

Parceria para Governo Aberto (OGP)

4º Plano de Ação

Documento de referência dos Marcos 2 e 3 do
Compromisso 2: Ecossistema de Dados Abertos

Modelos de Abertura de Dados: Potencialidades e Limites

Brasília

2019

SUMÁRIO

INTRODUÇÃO	3
PLATAFORMAS DE DADOS ABERTOS	9
1.1. Características desejáveis para as ODP	10
1.2. Plataformas de dados abertos mais utilizadas	13
1.2.1. CKAN	14
1.2.2. Socrata	14
1.2.3. ArcGIS Open Data	15
1.2.4. Frictionless Data	15
BENCHMARK INTERNACIONAL	16
2.1. Recomendações do W3C	16
2.2. Dados abertos na Europa	16
2.2.1. Portal de dados abertos da União Europeia	17
2.2.2. Portal europeu de dados	18
2.2.3. Portal do Centro de Conhecimento em Bioeconomia	18
2.2.4. Portal do Centro de Pesquisa Conjunta	19
2.3. Dados Abertos na América Latina	19
MAPEAMENTO DOS MODELOS DE ABERTURA NO BRASIL	20
3.1. Portais nacionais que fazem uso da plataforma CKAN	20
3.2. Dados geográficos na cidade de São Paulo - Geosampa	23
3.3. Dados de segurança pública no Rio de Janeiro - ISP Dados	23
3.4. Portal de Transparência do Governo do Distrito Federal	24
3.5. Pátio Digital - Transparência e dados abertos na secretaria de Educação de São Paulo	25
3.6. Modelos de abertura de dados do Instituto Brasileiro de Geografia e Estatística (IBGE)	26
3.6.1. SIDRA - Sistema IBGE de Recuperação Automática	26
3.6.2. Séries Históricas e Estatísticas	27
3.6.3. Portal do IBGE	27
3.6.4. Downloads	28
3.7. Portal de dados do Instituto de Pesquisa Econômica Aplicada (IPEA)	29
3.8. Dados Abertos da Câmara dos Deputados	30
3.9. Dados Abertos do Senado Federal	30
3.9.1. Página de Dados Abertos do Senado Federal	31
3.9.2. LexML Brasil	31
3.9.3. SIGA Brasil	32
3.10. Portal da Transparência da Controladoria-Geral da União (CGU)	33
3.11. TABNET - Departamento de Informática do SUS (DATASUS)	34
3.12. Repositório de Dados Eleitorais do Tribunal Superior Eleitoral (TSE)	35
3.13. Dados Abertos FEE (RS)	35
3.14. Portal da Transparência do Ministério Público Federal (MPF)	36

CONSOLIDAÇÃO DOS RESULTADOS	37
REFERÊNCIAS	38
EXPEDIENTE	40

INTRODUÇÃO

Dados são considerados **abertos** quando qualquer pessoa é livre para acessá-los, usá-los, modificá-los e compartilhá-los, sujeitos, no máximo, a medidas que preservem sua origem e sua publicidade (OKFN, 2006).

Internacionalmente, foram consensuados os “8 Princípios de Dados Abertos”, os quais definem que, para serem abertos, os dados devem ser:

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. Dados públicos são dados que não estão sujeitos a limitações válidas de privacidade, segurança ou controle de acesso, reguladas por estatutos.
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.
5. **Processáveis por máquina.** Os dados são razoavelmente estruturados para possibilitar o seu processamento automatizado.
6. **Acesso não discriminatório.** Os dados estão disponíveis a todas as pessoas, 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. **Licenças livres.** Os dados não estão sujeitos a restrições por 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.

Assim, qualquer campo de conhecimento ou prática pode realizar-se a partir dos princípios de dados abertos, sejam informações sobre políticas públicas, mas também, imagens, músicas, livros, conhecimento acadêmico, etc. Com isso, também, passaram a existir redes

articuladas em torno desta concepção, como “cidades abertas”, “governos abertos”, “educação aberta”, “transporte aberto”, “jornalismo de dados” e muitas outras.

No **Brasil**, a abertura de dados públicos existe há pelo menos 15 anos. Embora o acesso à informação sobre recursos públicos já estivesse previsto na Constituição de 1988, foi após a revolução digital e legislações específicas que este processo se intensificou em termos de procedimento de gestão pública. Podemos citar como marco o lançamento do Portal da Transparência do Governo Federal, em 2004.

Considerada um marco para o desenvolvimento de ações relativas à área de dados abertos no Brasil, a Lei da Transparência (L.C. nº 131/2009) alterou a redação da Lei de Responsabilidade Fiscal no que se refere à transparência da gestão fiscal e promoveu inovações ao determinar que sejam disponíveis, em tempo real, informações pormenorizadas sobre a execução orçamentária e financeira dos órgãos públicos.

Este processo se intensificou após o Decreto 7.185/2010, que dispõe sobre o padrão mínimo de qualidade do sistema integrado de administração financeira e controle, e a **Lei de Acesso à Informação** (Lei nº 12.527/2011). Além de regulamentar o fornecimento de informações sob demanda do cidadão, também determina que órgãos e entidades devem se antecipar aos pedidos e publicar seus dados e informações na internet, bem como exige que os dados sejam publicados em formatos abertos e não-proprietários. Dentre as diretrizes dispostas no art. 3º da Lei, destacam-se: observância da publicidade como preceito geral e do sigilo como exceção; a divulgação de informações de interesse público, independentemente de solicitações; a utilização de meios de comunicação viabilizados pela tecnologia da informação; o fomento ao desenvolvimento da cultura de transparência na administração pública; e o desenvolvimento do controle social.

Além disso, também em 2011, o Brasil tornou-se membro co-fundador da Parceria para Governo Aberto - em inglês *Open Government Partnership (OGP)* - e assumiu internacionalmente um compromisso de estruturação dos primeiros passos para uma política de dados abertos. O Brasil encontra-se em fase de implementação do 4º Plano de Ação Nacional da OGP e cabe mencionar que todos os planos anteriores tinham compromissos relacionados com o tema de dados abertos.

O Portal de Dados Abertos do Governo Federal, fruto de um dos compromissos formalizados no Primeiro Plano de Ação da OGP, consiste numa ferramenta disponibilizada pelo governo para que todos possam encontrar e utilizar os dados e as informações públicas. Também cabe destacar a instituição da Infraestrutura Nacional de Dados Abertos (INDA) em 2012, tendo como principal objetivo incentivar a oferta de dados abertos pelos seus produtores de forma padronizada e centralizada em catálogo governamental e sua utilização e agregação de valor pela sociedade. A Gestão INDA é exercida por um Comitê Gestor composto de órgãos públicos, sociedade civil e academia e que direciona os principais aspectos operacionais da temática de dados abertos públicos no país.

Em 2016, instituiu-se a Política de Dados Abertos do Poder Executivo Federal, por meio do Decreto 8.777/2016. A obrigatoriedade no cumprimento da Política contempla os órgãos e entidades da administração pública federal direta, autárquica e fundacional, que devem promover a publicação de dados sem natureza sigilosa contidos em bases de dados. Dessa forma, cada órgão ou entidade possui a obrigação de elaborar um Plano de Dados Abertos (PDA) com vigência de dois anos contendo o planejamento das ações de implementação e promoção da abertura de seus dados.

Para complementar o Decreto 8.777/2016, o Comitê Gestor da Infraestrutura Nacional de Dados Abertos aprovou a Resolução Nº03, de 13 de outubro de 2017, e assim estabeleceu normas sobre a elaboração e publicação dos PDAs. Dentre as normas contidas na referida Resolução, vale destacar que os órgãos devem utilizar mecanismos de participação social para conhecer a demanda pelos dados do órgão para então priorizar e otimizar os esforços de abertura de suas bases.

Lançado em 2018, o Programa Escala Brasil Transparente - EBT - Avaliação 360°, diferentemente das três edições anteriores, trouxe uma inovação em sua metodologia. A principal diferença é que além da transparência passiva, também foi avaliada a transparência ativa, incluindo a verificação da disponibilização de dados abertos pelos entes federados.

NORMATIVOS

Listamos, abaixo, o conjunto de normas que se relacionam direta ou indiretamente com o tema dados abertos:

- **Lei de Responsabilidade Fiscal** (LC 101/2000): estabelece normas de finanças públicas voltadas para a responsabilidade na gestão fiscal.
- **Lei da Transparência** (LC 131/2009): estabelece normas de finanças públicas voltadas para a responsabilidade na gestão fiscal, a fim de determinar a disponibilização, em tempo real, de informações pormenorizadas sobre a execução orçamentária e financeira da União, dos Estados, do Distrito Federal e dos Municípios.
- **Lei nº 10.650/2003**: dispõe sobre o acesso público aos dados e informações existentes nos órgãos e entidades integrantes do Sistema Nacional do Meio Ambiente - Sisnama.
- **Lei de Acesso à Informação** (LO 12.527/2011): regulamenta o direito constitucional de acesso às informações públicas.

- **Lei nº 13.473/2017:** dispõe sobre as diretrizes para a elaboração e execução da Lei Orçamentária de 2018 e dá outras providências.
- **Decreto s/n de 15/09/2011:** dispõe sobre o Plano de Ação Nacional sobre Governo Aberto e dá outras providências.
- **Decreto nº 5482/2005:** dispõe sobre a divulgação de dados e informações pelos órgãos e entidades da administração pública federal, por meio da Rede Mundial de Computadores - Internet.
- **Decreto nº 7.724/2012:** Regulamenta a Lei de Acesso à Informação (Lei nº 12.527, de 18 de novembro de 2011).
- **Decreto nº 8.638/2016:** dispõe sobre a Política de Governança Digital no âmbito dos órgãos e das entidades da administração pública federal direta, autárquica e fundacional.
- **Decreto nº 8.638/2016:** dispõe sobre a política de governança da administração pública federal direta, autárquica e fundacional.
- **Decreto nº 9.468/2018:** dispõe sobre o Conselho de Transparência Pública e Combate à Corrupção.
- **Instrução Normativa IN nº 04 MPOG de 12/04/2012:** dispõe sobre a Infraestrutura Nacional de Dados Abertos - INDA.
- **Resolução nº 03 do CGINDA:** aprova as normas sobre elaboração e publicação de Planos de Dados Abertos, conforme disposto no Decreto nº 8.777, de 11 de maio de 2016.

Hoje, além das diversas bases de dados abertas que concentram informações de todos os recursos federais, e também os documentos oficiais, praticamente todos os órgãos têm páginas de transparência em seus portais. Além disso, estados e municípios passaram a ser obrigados a disponibilizar os dados, sendo que alguns estão mais ou menos avançados que outros, até mesmo porque em municípios pequenos a própria questão tecnológica ainda é uma limitação. Ainda no que diz respeito à legislação, o **Marco Civil da Internet** é mais uma referência a ser considerada, pois inclui a transparência como um dos princípios que devem reger a estrutura de governança da internet.

