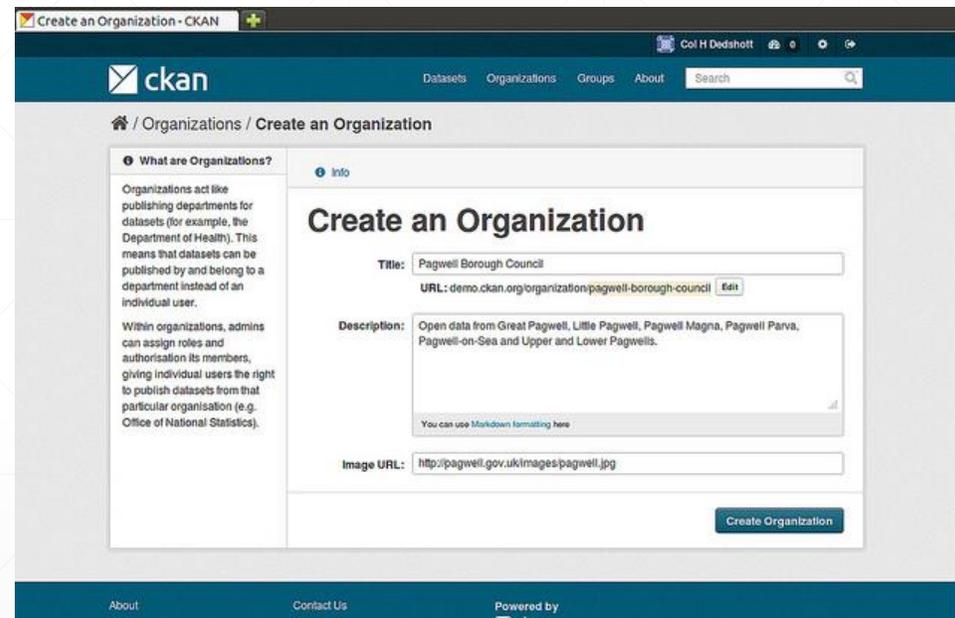


The screenshot shows the CKAN registration page. The browser title is "Register - CKAN". The page has a dark blue header with the CKAN logo and navigation links: "Datasets", "Organizations", "Groups", "About", and a search bar. The main content area is titled "Registration" and contains two columns. The left column is titled "Why Sign Up?" and contains the text "Create datasets, groups and other exciting things". The right column is titled "Register for an Account" and contains a registration form with the following fields: "Username:" (filled with "dedshott"), "Full Name:" (filled with "Col H Dedshott"), "Email:" (filled with "dedshott@catapults.org"), "Password:" (filled with "*****"), and "Confirm:" (filled with "*****"). A "Create Account" button is located at the bottom right of the form. The footer is dark blue and contains links for "About", "Pricing", "CKAN", "Contact Us", "Privacy Policy", "CKAN", and "Powered by CKAN". It also includes version information: "Using 416535e from branch release-v2.0.1. Last updated 2013-06-14 17:25:04." and a map data attribution: "Map data © OpenStreetMap contributors, CC-BY-SA, map imagery © CloudMade".

Criar sua organização

Executar os seguintes passos:

- Clique no link “Organizations” no topo de qualquer página
- Preencha com as informações solicitadas



The screenshot shows the CKAN 'Create an Organization' form. The browser title is 'Create an Organization - CKAN'. The CKAN logo is in the top left, and navigation links for 'Datasets', 'Organizations', 'Groups', and 'About' are in the top right. The breadcrumb trail is '/ Organizations / Create an Organization'. On the left, there is a sidebar with a 'What are Organizations?' section. The main content area has an 'info' icon and the title 'Create an Organization'. The form fields are: 'Title' (Pagwell Borough Council), 'URL' (demo.ckan.org/organization/pagwell-borough-council), 'Description' (Open data from Great Pagwell, Little Pagwell, Pagwell Magna, Pagwell Parva, Pagwell-on-Sea and Upper and Lower Pagwells), and 'Image URL' (http://pagwell.gov.uk/images/pagwell.jpg). A 'Create Organization' button is at the bottom right.

What are Organizations?

Organizations act like publishing departments for datasets (for example, the Department of Health). This means that datasets can be published by and belong to a department instead of an individual user.

