



Emprel

Sistemática de Atualização de Dados do Portal de Dados Abertos do Recife

Recife, 31 de outubro de 2023

Expediente

Prefeitura do Recife

João Henrique de Andrade Lima Campos
Prefeito do Recife

Isabella Menezes de Roldão Fiorenzano
Vice-prefeita do Recife

Felipe Martins Matos
Secretário de Planejamento, Gestão e Transformação Digital

Felipe Cadena Muniz
Secretario Executivo de Transformação Digital da Secretário de Planejamento, Gestão e Transformação Digital

Conselho de Administração

Felipe Martins Matos
Presidente

Maíra Rufino Fischer
Vice – Presidente

Joana Portela Florêncio
Conselheiro

Bernardo Juarez D’Almeida
Conselheiro

Rafael Figueiredo Bezerra
Conselheiro

Mesa Diretora da EMPREL

Bernardo Juarez D’Almeida
Presidente

Rafael Figueiredo Bezerra
Vice-Presidente

Vitor Pavesi
Diretor de Administração e Finanças

Alonso José da Silva Filho
Diretor de Planejamento e Atendimento ao Cliente

João Henrique de Lima Lobo
Diretor da Assessoria de Assuntos Jurídicos

Alyson carvalho Pereira de Matos
Diretor de Infraestrutura de Informática

Ana Carolina Alves Breda
Diretora de Transformação Digital

Rosana Carvalho Barbosa
Diretora de Sistemas Financeiros e Tributários

Breno Alencar Gonçalves
Diretor de Inovação Aberta e Governança de Dados

João Luís Carvalho Paes
Diretoria de Comunicação em *Streaming*

Clóvis Farias do Monte Júnior
Assessor de Negócios Corporativos e Auditoria Interna

Sumário

| | |
|---|----|
| 1 - Apresentação | 5 |
| 2 - Princípios e diretrizes relacionados aos dados abertos | 6 |
| 3 - Classificação dos dados | 7 |
| 4 - Atualização dos dados | 10 |
| 4.1 - Carregamento de dados pela interface gráfica do usuário | 10 |
| 4.2 - Atualização de dados através de execução de rotinas de carga de dados | 11 |

Lista de Figuras

| | |
|--|----|
| Figura 1 - Cópia da tela inicial do Portal de Dado Abertos da Prefeitura do Recife | 6 |
| Figura 2 - Cópia da tela de Grupos do Portal de Dado Abertos da Prefeitura do Recife . | 7 |
| Figura 3 - Cópia da tela de Conjuntos de Dados do Portal de Dado Abertos da Prefeitura do Recife | 8 |
| Figura 4 - Cópia da tela com os Formatos de Arquivos do Portal de Dado Abertos da Prefeitura do Recife | 9 |
| Figura 5 - Cópia da tela de Organizações do Portal de Dado Abertos da Prefeitura do Recife | 10 |

1 - Apresentação

Este documento descreve a sistemática de atualização dos dados do Portal de Dados Abertos do Recife, em Pernambuco, Brasil.

O Portal de Dados Abertos do Recife é o repositório onde a Prefeitura disponibiliza os dados públicos que podem ser acessados livremente e de forma gratuita pela população, através do endereço <http://dados.recife.pe.gov.br>.

Destaque-se que os dados estão dispostos em formatos adequados para serem consumidos e entendidos por máquina, para servirem como entrada em sistemas, sites, portais e painéis de *business intelligence*, entre outros, para os interessados, ou seja, pesquisadores, *startups* e empresas. Essas ferramentas poderão gerar visualizações finais para o cidadão através de um processamento.

As informações disponibilizadas pela Prefeitura do Recife no formato adequado para consumo direto por seres humanos estão disponíveis no Portal da Transparência da Prefeitura do Recife, no endereço <http://transparencia.recife.pe.gov.br>. Lá, poderão ser encontrados e baixados, também de forma livre e gratuita, consultas e relatórios da gestão.

2 - Princípios e diretrizes relacionados aos dados abertos

Os dados disponibilizados no Portal de Dados Abertos do Recife (Figura 1) seguem os seguintes princípios e diretrizes:



Figura 1 - Cópia da tela inicial do Portal de Dado Abertos da Prefeitura do Recife

Granularidade: O dado deve ser disponibilizado na sua maior granularidade possível. Isto significa que os dados não serão dispostos em formato consolidado ou sumarizado, pelo contrário, serão disponibilizados na sua forma atômica.

Sigilo: Todo o dado disponibilizado será aquele considerado não sigiloso, não sensível, ou seja, dado que não fira o direito à privacidade de qualquer pessoa física ou jurídica, segundo a LGPD - Lei Geral de Proteção de Dados.

Formato: Os dados abertos serão disponibilizados em formatos livres, abertos, e de forma gratuita.