A publicação “5 motivos para a abertura de dados na Administração Pública” elaborada pelo Tribunal de Contas da União (TCU)¹, apresenta razões para que as organizações públicas invistam em iniciativas de abertura de dados governamentais. Os cinco motivos para a abertura dos dados são:

1. Transparência na gestão pública;
2. Contribuição da sociedade com serviços inovadores ao cidadão;
3. Aprimoramento na qualidade dos dados governamentais;
4. Viabilização de novos negócios;
5. Obrigatoriedade por lei.

É neste contexto que se insere o Compromisso 2: Ecosistema de Dados Abertos, do 4º Plano de Ação do *Open Government Partnership (OGP)*, definido como: “Estabelecer, de forma colaborativa, modelo de referência de política de dados abertos que promova integração, capacitação e sensibilização entre sociedade e as três esferas de governo a partir do mapeamento das demandas sociais”.

Para o alcance deste compromisso, estão atuando conjuntamente os seguintes representantes do governo e da sociedade civil:

- Governo: Controladoria-Geral da União (CGU), Ministério da Economia, Ministério da Ciência, Tecnologia, Inovação e Comunicações, Ministério da Educação, Ministério da Saúde;
- Sociedade Civil: Diretoria de Análise de Políticas Públicas da Fundação Getúlio Vargas (FGV DAPP), Instituto de Estudos Socioeconômicos (Inesc), Centro de Estudos de Tecnologias Web/Ceweb.br do Núcleo de Informação e Coordenação do Ponto Br (NIC.br); Rede pela Transparência e Participação Social (RETPS) e Open Knowledge Brasil (OKBr).

Na ocasião da elaboração do desenho do Compromisso 2, foram definidos coletivamente 9 Marcos para o seu alcance:

Marco 1 - Identificação dos atores nas 3 esferas de governo e na sociedade

Marco 2 - Mapeamento dos modelos de abertura de dados existentes

Marco 3 - Identificação das potencialidades e dos limites dos modelos existentes

¹ Disponível em <https://portal.tcu.gov.br/biblioteca-digital/cinco-motivos-para-a-abertura-de-dados-na-administracao-publica.htm>. Acesso em 16/07/2019.

Marco 4 - Elaboração de estrutura de modelo de referência

Marco 5 - Elaboração de texto de cada tópico da estrutura de modelo de referência

Marco 6 - Realização de consulta pública sobre o texto do modelo

Marco 7 - Plano de comunicação e disseminação do modelo para as três esferas do governo e sociedade civil

Marco 8 - Lançamento do modelo de referência

Marco 9 - Disseminação do modelo de referência

Este relatório se refere ao desenvolvimento dos **Marcos 2 e 3** do Compromisso 2 - Ecossistema de dados abertos, ou seja, apresenta um **mapeamento dos modelos de abertura de dados existentes e identifica, nestes modelos, potencialidades e limitações**. Partindo de diagnósticos realizados no âmbito do Marco 1 com relação aos atores que fazem parte do Ecossistema de Dados Abertos no Brasil, o relatório pretende traçar fundamentos para elaboração de um modelo de referência de dados abertos no Brasil, tarefa a ser cumprida ao longo dos próximos marcos deste compromisso.

O documento encontra-se estruturado em quatro seções além desta introdução. A seção 1 apresenta uma discussão sobre plataformas de dados abertos comumente utilizadas por diversos órgãos para publicação de dados. Em seguida, a seção 2 traz um breve benchmark de experiências internacionais de publicação de dados abertos. A seção 3 é o resultado do mapeamento de modelos de abertura e aplicação das plataformas no Brasil, traçando potencialidades e limites de cada um. Por fim, a seção 4 faz uma interpretação dos resultados, traçando considerações para elaborar um modelo de referência aplicável à realidade brasileira.

1. PLATAFORMAS DE DADOS ABERTOS

Como apresentado na introdução do documento, dado aberto é definido como o dado que pode ser usado livremente, reutilizado e redistribuído por qualquer pessoa ou grupos de pessoas, limitado apenas aos requisitos de atribuição e compartilhamento. A abertura dos dados, neste contexto, significa que os dados devem estar disponíveis e não devem apresentar um custo de reprodução excessivo, de preferência possam ser baixados pela Internet de uma forma conveniente e modificável. A abertura dos dados indica também os termos sob os quais a reutilização e a redistribuição possam ser realizadas, incluindo a combinação com outros conjuntos de dados (OSAGIE et al., 2015). Nesse contexto, as plataformas de software para gerenciamento de dados abertos - ou simplesmente ODP (do inglês *Open Data Platform*) - possuem fundamental importância na implementação dos portais de dados abertos.

Para a definição do modelo de abertura de dados governamentais e mesmo para dados privados, é fundamental a escolha da ODP pelo provedor dos dados. Sendo assim, nesta seção apresenta-se uma breve descrição sobre as principais plataformas de gerenciamento de dados abertos existentes e as características mais relevantes a serem consideradas na escolha.

São inúmeras as soluções disponíveis no mercado, variando de plataformas comerciais (proprietárias) a aquelas disponíveis com licença livre (*free software*) ou de código aberto (*open source*). Essas soluções são utilizadas para implementar milhares de portais que, por sua vez, disponibilizam centenas de milhares de *datasets* (conjuntos de dados) de fontes públicas ou privadas com diversas finalidades.

Existem vários estudos sobre o grau de utilização e as facilidades oferecidas pelas ODP. Este trabalho está baseado nos estudos de Correa (2017), Correa, Zander & Silva (2018); IBRD (2014) e Osagie et al. (2015), considerados referências na área.

Não existe uma única definição para plataforma de software para gerenciamento de dados abertos ou, simplesmente, plataforma de dados abertos. Para este relatório é adotada a seguinte definição:

Plataforma de Dados Abertos, ODP, refere-se a uma infraestrutura de software que compreende alguns componentes e interfaces para a publicação de conjuntos de dados e para o fornecimento de metadados, catálogo, armazenamento, serviços de busca e descoberta para acessar e gerenciar dados abertos (OSAGIE et al., 2015)².

² É importante distinguir que o acrônimo ODP é utilizado também em pelo menos duas outras situações ou significados distintos: a) ODP como *Open Data Platform*, é uma iniciativa homônima do setor de Tecnologia da Informação (TI) que fornece um conjunto comum de ferramentas e

Segundo os autores previamente citados, as ODP devem apresentar as seguintes características:

1. As plataformas devem ser construídas especificamente para o gerenciamento de dados abertos e ter alguma base instalada. Não basta prover uma estrutura comum de portal na Web. Como exemplo de boa prática³, o site *Data Portals*⁴ traz uma extensa lista (não exaustiva) de portais de dados abertos existentes no mundo.
2. A documentação sobre a plataforma deve ser de domínio aberto e de fácil acesso a desenvolvedores, ou comunidade de desenvolvedores.
3. A plataforma deve disponibilizar também recursos avançados de configuração.

1.1. Características desejáveis para as ODP

Para avaliação das plataformas de dados abertos quanto à disponibilidade de recursos relacionados à transparência, Osagie *et al.* (2015) estabeleceram um conjunto de características desejáveis que as ODP devem apresentar. As características são:

1. **Metadados, padrões e esquemas de formato de dados e arquivos.**
Metadados são os dados a respeito dos dados, ou seja, apresentam a estrutura dos dados (chaves, índices, colunas), as informações sobre o conjunto de dados (título, autor, assuntos, palavras-chave) e as informações de proveniência (editor, histórico de revisões, mudanças, fonte dos dados). Eles ampliam os recursos de busca e permitem a interoperabilidade entre diferentes sistemas. Os formatos mais populares para apresentação dos dados são: XML⁵, CSV⁶, JSON⁷, XLS⁸, PDF⁹ e HTML¹⁰. Os formatos legíveis por

tecnologias para a plataforma *Apache Hadoop* (Disponível em: <https://hadoop.apache.org/>. Acesso em 03/07/2019). A iniciativa congrega vários fornecedores de TI que contribuem para a plataforma em gerenciamento de big data, análise, infraestrutura e recursos de proteção do Hadoop. b) ODP como *Open Data Portal*, é a denominação utilizada pelo portal de dados abertos da União Europeia (Disponível em: <https://data.europa.eu/euodp/en/home>. Acesso em 16/07/2019). O portal contém conjuntos de dados recolhidos e publicados pela UE.

³ W3C, *World Wide Web Consortium. Data on the Web Best Practices. W3C Recommendation - 31 January 2017*. Disponível em: <https://www.w3.org/TR/dwbp/>. Acesso em 28/06/2019. A tradução oficial está disponível em <https://w3c.br/traducoes/DWBP-pt-br/index.html>.

⁴ *DataPortals*. Disponível em <http://dataportals.org/>. Acesso em 10/07/2019.

⁵ XML, *Extensible Markup Language*. Formato baseado em texto para representar informações estruturadas de documentos, dados, configuração, livros, transações, faturas e outros documentos. É uma recomendação do W3C, <https://www.w3.org/XML/Schema>.

máquina mais comuns são: CSV, XML, Geo, XLS. O formato RDF¹¹ permite que os dados possam ser consultados via SPARQL¹².

2. **Facilidades de pesquisa flexível para conjuntos de dados.** As plataformas devem oferecer recursos flexíveis, acessíveis e fáceis de serem utilizados na busca e recuperação de conjuntos de dados a partir de palavras-chave fornecidas pelos usuários. A indexação fornece uma pesquisa mais eficiente e acelera o processo.
3. **Ferramentas de mídia social, colaboração e compartilhamento.** Representa uma coleção de mecanismos que permite a interação entre os usuários. Isso inclui ferramentas de mídia social (por exemplo: Facebook, Google+, Twitter) para comunicar e colaborar, para comentar, analisar e avaliar os conjuntos de dados e compartilhar links.
4. **Guia para publicação de dados.** São diretrizes e fluxo de trabalho do processo de publicação de conjuntos de dados. Podem incluir o refinamento de dados, a separação de conjuntos de dados públicos/privados, e oferecer recursos para subir de arquivos via interface do usuário da web, API ou vinculados a um arquivo existente na web, e controle de acesso e adição de metadados ao fluxo de trabalho para subir arquivos de dados.
5. **Facilidades para coleta, federação e catalogação.** A federação permite a replicação de dados em diferentes instâncias e fornece integração perfeita entre diferentes instâncias de portal independentes, realizando uma pesquisa em várias instâncias da plataforma. O recurso de coleta permite a extração de dados abertos dos portais de dados abertos ou outras fontes de

⁶ CSV, *Comma-Separated Values*. Representa arquivos de texto de formato aberto regulamentado pelo RFC 4180 (<https://tools.ietf.org/html/rfc4180>), que faz uma ordenação de bytes ou um formato de terminador de linha, separando valores com vírgulas. Usado em software de planilha e pacotes de análise de dados.

