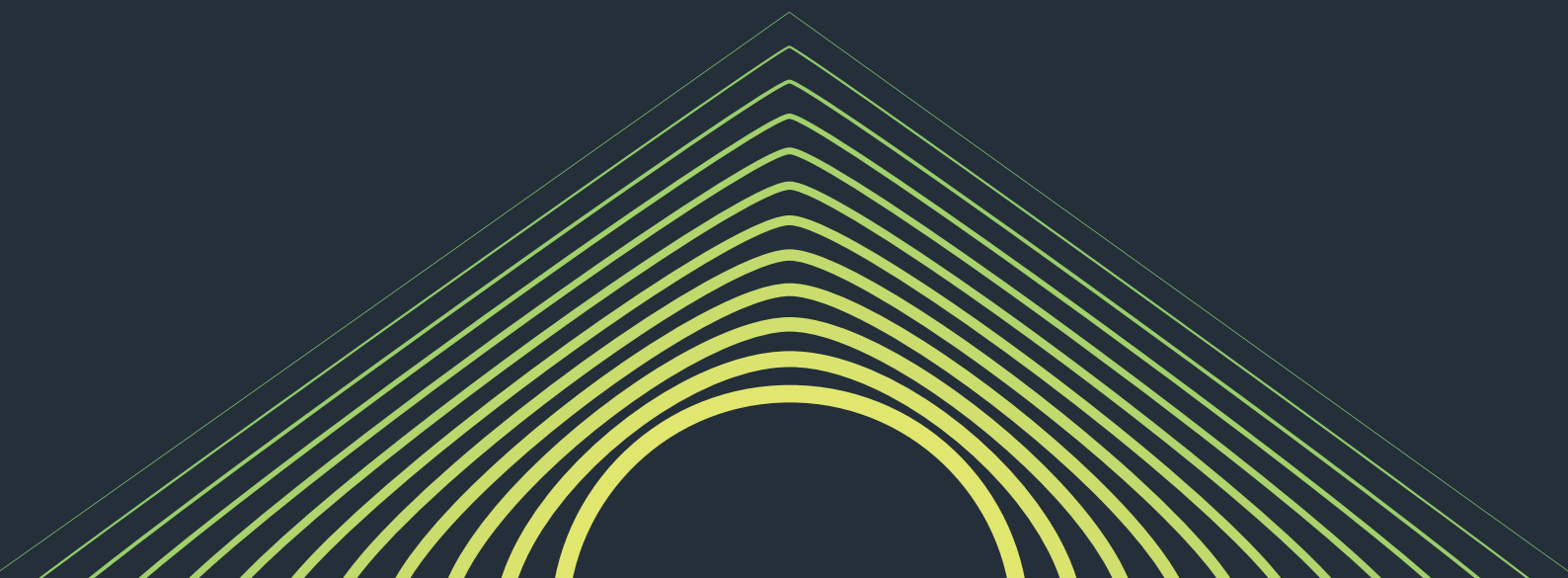


ÍNDICE DE DADOS ABERTOS PARA O BRASIL



FGV DAPP



ÍNDICE DE DADOS ABERTOS PARA O BRASIL

ABRIL | 2017





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.

Escritório

Praia de Botafogo 190, Rio de Janeiro
RJ - CEP 222509000
Caixa Postal 62.591 CEP 22257-970
Tel (21) 3799-5498 | www.fgv.br

Presidente Fundador

Luiz Simões Lopes

Presidente

Carlos Ivan Simonsen Leal

Vice-Presidentes

Sergio Franklin Quintella,
Francisco Oswaldo Neves Dornelles e
Marcos Cintra Cavalcante de Albuquerque



Diretor

Marco Aurelio Ruediger

Diretor-Executivo

Ariel Kogan

DAPP

(21) 3799-4300
www.dapp.fgv.br | dapp@fgv.br

Equipe

Comunidade OKBr

EXPEDIENTE

EQUIPE DE EXECUÇÃO

Pesquisadores

Amaro Grassi
Andressa Falconieri
Bárbara Barbosa
Janaina Fernandes
Wagner Oliveira

Projeto gráfico

Arielle Asensi
Humberto Ferreira
Luis Gomes
Rebeca Liberatori Braga



SUMÁRIO

- 04 SOBRE A FGV/DAPP
- 04 SOBRE A OPEN KNOWLEDGE BRASIL
- 05 SOBRE A PARCERIA FGV/DAPP E OKBR
- 06 O ÍNDICE DE DADOS ABERTOS GLOBAL
- 07 RESULTADOS
- 12 SOBRE AS BASES DE DADOS
- 28 CONCLUSÃO
- 21 FONTES