Formas de acesso aos dados: Os dados podem ser acessados de duas formas principais: através de *download* (baixa) de arquivos de dados ou consumidos diretamente por programas de computador via APIs (*Application Program Interface*).

Facilidade de localização do dado desejado: O Portal de Dados Abertos indexa todo o conteúdo armazenado, de forma a facilitar a busca de dados por palavra-chave, por conjunto de dados, por recurso de dados, por formato do arquivo de dados e por órgão detentor do dado.

Definição dos dados: Cada recurso de dados contém o seu metadado relacionado. O metadado é um descritor do conteúdo de um recurso de dados. Ele deve apresentar cada dado pertencente a um recurso de dado, descrevendo o seu nome, significado, formato (inteiro, alfanumérico, data, etc), tamanho, alguma restrição ou lei de formação, e, por fim, a periodicidade de atualização do seu conteúdo.

Serviço contínuo de dados: Uma vez disponibilizado um recurso de dados no portal, esse recurso deverá permanecer indefinidamente disponível no portal, pois a premissa é que esses dados estarão sendo utilizados por algum programa e/ou serviço e não deverão sofrer com indisponibilidades.

3 - Classificação dos dados

Os **dados atômicos** são agrupados em **Recursos de Dados** (ex.: tabelas), que por sua vez são agrupados em **Conjuntos de Dados**. Portanto, existe uma hierarquização dos dados, sendo o conjunto de dados a macro estrutura que agrupa conceitualmente vários recursos de dados relacionados, que por sua vez são, em última instância, a materialização do que é entregue aos usuários do portal.

Os conjuntos de dados estão agrupados em **Grupos** (Figura 2), que existem para criar e gerenciar coleções de conjuntos de dados. Isso pode ser feito para catalogar conjuntos de dados de um projeto ou time particular, ou em um tema particular, ou como uma forma simples de ajudar as pessoas a encontrar e buscar seus próprios conjuntos de dados.

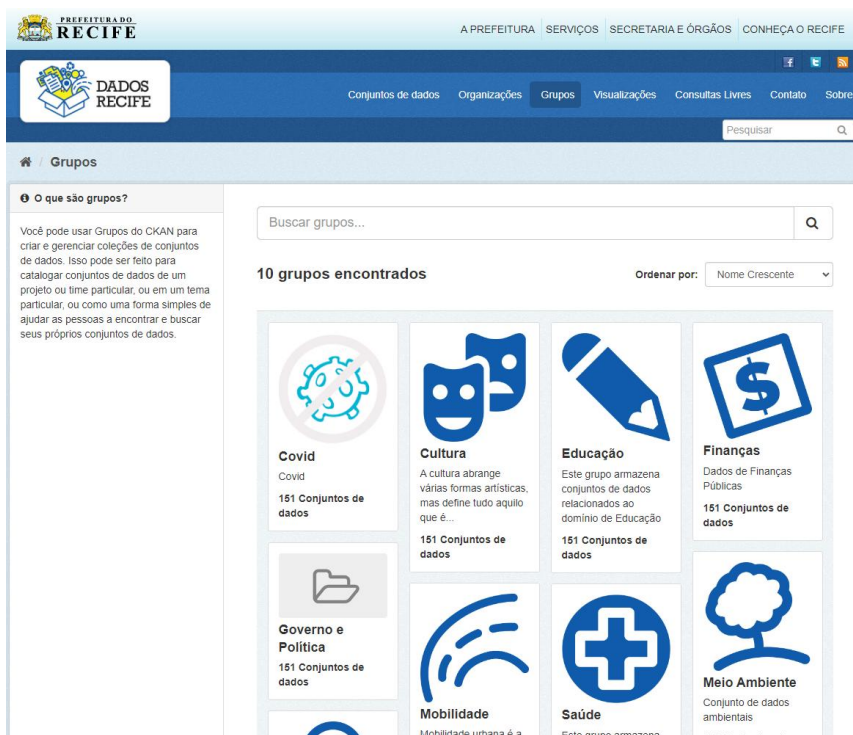


Figura 2 - Cópia da tela de Grupos do Portal de Dado Abertos da Prefeitura do Recife

Além da classificação em conjuntos de dados (Figura 3), os recursos de dados também são classificados pelo seu formato de arquivo (Figura 4).