⁷ JSON, *JavaScript Object Notation*. Padrão aberto para troca de dados entre sistemas (<https://www.json.org/>).

⁸ XLS. Formato proprietário de arquivo utilizado pelo software Excel/Microsoft em versões anteriores a 2012. Nas versões mais recentes o formato é o XLSX. (https://microsoft_excel.pt.downloadastro.com/).

⁹ PDF, *Portable Document Format*. Padrão gratuito para visualizar, imprimir, compartilhar e comentar documentos (<https://acrobat.adobe.com/br/pt/acrobat/pdf-reader.html>).

¹⁰ HTML, *HyperText Markup Language*. Linguagem de marcação utilizada na construção de páginas na Web. Documentos HTML podem ser interpretados por navegadores. (<http://www.w3c.br/Padroes/WebDesignAplicacoes>).

¹¹ RDF, *Resource Description Framework*. Modelo padrão para intercâmbio de dados na Web. Padrão mantido pelo W3C (<https://www.w3.org/RDF/>).

¹² SPARQL. Linguagem de consulta semântica para bancos de dados, permite recuperar e manipular dados armazenados no formato RDF. Padrão mantido pelo W3C (<https://www.w3.org/TR/sparql11-overview/>).

dados. Esse recurso inclui a conversão de dados para o formulário exigido pela plataforma. O catálogo descreve o mecanismo implementado para a navegação de conjuntos de dados.

6. **Ferramentas de análise de dados.** O suporte para análise de dados varia entre as plataformas e os recursos básicos de análise incluídos na maioria das plataformas. Algumas plataformas oferecem recursos mais avançados, como operações estatísticas, OLAP¹³, painéis de visualização, HTML (HyperText Markup Language). Linguagem de marcação utilizada na construção de páginas na Web. Documentos HTML podem ser interpretados por navegadores. (<http://www.w3c.br/Padroes/WebDesignAplicacoes>) ação e *widgets*¹⁴ de análise. Há plataformas que fornecem suporte para a linguagem de programação estatística R¹⁵.
7. **Ferramentas de visualização.** Suportam não apenas as visualizações básicas, como mapas e gráficos, como também serviços de mapas, comumente usados para criar visualizações complexas, como o *OpenStreetMaps*¹⁶, os mapas do Google e do Bing, entre outros; além de bibliotecas como *D3.js*¹⁷ e *Recline.js*¹⁸.
8. **Ferramentas de personalização** (no sentido de tornar a plataforma mais pessoal em vez de impessoal ou puramente profissional). Personalização é um conjunto de recursos que permite: (i) modificar a aparência do portal pelos administradores do portal (por exemplo, marca, logotipo, cores), (ii) personalizar a visualização do portal para os usuários.
9. **Ferramentas de customização** (no sentido de modificar a plataforma de acordo com os requisitos específicos dos usuários ou da organização). Customização é um conjunto de recursos que permite que os administradores do portal definam os padrões de metadados, as regras do

¹³ OLAP, *Online Analytical Processing*. Capacidade para manipular e analisar um grande volume de dados sob múltiplas perspectivas. As aplicações OLAP são usadas pelos gestores em qualquer nível da organização para lhes permitir análises comparativas que facilitem a sua tomada de decisões diárias.

¹⁴ *Widget*. Componente que pode ser utilizado em computadores, celulares, *tablets* e outros aparelhos para simplificar o acesso a um outro programa ou sistema. Eles geralmente contêm janelas, botões, ícones, menus, barras de rolagem e outras funcionalidades.

¹⁵ *The R Project for Statistical Computing*. Disponível em: <http://www.r-project.org/>. Acesso em 28/06/2019.

¹⁶ Mapa do mundo de licença aberta. Disponível em: <https://www.openstreetmap.org/>. Acesso em 29/06/2019.

¹⁷ *Data-Driven Documents*. Biblioteca JavaScript utilizada para manipular documentos baseados em dados. Disponível em: <http://d3js.org/>. Acesso em 29/06/2019.

¹⁸ *Recline.js*. Uma biblioteca para criar aplicativos de dados em *Javascript* e HTML. Disponível em: <http://okfnlabs.org/recline/>. Acesso em 29/06/2019.

portal, ativem ferramentas e recursos, bem como configurem o armazenamento de dados e os limites.

10. **Serviço de licenciamento de conjuntos de dados.** Descreve como as informações de licenciamento podem ser adicionadas aos conjuntos de dados e metadados.
11. **Acessibilidade.** Ou facilidade para acessar os dados. Uma das opções é a oferta de um conjunto de rotinas, protocolos e ferramentas para criar aplicativos de software (API). Normalmente, a API é exposta como serviços REST¹⁹ ou SOAP²⁰.
12. **Mecanismos de extensibilidade.** São expressos pelo número de recursos fornecidos nas plataformas para permitir a adaptação e a extensão (por exemplo, fornecimento de API e bibliotecas, suporte para marca do site e conectores, plugins e extensões).

1.2. Plataformas de dados abertos mais utilizadas

Nesta seção são apresentadas três das principais plataformas mais utilizadas e mais bem avaliadas pelos trabalhos de Correa, Zander & Silva (2018) e Osagie *et al.* (2015). As plataformas são: CKAN²¹, Socrata²² e ArcGIS Open Data²³, na ordem das mais utilizadas na construção de portais de dados abertos.

É importante mencionar também uma quarta iniciativa que, apesar de não ter uma ODP em si, pode ser vista como um *toolkit* para criar e manter um ODP. Trata-se da *Frictionless Data*²⁴, conjunto de padrões e ferramentas para abertura e uso de dados criado para facilitar o transporte de dados de alta qualidade entre diferentes ferramentas e plataformas para análise posterior. De lançamento mais recente, não foi analisada pelos autores acima referenciados. No entanto, numa breve análise, a *Frictionless* atende os principais critérios estabelecidos pelos autores.

¹⁹ REST, *Representational State Transfer*. Define as características para a criação de web services (serviços Web). Faz uso de JSON.

²⁰ SOAP, *Simple Object Access Protocol*. Protocolo para troca de informações estruturadas em uma plataforma descentralizada e distribuída. Faz uso de XML.

²¹ CKAN, Comprehensive Knowledge Archive Network. Disponível em: <https://ckan.org/>. Acesso em 03/06/2019.

²² Socrata, The Future of Connected Government. Disponível em: <https://www.tylertech.com/products/socrata>. Acesso em 01/07/2019.

²³ ArcGIS, Software de mapeamento e análise de dados. Disponível em: <https://www.esri.com/en-us/arcgis/about-arcgis/overview>. Acesso em 01/07/2019.

²⁴ Disponível em: <https://frictionlessdata.io/>. Acesso em 03/06/2019.

1.2.1. CKAN

CKAN (*Comprehensive Knowledge Archive Network*) é a plataforma mais utilizada por governos no mundo e a mais bem avaliada segundo os critérios de Osagie *et al.* (2015). É suportada pela *Open Knowledge Foundation*²⁵, e disponibilizada como uma estrutura *open source* para portais de dados abertos. É provavelmente a plataforma mais replicada em portais no Brasil, tanto no governo federal quanto em diversos estados e municípios.

Possui uma organização bastante objetiva, simples e intuitiva, focada nas dimensões da *Open Definition*, em especial o uso de formatos abertos, além da sua alta flexibilidade, dado que é disponibilizado em código aberto. Por outro lado, o CKAN não foi pensado para permitir seleções das bases de dados, não permite filtros e não gera visualizações por default. Isso se deve ao fato da orientação da plataforma, focada nos conjuntos de dados. No entanto, pode não ser amigável para um usuário que não tenha interesse em acessar os dados crus ou deseja uma informação de consumo mais rápido.

Por ser de código aberto, diversas modificações podem ser implementadas quando o modelo é aplicado. Na Seção 3.1 são apresentados alguns exemplos de implementação por unidades do governo brasileiro.

1.2.2. Socrata

Lançada em 2007, a plataforma foi considerada a principal solução para o gerenciamento de dados abertos antes do crescimento da CKAN. Com avaliação similar a CKAN, de acordo com Osagie *et al.* (2015), a Socrata também apresenta 11 das 12 características esperadas para uma ODP. É a segunda plataforma mais utilizada no mundo. É bastante utilizada pelo governo norte-americano, oferecendo dados como um serviço, similar à abordagem de software como serviço (SaS).

A análise de dados e visualização são as principais características da Socrata destacadas no estudo realizado por Osagie *et al.* (2015). O mesmo estudo indicou recursos limitados em relação à pesquisa e à indexação.

²⁵ Disponível em <https://okfn.org/>. Acesso em 03/06/2019.

1.2.3. ArcGIS Open Data

O ArcGIS *Open Data*, plataforma de mapeamento e análise de dados, desenvolvido pela ESRI²⁶, oferece recursos para trabalhar com dados abertos. No entanto, há poucas informações oficiais sobre os recursos disponíveis para gerenciamento de dados abertos. Não é um software de código aberto e é mais conhecido no mercado como um sistema de informações geográficas. De fato, os usuários do ArcGIS utilizam o software para publicar os dados abertos utilizados na construção dos mapas (CORREA, ZANDER & SILVA, 2018).

1.2.4. Frictionless Data²⁷

O Frictionless Data é um conjunto de ferramentas e estratégias focado em facilitar a interoperabilidade de dados, por meio da definição de padrões e critérios técnicos com o objetivo de otimizar o armazenamento e os usos de dados. Neste sentido, o projeto oferece uma série de bibliotecas, apps e instruções em diferentes linguagens para aplicar os padrões *Frictionless*, facilitando a gestão de metadados, limpeza, organização, extração, compartilhamento, entre outros recursos etc.

Uma interessante e inovadora aplicação do *Frictionless* está no uso de dados abertos para aplicações de aprendizado de máquina (*machine learning*). É o caso da OpenML²⁸, uma plataforma de serviços para aprendizado de máquina. Com uso dos recursos da *Frictionless*, o OpenML tem o objetivo de tornar o aprendizado de máquina e a análise de dados simples, acessível, colaborativa e aberta com uma divisão ideal de trabalho entre computadores e seres humanos. As pessoas podem fazer upload e compartilhar conjuntos de dados e perguntas (tarefas de predição) no OpenML que, em seguida, resolvem de forma colaborativa usando algoritmos de aprendizado de máquina.

²⁶ ESRI, Environmental Systems Research Institute. Disponível em: <https://www.esri.com/en-us/home>. Acesso em 01/07/2019.