Within organizations, admins can assign roles and authorization its members, giving individual users the right to publish datasets from that particular organisation (e.g. Office of National Statistics).

info

Create an Organization

Title:

URL: Edit

Description:

You can use [Markdown formatting here](#)

Image URL:

Create Organization

About Contact Us Powered by CKAN

Crie seu dataset

Executar os seguintes passos:

- Clique no link “Datasets” em qualquer página
- Em seguida clique em “Add Dataset”
- Preencha, no mínimo, “Title” e deixe o dataset público
- Clique no botão “Next: Add data”
- Faça upload do arquivo exemplo
 - DEINFO_ED_CENSOESCOLAR_2015.csv
- Digite um nome para o recurso e o formato

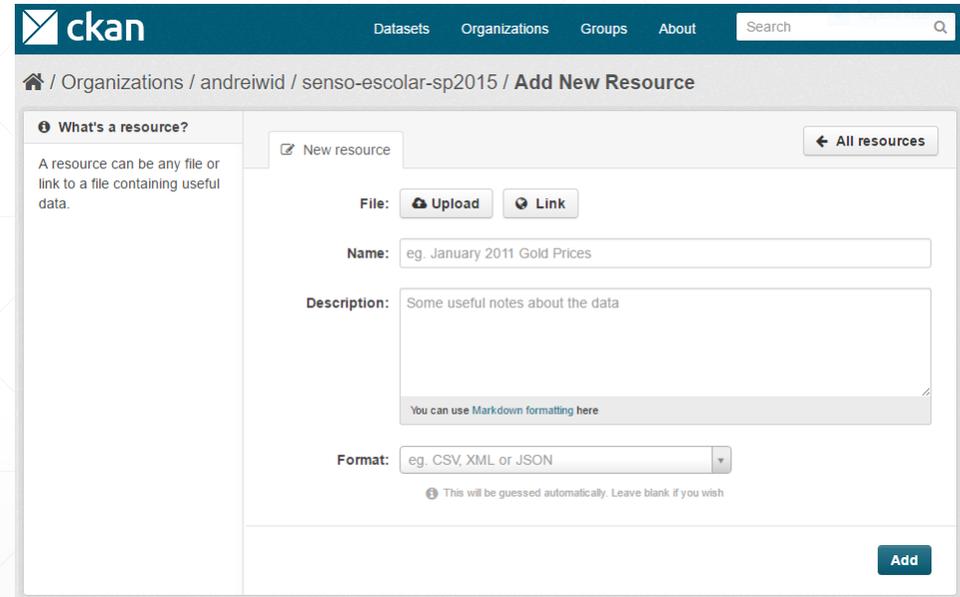
The screenshot shows the CKAN 'Create Dataset' form at step 1, 'What are datasets?'. The left sidebar contains a help text: 'A CKAN Dataset is a collection of data resources (such as files), together with a description and other information, at a fixed URL. Datasets are what users see when searching for data.' The main form area has a progress bar with '1 Create dataset' and '2 Add data'. The form fields are: Title (text input with 'eg. A descriptive title'), URL (text input with 'demo.ckan.org/dataset/<dataset>' and an 'Edit' button), Description (text area with 'eg. Some useful notes about the data' and a 'You can use Markdown formatting here' note), Tags (text input with 'eg. economy, mental health, government'), License (dropdown menu with 'Creative Commons Attribu...'), Organization (dropdown menu with 'andrewid'), Visibility (dropdown menu with 'Private'), and Source (text input with 'http://example.com/dataset.json'). A small info icon indicates that license definitions can be found at opendefinition.org.

The screenshot shows the CKAN 'Create Dataset' form at step 2, 'Add data'. The left sidebar contains a help text: 'A resource can be any file or link to a file containing useful data.' The main form area has a progress bar with '1 Create dataset' and '2 Add data'. The form fields are: File (with 'Upload' and 'Link' buttons), Name (text input with 'eg. January 2011 Gold Prices'), Description (text area with 'Some useful notes about the data' and a 'You can use Markdown formatting here' note), and Format (dropdown menu with 'eg. CSV, XML or JSON'). A small info icon indicates that the format will be guessed automatically. At the bottom, there are three buttons: 'Previous', 'Save & add another', and 'Finish'.