The screenshot shows the 'Conjuntos de dados' (Datasets) page of the Recife Open Data Portal. At the top, there is a navigation bar with links for 'A PREFEITURA', 'SERVIÇOS', 'SECRETARIA E ÓRGÃOS', and 'CONHEÇA O RECIFE'. Below this is a search bar and a menu with options like 'Conjuntos de dados', 'Organizações', 'Grupos', 'Visualizações', 'Consultas Livres', 'Contato', and 'Sobre'. The main content area displays a search bar with the text 'Search datasets...' and a magnifying glass icon. Below the search bar, it indicates '151 conjuntos de dados encontrados' (151 datasets found) and an 'Ordenar por:' (Sort by) dropdown menu set to 'Relevância' (Relevance). The datasets are listed with titles and descriptions, such as 'Demandas dos Cidadãos e Serviços - Dados Vivos Recife', 'Licenciamento Urbanístico', 'Licenciamento Ambiental', and 'Licenciamento Sanitário'. Each dataset entry includes icons for its file format, such as JSON, CSV, PDF, and GeoJSON. A sidebar on the left provides navigation options under 'Organizações' (Secretaria de Saúde: 52, Autarquia de Trânsi...: 24, Secretaria de Educação: 15, Empresa de Manutenc...: 7, Secretaria de Meio ...: 7, Secretaria de Infra...: 6, Secretaria de Turis...: 6, Secretaria de Finanças: 5, Emprel: 4, Secretaria de Desen...: 4) and 'Mostrar mais Organizações'. Below this, there are 'Grupos' (Groups) with counts: Saúde (45), Urbanismo (22), Mobilidade (21), Governo e Política (15), Covid (12), and Educação (11).

Figura 3 - Cópia da tela de Conjuntos de Dados do Portal de Dado Abertos da Prefeitura do Recife

Os formatos de arquivos dos recursos de dados no portal são os seguintes:

GEOJSON - Formato livre de dado espacial, que pode ser utilizado em aplicações de geoprocessamento. Cada arquivo geojson corresponde a uma camada de geoinformação, que contém a parte espacial do dado, podendo ser no formato de ponto, linha ou polígono, individualizado ou em múltiplos, que são definidos por um par de coordenadas (para o ponto) ou um conjunto de pares de coordenadas espaciais (linhas e polígonos), seguindo um determinado Sistema de Referência de Coordenadas (ex.: coordenadas geográficas, com latitude e longitude), como também podendo conter a parte descritiva dos dados, em formato de tabela, que estará atrelada ao dado espacial.

JSON - Formato livre de arquivo de dados descritivos, facilmente manipulados por linguagens de programação como java, php e outras.

CSV - Formato livre de arquivo de dados descritivos, onde os dados são delimitados por um separador definido (ex.: vírgula).

PDF - Formato de dado de leitura padrão de mercado utilizado para apresentação de dados. Esse formato só é utilizado dentro do Portal de Dados Abertos do Recife para metadados de recursos de dados e estão sendo gradativamente substituídos pelo formato JSON.

The screenshot displays the Open Data Portal interface. On the left is a sidebar with a list of categories and their respective counts: Governo e Política (15), Covid (12), Educação (11), Finanças (11), Turismo (9), Cultura (6), Meio Ambiente (4), and a section for 'Etiquetas' (Tags) including saúde (29), mobilidade (17), ctu (15), covid (14), transito (14), transporte (13), covid-19 (12), corona (11), velocidade (11), and Censo Escolar (10). Below the tags is a 'Formatos' section listing CSV (144), JSON (121), PDF (54), and GeoJSON (21).

The main content area on the right shows several data sets with their descriptions and available formats (CSV, PDF, JSON):

- Licitações**: Licitações – Processo Licitatório ou Licitação é um processo administrativo realizado para selecionar a melhor proposta de fornecimento de materiais e serviços para a... (Formats: CSV, PDF)
- Perfil das Pessoas Vacinadas - Covid-19**: Contém os microdados relativo a vacinação contra a COVID-19 (Formats: JSON, CSV)
- Central de Atendimento de Serviços da Emlurb – 156**: São os dados das demandas de serviços de arborização, Drenagem, Iluminação, Limpeza, Pavimentação, tapa buracos, entre outros serviços de reparação da Cidade do Recife,... (Formats: CSV, PDF, JSON)
- Relação de pessoas vacinadas - Covid 19**: Contém a relação de pessoas vacinadas bem como a categoria de grupos a qual pertencem. Idosos: Idosos de ILPIs, Idosos de 80 a 84 anos, Idosos com 85 anos ou mais... (Formats: CSV, JSON)
- Despesas Orçamentárias**: Despesas realizadas pelo governo municipal com serviços, obras e compras, com a finalidade de atender as necessidades coletivas da população. (Formats: CSV, PDF, JSON)
- Receitas**: Receita Pública é a soma de ingressos, impostos, taxas, contribuições e outras fontes de recursos, arrecadados para atender às despesas públicas. (Formats: CSV, PDF, JSON)
- Gestão de Dados Abertos**: Mapeamento dos conjuntos de dados disponíveis com a situação quanto à atualização. (Formats: CSV, JSON)

Figura 4 - Cópia da tela com os Formatos de Arquivos do Portal de Dado Abertos da Prefeitura do Recife

Por fim, os recursos de dados também são agrupados pelas organizações responsáveis (Figura 5) pelos recursos de dados.