²⁷ Disponível em: <https://frictionlessdata.io/>. Acesso em 03/06/2019.

²⁸ Disponível em: <https://www.openml.org/>. Acesso em 16/07/2019.

2. BENCHMARK INTERNACIONAL

Nesta seção são apresentados padrões internacionais e alguns modelos de abertura de dados utilizados em outros países que podem servir de inspiração para pensar um modelo de referência brasileiro. O objetivo aqui não é tentar ser exaustivo, mas trazer alguns casos de sucesso e recomendações internacionais que podem trazer *insights* sobre um modelo a ser aplicado no Brasil, com as devidas adaptações contextuais.

2.1. Recomendações do W3C²⁹

Uma das referências utilizadas para elaboração do modelo de referência é o documento de recomendação do World Wide Web Consortium (W3C), construído por especialistas internacionais da área, que relaciona 35 boas práticas de publicação e consumo de dados na World Wide Web (LÓSCIO, BURLE e CALEGARI, 2017)³⁰. Além disso, foi consultado um livreto que explica sobre os fundamentos para publicação de dados na Web, com base na recomendação Data on the Web Best Practices³¹.

2.2. Dados abertos na Europa

A União Europeia (UE), assim como outros países ou organizações, lida intensamente com grandes quantidades de dados. Esses dados podem e devem ser pronta e amplamente acessíveis e reutilizáveis sem restrições, exceto caso a legislação específica imponha limites considerando questões de segurança, privacidade e ética. Em outras palavras, as iniciativas de dados abertos na Europa adotam integralmente o conceito de dados abertos estabelecidos pela OGP.

Nesta seção são apresentados alguns dos portais administrados pela UE. A escolha assumiu os seguintes critérios: (i) facilidade de acesso; (ii) abrangência dos conjuntos de dados disponibilizados, (iii) quantidade de conjuntos de dados disponibilizados; (iv) material de apoio. Os portais escolhidos foram:

²⁹ Disponível em: <https://www.w3.org/TR/dwbp/>. Acesso em 03/06/2019.

³⁰ O documento está disponível em <https://www.w3.org/TR/dwbp/> e a tradução autorizada em: <https://w3c.br/traducoes/DWBP-pt-br/index.html>. Ambos acessados em 11/09/2019.

³¹ Disponível em: <https://ceweb.br/publicacao/livro-fundamentos-dados-web/>. Acesso em 11/09/2019.

- Portal de Dados Abertos da UE³².
- Portal Europeu de Dados³³.
- Portal do Centro de Conhecimento em Bioeconomia³⁴.
- Portal do Centro de Pesquisa Conjunta da UE³⁵.

Os portais foram desenvolvidos com a plataforma CKAN e são aderentes à especificação de perfis de aplicação DACT-AP³⁶. Esta orientação é para registro de metadados que atendam às necessidades das aplicações específicas de portais de dados na Europa, proporcionando, assim, interoperabilidade semântica com outros aplicativos com base na reutilização de vocabulários controlados estabelecidos (por exemplo, EuroVoc³⁷) e mapeamentos para vocabulários de metadados existentes (por exemplo, Dublin Core³⁸, SDMX³⁹, INSPIRE⁴⁰, entre outros)⁴¹.

2.2.1. Portal de dados abertos da União Europeia

O Portal de Dados Abertos da União Europeia “é o ponto de acesso único aos dados abertos publicados pelas instituições e outros organismos da UE”. Todos os conjuntos de dados hospedados no portal “podem ser utilizados e reutilizados livremente para fins comerciais e não comerciais”.

O portal está disponível nas 24 línguas oficiais da UE e está aberto à participação de especialistas em dados abertos que podem contribuir para orientação para o provimento, revisão e avaliação dos conjuntos de dados. O portal facilita acesso aos dados abertos

³² Disponível em: <https://data.europa.eu/euodp/pt/home>. Acesso em 31/07/2019.

³³ Disponível em: <https://www.europeandataportal.eu/pt/homepage>. Acesso em 31/07/2019.

³⁴ Disponível em: <https://data-bioeconomy.jrc.ec.europa.eu/about>. Acesso em 31/07/2019.

³⁵ Disponível em: <https://data.jrc.ec.europa.eu/>. Acesso em 31/07/2019.

³⁶ DCAT-AP, acrônimo de DCAT Application Profile, é a especificação usada pelos portais europeus para registro de metadados. A especificação está baseada no Data Catalogue Vocabulary (DCAT), especificação desenvolvida pelo W3C (disponível em: <https://www.w3.org/TR/vocab-dcat/>. Acesso em 31/07/2019).

³⁷ O EuroVoc é o tesouro multilíngue e multidisciplinar da UE, cujas palavras-chave são divididas em 21 domínios e 127 subdomínios, e são utilizadas para descrever o conteúdo da legislação europeia. Disponível em: <https://eur-lex.europa.eu/browse/eurovoc.html?locale=pt>. Acesso em 31/07/2019.

³⁸ Dublin Core é um esquema de metadados que visa descrever objetos digitais, tais como vídeos, sons, imagens, textos e sites na web. Aplicações de Dublin Core utilizam XML e RDF (Resource Description Framework). Disponível em: <https://www.dublincore.org/>. Acesso em 31/07/2019.

³⁹ SDMX, Statistical Data and Metadata eXchange, é uma iniciativa global para melhorar estatísticas e o intercâmbio de metadados. Disponível em: <https://sdmx.org/>. Acesso em 31/07/2019.

⁴⁰ INSPIRE Knowledge Base – Infrastructure for spatial information in Europe. Disponível em: <https://inspire.ec.europa.eu/>. Acesso em 31/07/2019.

⁴¹ Indicadores para avaliar o uso, a maturidade e o impacto dos dados abertos na Europa podem ser encontrados em <https://www.europeandataportal.eu/pt/dashboard#2018>. Acesso em 31/07/2019.

através de um catálogo normatizado, uma lista de ferramentas web e aplicações, editor de SPARQL, acesso mediante REST API⁴² e suporte para melhor utilizar o portal.

O portal também provê recursos para tratar dados conectados (ou ligados). São indicações de práticas para publicação e compartilhamento na web de dados estruturados sob uma ampla variedade de temas. Visa facilitar a interligação de informações oriundas de diferentes fontes proporcionando oportunidades de inovação e novos negócios. É oferecido também um catálogo com um conjunto de ferramentas para visualização dos dados.

2.2.2. Portal europeu de dados

O Portal Europeu de Dados reúne dados do setor público europeu. Os dados são também conhecidos como informações do setor público (ISP, ou no inglês, *Public Sector Information* – PSI). O portal recolhe os metadados das informações do setor público disponíveis nos portais de dados públicos em todos os países europeus. Inclui igualmente as informações relativas ao fornecimento de dados e às vantagens da sua reutilização.

O portal busca garantir a qualidade dos metadados de forma a ajudar provedores e usuários de dados a avaliar seus registros de dados considerando um conjunto padrão de critérios. Os critérios são os seguintes:

- a) Acessibilidade das distribuições,
- b) Legível por máquinas e humanos,
- c) Conformidade dos metadados em relação à especificação de perfil de aplicação DCAT-AP e
- d) Utilização de uma licença reconhecida pelo Portal Europeu de Dados.

O portal oferece conjuntos de dados nos formatos CSV e JSON. Os datasets são disponibilizados por categorias com dados das diversas regiões e estados na UE. As categorias são: Agricultura, Pesca, Silvicultura e Alimentação; Energia; Regiões e cidades; Transportes; Economia e Finanças; Questões internacionais; Governo e Setor Público; Justiça, Sistema Judiciário e Segurança Pública; Ambiente; Educação, Cultura e Desporto; Saúde; População e Sociedade; e Ciência e Tecnologia.

2.2.3. Portal do Centro de Conhecimento em Bioeconomia

O portal oferece um inventário de dados relativos à bioeconomia alimentado por diferentes fontes europeias e externas à UE. O destaque para o tema da bioeconomia se

⁴² REST API. Acrônimo de REpresentational State Transfer. É uma proposta de arquitetura para sistemas web com princípios, regras e restrições bem definidas para comunicação entre aplicações. Disponível em: <https://restfulapi.net/>. Acesso em 31/07/2019.

dá pela sua importância para o mundo atual. Tem por princípio o reconhecimento da interdependência da economia e dos ecossistemas naturais e seus recursos biológicos ao longo do espaço e do tempo.

Os datasets são disponibilizados em diversos formatos por categorias relacionadas às áreas de conhecimento englobadas pela bioeconomia. As categorias são: Agricultura; Alimentos; Bioenergia; Biotecnologia; Florestas; Lixo orgânico; Materiais e produtos químicos baseados em bioenergia; Pesca e aquicultura; Produtos e móveis de madeira baseados em papel de celulose; e Tecidos baseados em bioenergia. O portal oferece também documentos relacionados à bioeconomia e sua importância para o futuro da humanidade.

2.2.4. Portal do Centro de Pesquisa Conjunta

O portal abre acesso aos resultados de pesquisas financiadas pela UE. Apresenta 2.572 conjuntos de dados sobre diversos assuntos distribuídos em 10 áreas científicas. As áreas são: Agricultura e segurança alimentar; Ambiente e mudança climática; Energia e transporte; Inovação e crescimento; Normas e padrões; Saúde e proteção ao consumidor; Segurança e proteção; Segurança e proteção nuclear; Sociedade da informação; e União econômica e monetária. Os conjuntos de dados podem ser encontrados em diversos padrões. No entanto, nem todos os datasets são oferecidos com a padrão CSV. Aparentemente, para alguns datasets não há aderência integral à especificação DCAT-AP.

2.3. Dados Abertos na América Latina

Assim como no caso da União Europeia, é possível encontrar esforços interessantes para abertura de dados nos países da América Latina. Com menos acesso a recursos e uma articulação internacional menos robusta, boa parte dos países latinoamericanos têm desenvolvido suas políticas de dados abertos com apoio, principalmente, de mobilizações realizadas por atores não-estatais. A Iniciativa Latinoamericana por los Datos Abiertos (ILDA), a Condatos e a AbreLatam são bons exemplos de atividades que engajam tanto a sociedade civil quanto as entidades governamentais na temática da abertura de dados, formando hubs de debates qualificados sobre o assunto, possibilitando troca de experiências para solução de problemas comuns e a criação de padrões de qualidade e acesso.