Adicione novos recursos

Executar os seguintes passos:

- Edite o dataset que acabou de criar
- Na aba “Resources” clique no botão “Add new resource”
- Faça upload dos seguintes arquivos
 - DEINFO_ED_CENSOESCOLAR_2015.xls
 - cotaparlamentar01-05_2016_resumo.csv
- Seu dataset deverá ter, neste momento, três recursos



The screenshot shows the CKAN interface for adding a new resource. The top navigation bar includes 'ckan', 'Datasets', 'Organizations', 'Groups', 'About', and a search box. The breadcrumb trail is 'Organizations / andreiwid / senso-escolar-sp2015 / Add New Resource'. On the left, a 'What's a resource?' sidebar explains that a resource can be a file or a link. The main form has a 'New resource' tab and an 'All resources' link. It includes fields for 'File' (with 'Upload' and 'Link' buttons), 'Name' (with the example 'eg. January 2011 Gold Prices'), 'Description' (with a text area and a note about Markdown formatting), and 'Format' (with a dropdown menu showing 'eg. CSV, XML or JSON' and a note that it will be guessed automatically). An 'Add' button is at the bottom right.

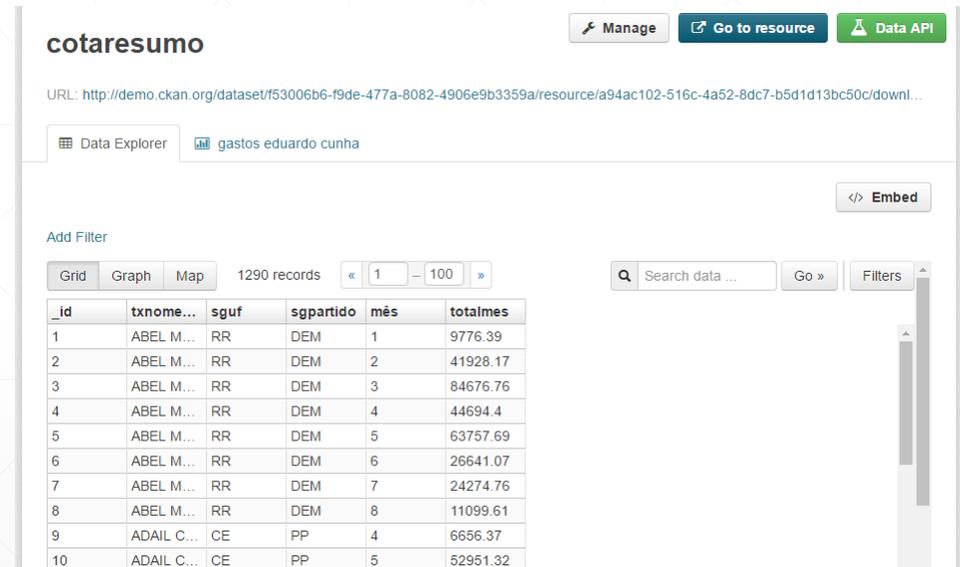
Exercícios adicionais para ambientação

- Faça a alteração dos metadados de um recurso
 - Altere a descrição
 - Faça a alteração dos metadados de um dataset
 - Altere a descrição
 - Adicione um tag
 - Crie um grupo de sua preferência e adicione seu dataset nele
-

Aprendendo a visualizar dados (tabular)

Executar os seguintes passos:

- Abrir seu dataset
- Na lista de recursos escolha o CSV referente à “cota parlamentar”
- No botão “Explore” selecione “Preview”
- No grid verifique a possibilidade de ordenar as colunas e aplicar filtros “on-the-fly”



The screenshot shows the Data Explorer interface for a dataset named "cotaresumo". The interface includes a "Data Explorer" tab, a search bar, and a table of data. The table has 1290 records and is currently displaying 10 records. The columns are: _id, txnome..., sguf, sgpartido, mês, and totalmes. The data is sorted by totalmes in descending order.

_id	txnome...	sguf	sgpartido	mês	totalmes
1	ABEL M...	RR	DEM	1	9776.39
2	ABEL M...	RR	DEM	2	41928.17
3	ABEL M...	RR	DEM	3	84676.76
4	ABEL M...	RR	DEM	4	44694.4
5	ABEL M...	RR	DEM	5	63757.69
6	ABEL M...	RR	DEM	6	26641.07
7	ABEL M...	RR	DEM	7	24274.76
8	ABEL M...	RR	DEM	8	11099.61
9	ADAIL C...	CE	PP	4	6656.37
10	ADAIL C...	CE	PP	5	52951.32

Aprendendo a visualizar dados (gráfico)

Executar os seguintes passos:

- No mesmo recurso do slide anterior aplique o seguinte filtro:
 - txnomeparlamentar = EDUARDO CUNHA
- Clique no botão “Graph”
- Nas opções do gráfico, configure o seguinte:
 - Graph Type = Lines and Points
 - Group Column (Axis 1) = mês
 - Series A (Axis 2) = totalmes



Aprendendo a visualizar dados (mapa)

Executar os seguintes passos:

- Agora acesse o recurso referente ao “censo escolar”
- Na tela de preview clique no botão “Map”
- Configure os campos de latitude e longitude
- Atualize o gráfico

Obs.: no dia 30/08 o demo.ckan.org estava apresentando problemas ao visualizar o mapa. Isso deve se normalizar em breve.