PREFEITURA DO RECIFE

A PREFEITURA | SERVIÇOS | SECRETARIA E ÓRGÃOS | CONHEÇA O RECIFE

DADOS RECIFE

Conjuntos de dados | **Organizações** | Grupos | Visualizações | Consultas Livres | Contato | Sobre

Pesquisar

Organizações

O que são Organizações?

Organizações do CKAN são usadas para criar, gerenciar e publicar coleções de conjuntos de dados. Usuários podem possuir diferentes papéis em uma organização, dependendo do seu level de autorização para criar, editar e publicar.

Pesquisar organizações...

23 organizações encontradas Ordenar por: Nome Crescente

CTTU
AUTARQUIA DE TRÂNSITO E TRANSPORTE URBANO DO RECIFE
Autarquia de Trânsito e Transporte Urbano do Recife - CTTU
A Autarquia de Trânsito e Transporte Urbano do Recife (CTTU)...
151 Conjuntos de dados

DADOS RECIFE
COMDICA - Conselho Municipal de Defesa e Promoção dos Direitos da Criança e do Adolescente
O Conselho Municipal de Defesa e Promoção dos Direitos da Criança e do...
151 Conjuntos de dados

Controladoria Geral do Município
A Controladoria Geral do Município (CGM) é o órgão integrante da...
151 Conjuntos de dados

Emprel
Empresa Municipal de Informática
151 Conjuntos de dados

Empresa de Manutenção e Limpeza Urbana - EMLURB
A EMLURB é uma Empresa Pública, constituída em 26 de abril de 1979, pelo...
151 Conjuntos de dados

Figura 5 - Cópia da tela de Organizações do Portal de Dado Abertos da Prefeitura do Recife

A plataforma do Portal disponibiliza facilidades de consultas por cada tipo de classificação citado acima, ou seja, o usuário poderá buscar os recursos de dados via palavra-chave ou pesquisando os recursos de dados de um grupo ou de um conjunto de dados ou pesquisando os recursos de dados por formato de arquivos ou, ainda, pesquisando os recursos de dados de uma determinada organização detentora de recursos de dados.

4 - Atualização dos dados

Existem basicamente duas formas de atualização dos dados do Portal de Dados Abertos:

1. Carregamento de dados pela interface gráfica do usuário do portal, e
2. Atualização de dados através de execução de rotinas de carga de dados oriundos de algum sistema computacional detentor dos dados.

4.1 - Carregamento de dados pela interface gráfica do usuário

Essa modalidade de atualização é utilizada por um usuário administrador do portal e é utilizada em duas situações: 1) para carregamento de metadados de recursos de dados; e 2)

para carregamento de dados originais no formato de planilhas eletrônicas de dados, que serão convertidos e armazenados no formato de recursos de dados no portal.

O metadado é criado quando o recurso de dados correspondente é disponibilizado pela primeira vez no portal, e somente será atualizado se houver alguma mudança na estrutura de algum dado do recurso de dados associado.

O recurso de dados deve ser atualizado manualmente nas datas correspondentes à periodicidade definida no arquivo de metadado do recurso de dados associado. Para isso, o responsável pelo recurso de dados, ou seja, o representante da secretaria, fundação, empresa, autarquia da Prefeitura responsável pelos dados, deverá fornecer a correspondente planilha eletrônica de dados também na periodicidade acordada com o administrador do portal, de forma a permitir que o recurso de dados seja carregado no prazo adequado, definido no metadado correspondente. Esse envio da planilha ocorre via e-mail para o grupo de e-mails da administração do portal (dadosabertos@recife.pe.gov.br).

4.2 - Atualização de dados através de execução de rotinas de carga de dados

A atualização de dados no Portal de Dados Abertos do Recife via execução de rotinas de carga de dados acontece apenas para dados que estão armazenados originalmente em sistemas computacionais, de forma que a atualização pode ser automatizada.

Neste caso, se faz necessário construir uma rotina que irá extrair os dados de sistemas transacionais ou de painéis de BI (business intelligence) e importar esses dados diretamente para o CKAN (plataforma computacional livre que suporta e implementa o Portal de Dados Abertos do Recife), como também se faz necessário programar e comandar a periodicidade de execução dessa rotina via um agendador de tarefas.

O agendador de tarefas, por sua vez, pode ser o agendador de tarefas do sistema operacional do servidor onde o sistema computacional detentor do dado original está em execução (ex.: Windows ou Linux) ou através de software orquestrador de execução de tarefas, que no caso específico, esse software é o JENKINS (software livre). Essa periodicidade deve ser aquela definida no arquivo de metadado do respectivo recurso de dados que está sendo carregado no portal.