3. MAPEAMENTO DOS MODELOS DE ABERTURA NO BRASIL

O levantamento a seguir compila as principais referências encontradas de modelos de abertura de dados no Brasil e em seus estados e municípios, conforme levantado por meio de diversas fontes, principalmente a partir do Open Data Index para cidades, realizado em 2018 e pelo Open Data Index para o Brasil de 2016. A lista não é exaustiva e tampouco avalia o conteúdo de cada modelo. Ressalta, portanto, as boas práticas encontradas, as limitações e os principais modelos adotados e replicados, do ponto de vista do usuário, com foco no design, na usabilidade e no grau de detalhamento da informação disponível.

3.1. Portais nacionais que fazem uso da plataforma CKAN

A seguir serão apresentados alguns exemplos de utilização⁴³ do CKAN por instituições brasileiras.

a) Portal Brasileiro de Dados Abertos⁴⁴

É o principal repositório de dados abertos do setor público brasileiro, reunindo bases de dados da maioria dos entes governamentais a nível federal, assim como de alguns Estados que aderiram ao portal. Os dados estão disponíveis em uma variedade de formatos, de acordo com o enviado pela entidade responsável, dentro da infraestrutura do CKAN. Não há visualização, filtragem ou extração de dados mais simples. As bases são disponibilizadas com licenças diferentes, mas é fácil identificar a licença usada e até navegar por tipo de licença. De forma única dentre os sites listados nessa seção, o Dados.gov.br dedica uma seção aos projetos e aplicativos montados com seus dados.

O Portal faz parte da Política de Dados Abertos consolidada no Decreto n.º 8.777, de 2016, e é parte integral da Infraestrutura Nacional de Dados Abertos.

⁴³ Há outros órgãos que usam o CKAN, como o próprio Dados.gov (governo federal), UFRN, Tesouro Nacional, FNDE, governos estaduais, entre outros.

⁴⁴ Disponível em: <http://dados.gov.br/>. Acesso em 10/06/2019.

b) Dados Abertos da Prefeitura de Belo Horizonte⁴⁵

Portal para os dados abertos publicados pela Prefeitura de Belo Horizonte que centraliza a distribuição de todos os dados de competência do poder executivo municipal (não incluindo dados sobre atividade legislativa, por exemplo). As bases são organizadas por órgão, área e etiquetas, disponíveis para download em formatos abertos. Possui uma pequena seção de visualização de alguns dos datasets no portal.

c) Dados Abertos do DF⁴⁶

Portal para reunir os dados abertos publicados pelo Governo do Distrito Federal. Adere formalmente à Infraestrutura Nacional de Dados Abertos e provê todos os dados em formatos abertos e com licença aberta. Aplica ao CKAN uma estrutura de visualização e interação de excelente design. Como ressaltado anteriormente, no entanto, não congrega todas as informações disponibilizadas em outros portais, como é o caso das remuneração de servidores disponível no Portal da Transparência. O portal é uma iniciativa que faz parte da política de Dados Abertos do DF⁴⁷.

d) Dados Abertos da Prefeitura de São Paulo⁴⁸

Portal para os dados abertos publicados pela Prefeitura de São Paulo. Também utiliza uma interface gráfica muito amigável, mas o portal não congrega um volume de informações compatível com a disponibilidade de dados de cada secretaria. Ou seja, parece faltar ainda um esforço de consolidação das informações num portal único.

e) Data.rio⁴⁹

Portal para os dados abertos publicados pela Prefeitura do Rio de Janeiro. O portal compatibiliza o CKAN com vários aplicativos para visualização de dados, além de uma seção infantil e uma série de estudos cariocas compilados pelo Instituto Pereira Passos. O portal também conta com uma seção específica para dados geográficos. A principal limitação do Data.Rio é a redundância: diversas informações

⁴⁵ Disponível em: <https://dados.pbh.gov.br/>. Acesso em 03/06/2019.

⁴⁶ Disponível em: <http://www.dados.df.gov.br/>. Acesso em 03/06/2019.

⁴⁷ O DF instituiu política de Dados Abertos, inspirada na política do Poder Executivo Federal: http://www.tc.df.gov.br/SINJ/Norma/2a90db6875624a65936a47e18e1c337b/Decreto_38354_24_07_2017.html. Acesso em 04/06/2019

⁴⁸ Disponível em: <http://dados.prefeitura.sp.gov.br/>. Acesso em 03/06/2019.

⁴⁹ Disponível em: <http://www.data.rio/>. Acesso em 03/06/2019.

são também apresentadas da mesma forma em outros links, o que dificulta a centralização do portal como fonte única.

f) Dados Recife⁵⁰

Portal para dados abertos e outros instrumentos de transparência da Prefeitura de Recife. O portal aplica o CKAN mas também provê outras informações no mesmo site, incluindo uma área de visualização, consulta livre e informações sobre os aplicativos da prefeitura. Em comparação com os portais do Rio, São Paulo e Distrito Federal, o portal de Recife é menos intuitivo e possui uma visualização menos amigável.

g) Portal da Transparência do Estado de Minas Gerais⁵¹

O Portal de Transparência de Minas é muito bem desenhado, inclui diversos tipos de informações que não são agregadas no mesmo lugar pelos outros portais desta lista e também tem uma área específica de dados abertos que aplica o CKAN. O portal direciona para outros portais com informações mais específicas e detalhadas como o DataViva⁵² e o Minas em Números⁵³. Este último, porém, é de difícil visualização e possui um design defasado em relação ao portal inicial, apesar de carregar e funcionar adequadamente.

h) Banco Central do Brasil⁵⁴

Aplica o CKAN de forma simples e direta, com apenas uma barra de pesquisa e listas das bases recentemente atualizadas e mais populares na página principal, incluindo também as bases no portal dados.gov.br (sendo uma das instituições com maior número de bases catalogadas). Reúne todos os dados publicados por 48 departamentos dentro do BCB. Sua implementação faz parte de um Plano de Dados Abertos interno, participante da Política de Dados Abertos do Governo Federal.

⁵⁰ Disponível em: <http://dados.recife.pe.gov.br/>. Acesso em 03/06/2019.

⁵¹ Disponível em: <http://www.transparencia.mg.gov.br/>. Acesso em 03/06/2019.

⁵² Disponível em; <http://dataviva.info/pt/>. Acesso em 03/06/2019.

⁵³ Disponível em: <http://www.numeros.mg.gov.br/QvAJAZfc/opendoc.htm?document=MapaResultados.qvw&host=QVS%40vm13532&anonymous=true>. Acesso em 03/06/2019.

⁵⁴ Disponível em: <https://dadosabertos.bcb.gov.br/>. Acesso em 03/06/2019.

3.2. Dados geográficos na cidade de São Paulo - Geosampa⁵⁵

O Geosampa é um portal que reúne dados geográficos da cidade de São Paulo, incluindo um mapa interativo com várias camadas de dados para visualização das informações cruzadas. É possível baixar cada camada disponível individualmente em diversos formatos abertos (.SHP, .KML, .DXF, entre outros) e baixar todos os metadados em .XML. Inclui um documento tutorial que ensina como usar a visualização e realizar os downloads. A interface é de uso simples e o tutorial disponível é de qualidade. Reúne em um lugar só todas as informações relacionadas à geografia da cidade e apresenta de uma forma interconectada, porém a visualização de dados é pesada e demora a carregar.

Potencialidades:

- Interface simples;
- Permite cruzamento de informações de forma rápida;
- Acesso fácil aos metadados e documentação.
- Há formulário para contato, permitindo interação com o usuário, ainda que não seja em tempo real.

Limites

- Visualização demora para carregar;
- Usabilidade restrita a dados geográficos, o que faz com que seja um modelo não aplicável a uma interface geral de dados abertos.

3.3. Dados de segurança pública no Rio de Janeiro - ISP Dados⁵⁶

O ISP Dados é um portal que reúne os dados de segurança do Estado do Rio de Janeiro, provendo inclusive informações por município. Os dados são atualizados com frequência e contém uma interface de visualização para quase todos os datasets. Num país que tem a criminalidade como um dos seus problemas sociais mais agravados, a disponibilização de estatísticas de qualidade num portal único e amigável é um avanço considerável, tendo sido inclusive reconhecida na avaliação do ODI para cidades (RUEDIGER e MAZOTTE, 2018).

Os dados disponíveis são detalhados e podem ser visualizados e baixados em diferentes níveis. A ferramenta de visualização é bem desenhada e carrega adequadamente. A plataforma apresenta seções para visualização de dados com recortes de análise pré definidos como grupos vulneráveis, mulher, entre outros. Por outro lado, a ferramenta

⁵⁵ Disponível em: <http://geosampa.prefeitura.sp.gov.br>. Acesso em 03/06/2019.

⁵⁶ Disponível em: <http://www.ispdados.rj.gov.br/>. Acesso em 03/06/2019.

ainda esbarra em falta de clareza no processo de download das bases, que poderia ser mais simplificado e melhor integrado com a visualização.

Potencialidades:

- Apesar de ser focada em índices de criminalidade, apresenta fundamentos replicáveis em estruturas mais gerais de dados;
- Processo de carregamento adequado;
- Alto grau de detalhamento dos dados disponibilizados;
- Acesso fácil aos metadados e documentação;
- Recortes pré definidos associados a problemáticas usuais deste tipo de dado;
- Há formulário para contato e e-mail, permitindo interação com o usuário, ainda que não seja em tempo real. Também disponibiliza números de telefone.

Limites

- Falta de clareza no processo de download dos dados;
- Portal de visualização poderia ser mais bem integrado com acesso a bases completas;
- Portal poderia usar melhor estrutura de ícones e experiência do usuário.

3.4. Portal de Transparência do Governo do Distrito Federal⁵⁷

O Portal de transparência dos recursos públicos do Governo do DF inclui as informações acerca de receitas, despesas, servidores, licitações e contratos, convênios, entre outras. Conta ainda com geolocalização dos dados e visualização da série histórica de despesas e receitas, algo que costuma ser difícil em dados orçamentários, tipicamente associados a cada Lei Orçamentária Anual (LOA).

O portal é bem organizado e intuitivo, os dados são apresentados com alto nível de detalhamento e a interface do site permite diversos filtros, além de apresentar painéis de visualização bastante informativos. No entanto, o portal não apresenta dados em formato aberto segundo a Open Definition. O Distrito Federal possui um portal específico para isso, como se verá adiante.

Potencialidades:

- Apesar de ser focada em questões orçamentárias e financeiras, apresenta fundamentos replicáveis em estruturas mais gerais de dados;

⁵⁷ Disponível em: <http://www.transparencia.df.gov.br/>. Acesso em 03/06/2019.

- Alto grau de detalhamento dos dados disponibilizados;
- Painéis com alto conteúdo informativo e visualização interativa.
- Há um link para contato/interação com o usuário, o da Ouvidoria do Governo do Distrito Federal, que oferece opções temáticas fechadas para comunicação: <https://www.ouv.df.gov.br/#/>

Limites

- Não apresenta bases completas para download em formato aberto;
- Não é completamente integrado com portal de dados abertos (algumas bases estão disponíveis, mas não todas).
- Faltam informações importantes na dimensão orçamentária (como valores pagos e restos a pagar pagos, por exemplo).