The screenshot shows a web interface for visualizing data on a map. At the top, there are tabs for 'Grid', 'Graph', and 'Map', with 'Map' selected. Below the tabs, it indicates '300 records' and a pagination control showing '1' of 100 records. A search bar with the text 'Search data ...' and a 'Go' button is visible. On the right side, there is a 'Filters' panel with the following options:

- Latitude / Longitude fields
- GeoJSON field
- Latitude field: LAT
- Longitude field: LONG
- Update button
- Auto zoom to features
- Cluster markers

The main area of the interface shows a map with numerous blue location pins. The background of the map area is a grid of text, which is partially obscured by the pins. The text visible includes phrases like '2016, direct tile access has been discontinued' and 'Please visit our blog post for more information: http://goo.gl/xB0xT'. At the bottom left, there is a 'Resources' tab.

Exercícios adicionais de visualização (persistência)

- Agora você deverá criar e salvar as visões personalizadas que fez nos últimos exercícios. Relembrando:
 - Gráfico dos gastos de Eduardo Cunha
 - Mapa com as escolas do censo escolar de 2015
 - Dica: para fazer isso, edite o recurso e acesse a aba “Views”
-

Trabalhando com APIs

Executar os seguintes passos:

- Acesse o recurso referente às “cotas parlamentares”
- Clique no botão verde do lado direito superior denominado “Data API”
- Verifique com atenção todas as opções disponíveis
- Acesse pelo menos algum exemplo diretamente de seu navegador
- Lembre-se: seu dataset deve estar público para fazer este exercício

Organizations / andreiwid / senso-escolar-sp2015 / cotaresumo

cotaresumo Manage Go to resource Data API

URL: <http://demo.ckan.org/dataset/f53006b6-f9de-477a-8082-4906e9b3359a/resource/a94ac102-516c-4a52-8dc7-b5d1d13bc50c/download...>

Data Explorer gastos eduardo cunha Embed

Add Filter

Grid Graph Map 1290 records « 1 – 100 » Go » Filters

_id	txnome...	sguf	sgpartido	mês	totalmes
1	ABEL M...	RR	DEM	1	9776.39
2	ABEL M...	RR	DEM	2	41928.17
3	ABEL M...	RR	DEM	3	84676.76
4	ABEL M...	RR	DEM	4	44694.4
5	ABEL M...	RR	DEM	5	63757.69
6	ABEL M...	RR	DEM	6	26641.07
7	ABEL M...	RR	DEM	7	24274.76

CKAN Data API

Access resource data via a web API with powerful query support. Further information in the [main CKAN Data API](#) and [DataStore](#) documentation.

Endpoints »

The Data API can be accessed via the following actions of the CKAN action API.

Create	http://demo.ckan.org/api/action/datastore_create
Update / Insert	http://demo.ckan.org/api/action/datastore_upsert
Query	http://demo.ckan.org/api/action/datastore_search
Query (via SQL)	http://demo.ckan.org/api/action/datastore_search_sql

Querying »

Query example (first 5 results)

```
http://demo.ckan.org/api/action/datastore_search?resource_id=a94ac102-516c-4a52-8dc7-b5d1d13bc50c&limit=5
```

Exercícios adicionais com APIs

- Agora faça uma consulta SQL utilizando a API do seu dataset
- A partir do dataset sobre cotas parlamentares, mostre os valores mensais mais altos que os deputados tiveram no período. Ordene
- Resultado em formato de tabela:

- O resultado encontra-se nos arquivos de apoio

	txnomeparlamentar	maximo
1	CARLOS ANDRADE	93895,28
2	ELCIONE BARBALHO	91644,17
3	CARLOS BEZERRA	89210,83
4	ENIO VERRI	86938,18
5	DÉCIO LIMA	85350,53
6	CARLOS MANATO	85019,96
7	ABEL MESQUITA JR.	84676,76
8	CARLOS MELLES	84236,22
9	AUGUSTO COUTINHO	79922,86