3.5. Pátio Digital - Transparência e dados abertos na secretaria de Educação de São Paulo⁵⁸

O Pátio Digital é uma iniciativa da SME (Secretaria Municipal de Educação) da prefeitura de São Paulo. Trata-se de um espaço de governo aberto com previsões de transparência, participação e criação de facilidades e aplicativos relacionados à gestão da educação no Município. O portal da iniciativa apresenta, além de abertura sistemática de dados, painéis informativos e uma estratégia de comunicação que privilegia a clareza e a informatividade. O portal é amparado pela Portaria SME 3786⁵⁹, que estabelece a obrigatoriedade de implementação do Plano Anual de Transparência Ativa e Dados Abertos da Secretaria Municipal de Educação.

Apesar de ser uma iniciativa mais ampla de governo aberto, um dos eixos de atuação do Pátio é “Transparência e Dados Abertos”, que foi utilizado para levantamento de potencialidades e limitações como modelo de referência.

Potencialidades:

- Linguagem clara e acessível;
- Consolidação de informações e aplicativos sobre a temática da educação em um lugar;
- Estratégia de comunicação aderente com o perfil do usuário;
- Há formulário para contato, permitindo interação com o usuário, ainda que não seja em tempo real.

⁵⁸ Disponível em: <http://patiodigital.prefeitura.sp.gov.br/>. Acesso em 03/06/2019.

⁵⁹ Disponível em: <http://legislacao.prefeitura.sp.gov.br/leis/portaria-secretaria-municipal-de-educacao-sme-3786-de-17-de-abril-de-2017>. Acesso em 03/06/2019.

- Redireciona para o portal de dados abertos da prefeitura, estimulando a consolidação da plataforma como centralizadora dos dados oficiais e possibilitando análises históricas.

Limites:

- Estrutura pouco replicável em estruturas mais gerais de dados, ainda que os fundamentos possam ser vistos como boas práticas;

3.6. Modelos de abertura de dados do Instituto Brasileiro de Geografia e Estatística (IBGE)

O IBGE disponibiliza seus dados através de diversos portais e subdomínios dentro do seu site, com os metadados reunidos num portal próprio. Muitos dos dados são encontrados em mais de um portal e existe uma página específica dedicada a organizar os metadados das pesquisas e estudos. Parte dos dados do IBGE também estão disponíveis no Portal Brasileiro de Dados Abertos⁶⁰.

Dentre experiências mais ou menos exitosas, que serão delineadas abaixo, o IBGE carece de uma centralização das informações, devido ao grande número de redundâncias geradas pela profusão de portais que disponibilizam dados e pesquisas. Esse caso é importante de ser analisado exatamente para se observar as possíveis perdas que se tem ao disponibilizar informações semelhantes em portais diferentes, o que pode acabar confundindo o usuário.

3.6.1. SIDRA - Sistema IBGE de Recuperação Automática⁶¹

O SIDRA é o principal sistema para se obter os dados das pesquisas mais conhecidas do IBGE, incluindo os Censos, o IPCA e a PNAD contínua. O sistema é organizado em abas, listando em cada pesquisa as tabelas disponíveis e incluindo visualizações. O sistema permite cruzar informações disponíveis, resultando em tabelas customizadas. Os dados brutos podem ser obtidos em .XLSX ou .ODS. Cada tabela pode ser baixada por inteiro ou fracionada dentro da interface do site de acordo com a preferência do usuário.

⁶⁰ Disponível em: <http://dados.gov.br/organization/instituto-brasileiro-de-geografia-e-estatistica-ibge>. Acesso em 03/06/2019.

⁶¹ Disponível em: <https://sidra.ibge.gov.br/home/ipca15/brasil>. Acesso em 03/06/2019.

Potencialidades:

- Alto grau de detalhamento dos dados disponibilizados;
- Acesso fácil aos metadados e documentação;
- Apresenta algumas bases completas para download em formato aberto;
- Há formulário para contato, permitindo interação com o usuário, ainda que não seja em tempo real.

Limites:

- Interface pouco intuitiva para quem não possui conhecimento prévio;
- Para pesquisas maiores, o portal não dá acesso direto aos microdados.

3.6.2. Séries Históricas e Estatísticas⁶²

Portal que reúne séries históricas e estatísticas, separadas por temas, fontes e níveis geográficos. O site gera visualizações com Flash e disponibiliza os downloads em XLS. Não é possível filtrar ou organizar por período, nem fracionar a base para download fora do nível geográfico.

Potencialidades:

- Alto grau de detalhamento dos dados disponibilizados;
- Visualização integrada;
- Apresenta algumas bases completas para download em formato aberto.

Limites:

- Não permite seleções, filtros ou extração de dados mais simples;
- Ferramenta de busca limitada.

3.6.3. Portal do IBGE⁶³

O site do IBGE foi reformulado em 2017 e uma nova seção de estatísticas foi implementada. É possível navegar por área, nível geográfico e por pesquisa, e cada estudo possui uma página própria disponibilizando os downloads de tabelas, dos microdados e das informações técnicas, além de visualizações das principais tabelas. Os downloads são organizados em pastas com download direto da base em XLSX ou ODS. Os dados geográficos também estão disponíveis em uma seção

⁶² Disponível em: <https://seriesestatisticas.ibge.gov.br/>. Acesso em 03/06/2019.

⁶³ Disponível em: <https://www.ibge.gov.br/estatisticas/todos-os-produtos-estatisticas.html>. Acesso em 04/06/2019.

equivalente. A iniciativa é o que mais se aproxima de uma tentativa de centralização das informações num portal único.

Potencialidades:

- Alto grau de detalhamento dos dados disponibilizados;
- Acesso fácil aos metadados e documentação;
- Apresenta bases completas para download em formato aberto;
- Facilita o entendimento de quem não tem conhecimento prévio das pesquisas do IBGE;
- Conseguimos a partir deste portal chegar em praticamente qualquer estatística disponibilizada pelo IBGE;
- O link para contato não aparece na página específica, é preciso ir à página do IBGE, no link Atendimento, onde existem diversas formas de entrar em contato, tanto formulário (com divisões por públicos e regiões - <https://www.ibge.gov.br/atendimento.html>), como por telefones.

Limites:

- Não permite seleções, filtros ou extração de dados mais simples.

3.6.4. Downloads⁶⁴

Portal de download direto das bases de dados estatísticas e geográficas do IBGE, organizadas numa árvore de pastas. Não possui metadados ou descrições dos conjuntos de dados, mas é o portal que reúne de forma mais rápida e acessível as bases do IBGE. Disponibiliza um programa próprio para ler e manipular os dados publicados, além de disponibilizar para download como CSV. Por outro lado, não é possível baixar as pesquisas inteiras, apenas as tabelas individuais e a navegação é muito pouco intuitiva.

Potencialidades:

- Acesso rápido e simples às planilhas;
- Apresenta bases completas para download em formato aberto;
- Há uma aba específica de participação social (Participe: <http://brasil.gov.br/participacao-social>), com possibilidades para consultas, ouvidoria e audiências públicas.

⁶⁴ Disponível em: <https://downloads.ibge.gov.br/>. Acesso em 03/06/2019.

Limites:

- Interface pouco intuitiva para quem não tem conhecimento prévio;
- Não é possível baixar mais de uma planilha por vez nem baixar toda uma pasta;
- Não contém metadados e descrições;
- Não permite seleções, filtros ou extração de dados mais simples;
- Funciona essencialmente como um repositório de arquivos, sendo pouco eficiente como modelo de abertura de dados.

3.7. Portal de dados do Instituto de Pesquisa Econômica Aplicada (IPEA)⁶⁵

O portal de dados do IPEA é o Ipeadata, que reúne tanto os dados publicados diretamente pelo IPEA quanto por outras entidades. É dividido em três seções: macroeconômico, regional e social, que podem ser organizadas por tema, fonte ou nível geográfico. O site permite operar alguns dos dados antes de baixá-los, sendo possível mudar a frequência, obter a variação e a média móvel, e aplicar logaritmo. Os metadados de todas as séries são disponibilizados no portal. O Ipeadata possui também uma *API* e um pacote de R no CRAN para facilitar o acesso aos dados publicados.

O IPEA possui um Plano de Dados Abertos interno, que prevê a integração das bases de dados no Portal Brasileiro de Dados Abertos.⁶⁶

Potencialidades:

- *API* e pacote do CRAN facilitam o uso por pesquisadores;
- Apresenta bases completas para download em formato aberto;
- Permite diversos tipos de operação nas bases antes de realizar o download;
- Há uma aba específica de participação social (Participe: <http://brasil.gov.br/participacao-social>), com possibilidades para consultas, ouvidoria e audiências públicas.

Limites:

- Os metadados são pouco conectados aos dados, estando disponíveis em seções diferentes;
- Estrutura de categorias limita navegação do usuário;
- Visualização poderia privilegiar melhor a experiência do usuário (*user experience*).

⁶⁵ Disponível em: <http://www.ipeadata.gov.br/Default.aspx>. Acesso em 03/06/2019.

⁶⁶ Disponível em: <http://www.ipea.gov.br/acessoainformacao/dados-abertos>. Acesso em 03/06/2019.

3.8. Dados Abertos da Câmara dos Deputados⁶⁷

Além dos modelos de abertura de dados do Poder Executivo Federal e de estados e municípios, o Brasil possui uma série de modelos aplicados a dados legislativos. A Câmara dos Deputados destaca-se por disponibilizar uma *API* como principal forma de acesso aos seus dados, e disponibiliza as bases completas para download no site. É possível baixar as bases em CSV, XLSX, ODS, XML ou JSON, tanto inteiras quanto separadas por ano, mas não é possível fracionar ou filtrar previamente.

Potencialidades:

- *API* facilita o uso por pesquisadores e desenvolvedores;
- Apresenta bases completas para download em formato aberto;
- Interface simplificada;
- Linguagem clara e acessível;
- Grande variedade de formatos para download;
- Há uma aba específica “Entre em contato”, onde além do e-mail e telefones para interação com o usuário, há uma mensagem convidando os mesmos a participar (inclusive pessoalmente), o que pode gerar maior disposição dos cidadãos a contribuir. <https://dadosabertos.camara.leg.br/contact/contact.html>

Limites:

- Não permite seleções, filtros ou extração de dados mais simples;
- Não possui nenhuma forma de visualização;
- Pode ser restrito a um público específico.

3.9. Dados Abertos do Senado Federal

O Senado Federal disponibiliza todos os seus dados e informações legislativas e administrativas no seu portal de Dados Abertos, separados por área. Disponibiliza também uma *API* no padrão da utilizada pela Câmara, porém com documentação e acesso mais confusa do que a Câmara dos Deputados, não sendo apresentada de forma direta como no caso anterior. Além da página geral de dados abertos, o Senado também disponibiliza duas plataformas de acesso a dados: o LexML, para informações legislativas e jurídicas, e o SIGA, que reúne dados orçamentários provenientes do SIAFI (Sistema Integrado de Administração Financeira).

⁶⁷ Disponível em: <https://dadosabertos.camara.leg.br/swagger/api.html#api>. Acesso em 03/06/2019.

3.9.1. Página de Dados Abertos do Senado Federal⁶⁸

A página de Dados Abertos do Senado apresenta informações de cada Senador e seus gabinetes, funcionários e gastos, assim como os dados relativos à administração do Senado e à composição dos órgãos, conselhos e comissões da Casa.

Potencialidades:

- Linguagem clara e acessível;
- Apresenta bases completas para download em formato aberto;
- Design e navegação simples e intuitiva;
- Há diversas possibilidades de interação já na página inicial: comunicação sobre o portal em si ("[Informe problemas no portal](#)"); ouvidoria (<https://www12.senado.leg.br/institucional/falecomosenado/formulario>); fale com os senadores, onde há um PDF com o e-mail de cada um (http://www.senado.leg.br/transparencia/LAI/secrh/parla_inter.pdf); e redes sociais (Facebook, Twitter, Youtube e Flickr); e uma ferramenta específica E-cidadania, para envio de propostas; além do E-sic estar apresentado neste mesmo conjunto.

Limites:

- Muitas das categorias são dependentes da estrutura de *API* e arquivos JSON ou XML, que não são de fácil compreensão para o público geral;
- Não contém metadados e disponibiliza apenas descrições curtas.

3.9.2. LexML Brasil⁶⁹

Reúne leis, normas, pareceres e acórdãos de diversas instituições legislativas e jurídicas do país, incluindo o Congresso Nacional, as últimas instâncias do Judiciário e algumas Assembleias Legislativas e Câmaras Municipais. O acesso pode ser feito através de pesquisa direta no site ou da *API* do Senado. O código-fonte do projeto está disponível no portal de Dados Abertos do Senado, junto com listas de vocabulário e instruções de pesquisas através de *CQL (Contextual Query Language)*. Não é possível realizar o download de bases completas.

Potencialidades:

- *API* facilita o uso por pesquisadores e desenvolvedores;

⁶⁸ Disponível em: <https://www12.senado.leg.br/dados-abertos>. Acesso em 03/06/2019.

⁶⁹ Disponível em: <https://www.lexml.gov.br/>. Acesso em 03/06/2019.

- Alto grau de detalhamento dos dados disponibilizados;
- Em FAQ, há a possibilidade de comunicação com o usuário por meio de envio de formulário, e redes sociais (Twitter e Youtube).

Limites:

- Não é permitido fazer download de mais de um arquivo por vez;
- Não apresenta bases completas para download em formato aberto.

3.9.3. SIGA Brasil⁷⁰

Reúne dados orçamentários do SIAFI disponíveis para acesso através de painéis ou relatórios. Os painéis são ferramentas de visualização dos dados, estando disponíveis o Painel Cidadão, voltado para o público geral com informações simplificadas, o Painel Emendas, para a execução de emendas parlamentares, e o Painel Especialista, com visualizações detalhadas e filtros disponíveis para quase todos os campos.

Para se obter dados brutos, é necessário usar os Relatórios, ferramenta onde se realiza uma consulta específica aos dados do SIAFI, podendo ser exportada para diversos tipos de arquivo incluindo XLSX, CSV e PDF. Devido às atualizações constantes do SIAFI, o SIGA disponibiliza dados atualizados em tempo real. Não é possível baixar a base inteira e consultas muito abrangentes podem causar o travamento do sistema, e cada consulta só pode ser realizada em um universo por vez, o que impossibilita consultas em mais de um ano orçamentário.

Potencialidades:

- A pesquisa permite a seleção de dados bastante refinados, de acordo com a preferência do consultante;
- Alto grau de detalhamento dos dados disponibilizados.

Limites:

- Não apresenta bases completas para download em formato aberto;
- Falta de clareza com relação à disponibilização dos metadados;
- Ferramenta apresenta problemas de funcionamento;
- O uso do sistema é complicado para públicos mais gerais;
- Não há um canal específico de comunicação do Siga Brasil.

⁷⁰ Disponível em: <https://www12.senado.leg.br/orcamento/sigabrasil>. Acesso em 03/06/2019.

3.10. Portal da Transparência da Controladoria-Geral da União (CGU)⁷¹

O Portal da Transparência da CGU é a principal ferramenta de transparência do Governo Federal, tendo passado por uma renovação em 2018, que incluiu a completa integração das bases de dados governamentais disponibilizadas nas consultas. É um dos portais mais utilizados pelos cidadãos brasileiros, com uma média mensal de 1,6 milhão de visitas (sessões) nos últimos cinco anos. Disponibiliza informações desde dados orçamentários, até vínculos de pessoas físicas com órgãos federais, a exemplo de servidores públicos e beneficiários de programas sociais. Os dados podem ser acessados através de painéis, com visualizações interativas ou consultas, onde se pode baixar os dados brutos. Entretanto, o download das consultas é limitado a apenas 1000 registros, o que impõe uma forte restrição à capacidade do usuário de acessar os dados disponíveis neste portal, sendo necessário usar os filtros disponíveis para selecionar um universo pequeno suficiente para ser baixado. É possível baixar os dados completos clicando no item “Dados Abertos” no final da página ou através da API^{72,73}. O sistema de visualizações tanto nos painéis quanto nas consultas é intuitivo e bastante flexível.

O portal disponibiliza também aos usuários diversas consultas via API - Interface de Programa de Aplicativos (do inglês, *Application Programming Interface*), incluindo: benefícios sociais, contratos, convênios, empresas, licitações, sanções, servidores, entre outras. As APIs do Portal retornam os mesmos dados apresentados em tela, com a vantagem da flexibilidade característica da ferramenta, ao permitir que máquinas ou robôs obtenham acesso automático aos dados para a criação de aplicativos, análises de dados e outros usos.

Há formas dinâmicas de interação com os usuários, a exemplo do “Fique Olho”, que podem fomentar o controle social dos recursos públicos. Por meio dessa ferramenta, o cidadão pode indicar, por exemplo, se uma empresa funciona no endereço indicado, se o objeto de um convênio foi entregue e está compatível com o valor investido, se uma pessoa física recebe um benefício indevidamente, etc. O conjunto das informações é analisado pela CGU, podendo auxiliar em futuras ações de controle.

Potencialidades:

- API facilita o uso por pesquisadores e desenvolvedores;
- Alto grau de detalhamento dos dados disponibilizados;
- Design e navegação simples e intuitiva;

⁷¹ Disponível em: <http://www.portaltransparencia.gov.br/>. Acesso em 03/06/2019.

⁷² Disponível em: <http://www.portaltransparencia.gov.br/api-de-dados>. Acesso em 05/06/2019.

⁷³ Disponível em: <http://www.portaltransparencia.gov.br/download-de-dados>. Acesso em 04/06/2019.

- Apresenta bases completas para download em formato aberto;
- Painéis com alto conteúdo informativo e visualização interativa;
- Link para a ouvidoria, onde se pode registrar denúncias, fazer perguntas, etc. <https://sistema.ouvidorias.gov.br/publico/Manifestacao/RegistrarManifestacao.aspx?ReturnUrl=%2f>

Limites:

- Limite de download muito restritivo nas seleções;
- Carregamento lento das informações em alguns casos.

3.11. TABNET - Departamento de Informática do SUS (DATASUS)⁷⁴

O TABNET é o portal de acesso aos dados de saúde publicados pelo SUS. Os dados podem ser acessados através da navegação do site, porém a melhor forma de acesso é através do programa gratuito TabWIN. Navegando pelo site, é possível baixar os dados em CSV ou copiar para o Excel. Os dados disponíveis são bastante detalhados e é possível fazer seleções refinadas para baixar. O site conta com estatísticas de acesso, podendo ser visualizado o histórico de grupo de informações mais acessados.

Potencialidades:

- Alto grau de detalhamento dos dados disponibilizados;
- Notas técnicas facilmente acessíveis e detalhadas;
- As bases são devidamente padronizadas, permitindo comparabilidade dos dados.

Limites:

- Dependente do programa para Windows, o TabWIN, para operações mais avançadas, o que pode limitar consideravelmente o acesso;
- O uso do site é pouco intuitivo para quem não possui conhecimento prévio dos indicadores;
- Não apresenta bases completas para download em formato aberto.

⁷⁴ Disponível em: <http://www2.datasus.gov.br/DATASUS/>. Acesso em 03/06/2019.

3.12. Repositório de Dados Eleitorais do Tribunal Superior Eleitoral (TSE)⁷⁵

Também no judiciário federal existem experiências de abertura de dados, dentre as quais cabe destacar o repositório de dados eleitorais do TSE. O portal é a principal fonte de informações primárias sobre os resultados de eleições no Brasil. Navegando pelas categorias, é possível obter as bases completas, comprimidas em ZIP, no formato CSV. Contém os dados de candidatos, partidos e resultados, além das prestações de contas individuais e partidárias. A partir de 2012, também disponibiliza informações sobre pesquisas eleitorais conduzidas por ciclo eleitoral e UF.

Potencialidades:

- Alto grau de granularidade dos dados disponibilizados;
- Apresenta bases completas para download em formato aberto;
- Há uma aba “Fale conosco”, com formulário: <http://datasus.saude.gov.br/fale-conosco>

Limites:

- Não permite seleções, filtros ou extração de dados mais simples;
- Não possui nenhuma forma de visualização.

3.13. Dados Abertos FEE (RS)⁷⁶

A Fundação de Economia e Estatística do Rio Grande do Sul divulga um portal de dados abertos que possui dados diversos em formato aberto e licença aberta, mas a interface é limitada a uma árvore de categorias, sem opção de pesquisa. Permite seleção de recortes dentro da base de dados antes de baixar.

O Governo do Rio Grande do Sul, por intermédio da Subchefia de Ética, Controle Público e Transparência da Casa Civil, inseriu a temática de dados abertos no escopo da implementação da Lei de Acesso à Informação (Lei nº 12.527) e no contexto das políticas públicas voltadas à transparência e ao controle social. Política estabelecida por meio do Decreto nº 53.523/2017. Foi publicada cartilha e manuais temáticos no site: <https://dados.rs.gov.br/about>

⁷⁵ Disponível em: <http://www.tse.jus.br/eleicoes/estatisticas/repositorio-de-dados-eleitorais-1/repositorio-de-dados-eleitorais>. Acesso em: 04/06/2019.