Caminho das pedras para implementação de uma instância CKAN

Componentes do CKAN

- Servidor web para o WSGI.
 - Mais usados: Apache ou Nginx.
 - SGBD PostgreSQL para metadados.
 - Apache Solr para indexação.
 - Python como base e para ambiente virtual.
-

Componentes do CKAN

- Possui suporte para os principais CMS existentes, como Drupal, Joomla, Wordpress e outros.
 - Para tal, é necessária a instalação de plugins que façam esse suporte.
-

Como instalar

Há as seguintes formas de instalação do CKAN:

- Por pacote.
 - Pelo fonte.
-

Como instalar - Pacote

Requisito – Ubuntu 12.04 ou 14.04 de 64 bits.

- Instalar os pacotes do CKAN e dependências.
 - Instalar PostgreSQL e o Solr.
 - Reiniciar Apache e o Nginx.
-

Como instalar – Pacote - Passos

- `sudo apt-get update`
 - `sudo apt-get install -y nginx apache2 libapache2-mod-wsgi libpq5`
 - `wget http://packaging.ckan.org/python-ckan_2.2_amd64.deb`
 - `sudo dpkg -i python-ckan_2.2_amd64.deb`
 - `sudo apt-get install -y postgresql solr-jetty`
 - `sudo service apache2 restart`
 - `sudo service nginx restart`
-

Como instalar - Fonte

Pelo fonte, dependerá da versão do CKAN e do SO. Para ver os passos, pode-se acessar os seguintes links:

- <https://github.com/ckan/ckan/wiki/How-to-Install-CKAN>
 - <http://docs.ckan.org/en/latest/maintaining/installing/install-from-source.html>
-

Como instalar - Fonte

Basicamente, os passos são

- Instalar pacotes e suas dependências.
 - Instalar Python.
 - Baixar o fonte e instalá-lo no ambiente virtual do Python.
 - Baixar e configurar SGBD (PostgreSQL).
 - Criar arquivo de configuração (production.ini).
 - Configurar o Solr.
 - Criar o banco de dados
-

Pagina Inicial

The screenshot shows the homepage of the IMA (Informática de Municípios Associados) open data portal. The page features a dark red header with the IMA logo and navigation links. A search bar is prominently displayed, and the main content area is divided into two columns. The left column contains a welcome message and a list of data sets for 'Balança' (Balance) for the years 2012 and 2013, each with PDF and XLS download options. The right column features a search box, a list of common tags, and a detailed view of the 'Balança de 2012' dataset, including its title, description, and download links. The footer contains links to the portal, API, and Open Knowledge Foundation, as well as logos for CKAN and the IMA.

Entrar Registrar

Dados Abertos
ima
INFORMÁTICA
de Municípios Associados S/A

Conjuntos de dados Organizações Grupos Sobre Pesquisar conjuntos de dados

Procurar seus dados

Pesquisar

Etiquetas Mais Comuns

Bem vindo ao portal de Dados Abertos da IMA

Balanço

Balanço de 2012

PDF XLS

Balanço de 2013

PDF XLS

IMA
IMA - Informática de Municípios Associados

Balanço de 2012

PDF XLS

Remunerações - Dezembro de 2014
Remunerações de dezembro de 2014 (inclui salários e 13º).

PDF

Remuneração - Novembro 2014
Remunerações de novembro de 2014

PDF

Balanço de 2013

PDF XLS

O portal

Sobre o portal
API do CKAN
Open Knowledge Foundation
OPEN DATA

Links de interesse

Dados Recife
Dados Rio
#datapoa
Dados abertos - Distrito Federal

Impulsionado por
ckan

Desenvolvido por
ima
INFORMÁTICA
de Municípios Associados S/A

Adicionar (e remover) usuário

```
paster --plugin=ckan user add usuarioteste --config=/etc/ckan/std/std.ini
```

```
paster --plugin=ckan user remove usuarioteste --config=/etc/ckan/std/std.ini
```

Sim, é por meio de plugins. Mais plugins podem ser vistos neste link:

<http://extensions.ckan.org/>

Informações e ajuda

- Documentação - <http://docs.ckan.org>
 - Lista de discussão - <https://lists.okfn.org/mailman/listinfo/ckan-pt>
-

Mãos à obra

- O CKAN oferece um lugar para quem quer começar a brincar com o programa. O lugar é este: <http://demo.ckan.org/>
-

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.