⁷⁶ Disponível em: <https://dados.fee.tche.br/index.php>. Acesso em 03/06/2019.

Potencialidades:

- Acesso rápido e simples às planilhas;
- Apresenta bases completas para download em formato aberto;
- Há um formulário para contato (<https://www.fee.rs.gov.br/contato/>), e telefones também.

Limites:

- Os metadados são pouco conectados aos dados, estando disponíveis em seções diferentes;
- Não é possível baixar mais de uma planilha por vez nem baixar toda uma pasta.

3.14. Portal da Transparência do Ministério Público Federal (MPF)

⁷⁷

O Portal reúne os dados de transparência orçamentária do MPF, incluindo contratos, licitações e dados de pessoal, e também publicações e registros judiciais da sua alçada. A navegação é clara e intuitiva, e todos os dados são disponibilizados por ano e mês em planilhas de Excel ou relatórios em PDF.

Estão disponibilizados no site mais de 3 milhões de registros judiciais, separados em autos judiciais, procedimentos extrajudiciais, inquéritos policiais, manifestações extrajudiciais e documentos gerais. As consultas nesta base possuem filtros detalhados e é possível baixar consultas de até 5000 registros em bloco, exportando para XLS e PDF.

Potencialidades:

- Design e navegação simples e intuitiva;
- Permite download em bloco de consultas nas bases jurídicas de até 5000 registros;
- Há uma “Sala de atendimento ao cidadão” (<https://www.fee.rs.gov.br/contato/>), com formulário para envio de comunicação; além de redes sociais (facebook, Twitter, Flickr, Youtube e Instagram).

Limites:

- Não permite seleções, filtros ou extração de dados mais simples;
- Não possui nenhuma forma de visualização;
- Não apresenta bases completas para download em formato aberto.

⁷⁷ Disponível em: <http://www.transparencia.mpf.mp.br>. Acesso em 11/06/2019.

4. CONSOLIDAÇÃO DOS RESULTADOS

Este documento apresentou uma miríade de modelos de referência nacionais e internacionais de abertura de dados, bem como as plataformas tipicamente utilizadas para hospedar esses dados on-line. Cada modelo possui uma série de potencialidades que podem vir a ser exploradas na elaboração de um padrão de abertura replicável, mas também apresentam limitações específicas de cada contexto de aplicação.

Ainda que o documento não se proponha a uma análise exaustiva e totalmente representativa de todo o universo de modelos de abertura existentes, o levantamento realizado permite traçar algumas inferências acerca dos maiores potenciais e das limitações mais críticas nos portais analisados. Por exemplo: a questão da completude das bases disponibilizadas para download, apesar de ser um critério explícito da *Open Definition*, está presente em pouco mais da metade dos portais analisados.

Os portais analisados, em média, apresentam um bom grau de detalhamento das informações disponibilizadas, buscam fazer esforços de consolidação em um local só e apresentam boas práticas de *User Experience*. Por outro lado, questões mais técnicas como disponibilização de API, agilidade no carregamento e variedade de formatos disponíveis foram ressaltados como potencialidades em poucos portais.

Em relação às limitações, verificou-se que, apesar de grande parte dos portais se preocupar em disponibilizar portais de extração e visualização de dados além de apenas as bases de dados cruas, essas ferramentas eram muitas vezes restritas a um público muito específico, ou com recortes limitados, com baixa capacidade de cruzamento de informações. Além disso, muitos ainda apresentam problemas de falta de clareza e pouca disponibilidade de metadados.

Um dos grandes desafios de pensar um modelo de referência para abertura de dados é como atender ao mesmo tempo interesses e desejos diferentes de acordo com o perfil do usuário de dados. A análise dos modelos de abertura deixa claro que as iniciativas mapeadas privilegiam mais ou menos a capacidade de dialogar com um determinado perfil do usuário, o que acaba limitando as possibilidades para outros perfis.

Sendo assim, isso sugere que o modelo ideal deve permitir formas diferentes de acesso ao mesmo dado, entregando tanto uma informação previamente trabalhada e de consumo mais rápido para usuários que buscam tal tipo de conteúdo (cidadão comum, jornalistas, etc) quanto às bases de dados compatíveis com a Open Definition (OKFN, 2006), facilitando o uso dos dados crus para análises, desenvolvimento de soluções, etc. Dessa forma, o modelo atenderia de forma mais adequada um espectro maior de tipos de usuário.

REFERÊNCIAS

CORREA, A. S. Uma arquitetura de referência colaborativa para estruturação de dados abertos governamentais. 2017. 222 f. Tese (Doutorado em Sistemas Digitais). Escola Politécnica, Universidade de São Paulo, São Paulo. Disponível em: <http://www.teses.usp.br/teses/disponiveis/3/3141/tde-28062017-101825/pt-br.php>, acesso em 2019-06-28.

CORREA, A. S.; ZANDER, P. O.; SILVA, F. C. Investigating open data portals automatically: a methodology and some illustrations. In Proceedings of the 19th Annual International Conference on Digital Government Research: Governance in the Data Age - dg.o'18. Article No. 82. ISBN: 978-1-4503-6526-0. doi>10.1145/3209281.3209292.

EDP, Portal Europeu de Dados Abertos. Disponível em <https://www.europeandataportal.eu/pt/>, acesso em 07-03-2019.

IBRD/The World Bank, International Bank for Reconstruction and Development. Technical Assessment of Open Data Platforms for National Statistical Organisations, 2014. Disponível em: <http://documents.worldbank.org/curated/pt/744241468334210686/pdf/928100WP0Techn0Box385378B000PUBLIC0.pdf>, acesso em: 2019-06-28.

LÓSCIO, Bernadette Farias; BURLE, Caroline; CALEGARI, Newton. Data on the Web Best Practices. 2017. Disponible en: <<https://www.w3.org/TR/dwbp/>>. Accessed on May 29th, 2019.

OKFN - OPEN KNOWLEDGE FOUNDATION (2006). *Open Definition*. Available at: <https://okfn.org/>. Accessed on May 29th, 2019.

OSAGIE, E.; MOHAMMAD, W.; STASIEWICZ, A.; HASSAN, I. A.; PORWOL, L.; OJO, A. State-of-the-art Report and Evaluation of Existing Open Data Platforms – Deliverable 2.1, VERSION 1.1, 19/05/2015. Raising Open and User-friendly Transparency – Enabling Technologies for Public Administrations, Project number 645860, H2020-INSO-2014. Disponível em: <https://project.routetopa.eu/deliverable-2-1-state-of-the-art-report-and-evaluation-of-existing-open-data-platforms-now-available/>, acesso em 2019-06-28.

RUEDIGER, M. A.; MAZZOTTE, N. (Ed.). *Índice de dados abertos para cidades*. Rio de Janeiro: FGV DAPP, 2018.

TYGEL, A.; AUER, S.; DEBATTISTA, J.; ORLANDI, F.; CAMPOS, M. L. M. Towards Cleaning-up Open Data Portals: A Metadata Reconciliation Approach. 2015. Disponível em: [ArXiv:1510.04501 \[cs.IR\]](https://arxiv.org/abs/1510.04501), em 03-07-2019.

EXPEDIENTE

FGV DAPP

Diretor

Marco Aurelio Ruediger

Diretoria de Análise
de Políticas Públicas (DAPP)
DAPP Contato:
(21) 3799-4300
www.dapp.fgv.br
dapp@fgv.br

EQUIPE DE EXECUÇÃO

Coordenação

Wagner Oliveira

Pesquisadores

Leonor Jungstedt

EDITORIAL

Projeto Gráfico

Luís Gomes



Instituição de caráter técnico-científico, educativo e filantrópico, criada em 20 de dezembro de 1944 como pessoa jurídica de direito privado, tem por finalidade atuar, de forma ampla, em todas as matérias de caráter científico, com ênfase no campo das ciências sociais: administração, direito e economia, contribuindo para o desenvolvimento econômico-social do país.

OPEN KNOWLEDGE BRASIL (OKBR)



A Open Knowledge Brasil (OKBr) é uma Organização da Sociedade Civil (OSC) sem fins lucrativos e apartidária, fundada em 2013, que representa oficialmente a Open Knowledge Internacional no país. Sua missão é promover o conhecimento livre nos diversos campos da sociedade, além de buscar, na esfera política, tornar a relação entre governo e sociedade mais transparente e a participação social, mais efetiva e aberta.

Camille Moura

Pesquisa e Operações

camille@ok.org.br

CONTROLADORIA-GERAL DA UNIÃO (CGU)



A CGU busca promover o aperfeiçoamento e a transparência da Gestão Pública, a prevenção e o combate à corrupção, com participação social, por meio da avaliação e controle das políticas públicas e qualidade do gasto.

Marcelo de Brito Vidal

Coordenador-Geral de Governo Aberto e Transparência
marcelo.vidal@cgu.gov.br

Aureliano Vogado Rodrigues Junior

Coordenador-Geral de Governo Aberto e Transparência Substituto
aureliano.junior@cgu.gov.br

Ministério da Ciência, Tecnologias, Inovações e comunicações (MCTIC) Centro de Tecnologia da Informação Renato Archer (CTI)



O Centro de Tecnologia da Informação Renato Archer - CTI é uma unidade de pesquisa do Ministério da Ciência, Tecnologia, Inovações e Comunicações (MCTIC). Foi inaugurado em 1982 e, desde então, atua na pesquisa e no desenvolvimento em tecnologia da informação.

Jarbas Lopes Cardoso Junior
Tecnologista Senior
jarbas.cardoso@cti.gov.br

INSTITUTO DE ESTUDOS SOCIOECONÔMICOS (INESC)



Fundado em 1979, o Inesc é uma organização não governamental, sem fins lucrativos, não partidária e com sede em Brasília. Atua junto a sociedade civil e movimentos sociais na defesa da democracia e dos direitos, com foco no monitoramento do orçamento público. <https://www.inesc.org.br/>

Carmela Zigoni
Assessora Política
carmela@inesc.org.br

Centro de Estudos sobre Tecnologias Web (Ceweb.br) do Núcleo de Informação e Coordenação do Ponto BR (NIC.br)



Os objetivos do Ceweb.br são fomentar a inovação na Web; estimular o melhor uso da Web; mostrar o potencial da Web a diversos segmentos da sociedade e contribuir para a evolução da Web. Os resultados do Centro serão consolidados em plataformas, cursos, estudos, recomendações e publicações. Os mecanismos para alcançar esses resultados são acordos de cooperação e espaços de discussão e colaboração permanente. <https://ceweb.br>

Caroline Burle
Relações Institucionais e Internacionais
cburle@nic.br