SOBRE A FGV/DAPP

A Diretoria de Análise de Políticas Públicas da Fundação Getúlio Vargas (FGV/DAPP) é um centro de pesquisa social aplicada voltado à inovação para políticas públicas, produzindo análise de ponta com uso intensivo de redes sociais e conhecimento interdisciplinar. Tem como missão aprimorar a gestão pública brasileira e qualificar o debate público na sociedade em rede, por meio da transparência e do diálogo entre o Estado e a cidadania.

A FGV/DAPP desenvolve uma agenda de pesquisa a partir de metodologia própria de análise, aprimorada desde a sua criação, em 2012, e que reúne métodos quantitativos e qualitativos tradicionais com recursos inovadores de processamento e análise de bancos de dados públicos e de redes sociais. E disponibiliza à sociedade, afinal, um conjunto de ferramentas de visualização e análise de dados de fácil compreensão e acesso.

As análises de políticas públicas são produzidas por uma equipe interdisciplinar e diversificada, de formação em áreas como Sociologia, Ciência Política, Antropologia, Linguística, Economia, Administração Pública, Relações Internacionais, Estatística, Matemática e Comunicação Social, aliadas à Tecnologia da Informação e ao Design.

SOBRE A OPEN KNOWLEDGE BRASIL

A Open Knowledge Brasil (OKBr) é uma Organização da Sociedade Civil (OSC) sem fins lucrativos e apartidária, fundada em 2013. A OKBr representa no Brasil a Open Knowledge Internacional a partir de um acordo firmado também em 2013.

A OKBr utiliza e desenvolve ferramentas cívicas, faz análises de políticas públicas, trabalha com jornalismo de dados e promove o conhecimento livre para tornar a relação entre governo e sociedade mais transparente e para que haja uma participação política mais efetiva e aberta.

A OKBr acredita em um mundo onde as informações estão livres e disponíveis para permitir decisões e escolhas mais conscientes e inteligentes. Um mundo onde a informação e o conhecimento estão ao alcance de todos.

SOBRE A PARCERIA FGV/DAPP E OKBR

Um dos principais eixos de atuação da FGV/DAPP é a promoção da Transparência, entendida não apenas como a divulgação de dados públicos, mas como a capacidade de facilitar o entendimento das informações provenientes desses dados. A partir da integração do uso de tecnologia e design, a FGV/DAPP divulga informações de maneira simples e acessível sobre diversos temas, como o destino dos recursos públicos, a origem das doações de campanha eleitoral, a situação dos serviços públicos (educação, saúde, transporte, segurança), entre outros.

Tendo em vista seu papel de destaque na promoção dessa agenda, a FGV/DAPP firmou na segunda metade de 2016 uma parceria com a Open Knowledge Foundation no Brasil (OKBr) com o objetivo de contribuir para o levantamento anual das informações do Índice de Dados Abertos nacional e para detalhar e aprimorar a metodologia do índice aplicado em nível subnacional – cidades.

O ÍNDICE DE DADOS ABERTOS GLOBAL

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 publicidade¹. A Open Knowledge Foundation desenvolveu o Índice de Dados Abertos (ODI) global para comparar países em termos da sua capacidade de abrir dados aos cidadãos, à mídia e à sociedade civil. As informações que compõem o índice são levantadas por meio de uma plataforma *crowd-sourced*, ou seja, a partir das contribuições de agentes locais ligados à rede da OKFN em cada país, e depois revisadas por especialistas em dados abertos em cada país, gerando como resultado final um ranking de países.

O objetivo do índice é avaliar o estado da política de dados abertos de cada país, levando em consideração todas as suas características: quais dados são divulgados sobre determinado assunto, como eles são divulgados em relação ao acesso, ao formato, à granularidade, aos custos e ao ritmo de sua atualização. O índice avalia diversas dimensões, tais como finanças públicas, estatísticas socioeconômicas, legislativas e eleitorais, indicadores ambientais e informações administrativas e geográficas.

O índice oferece, assim, um parâmetro de referência para a capacidade de fornecer dados abertos para todos os países, apresentando tal informação de forma clara e de fácil entendimento e usabilidade, algo que está em consonância com a missão de Transparência da FGV/DAPP.

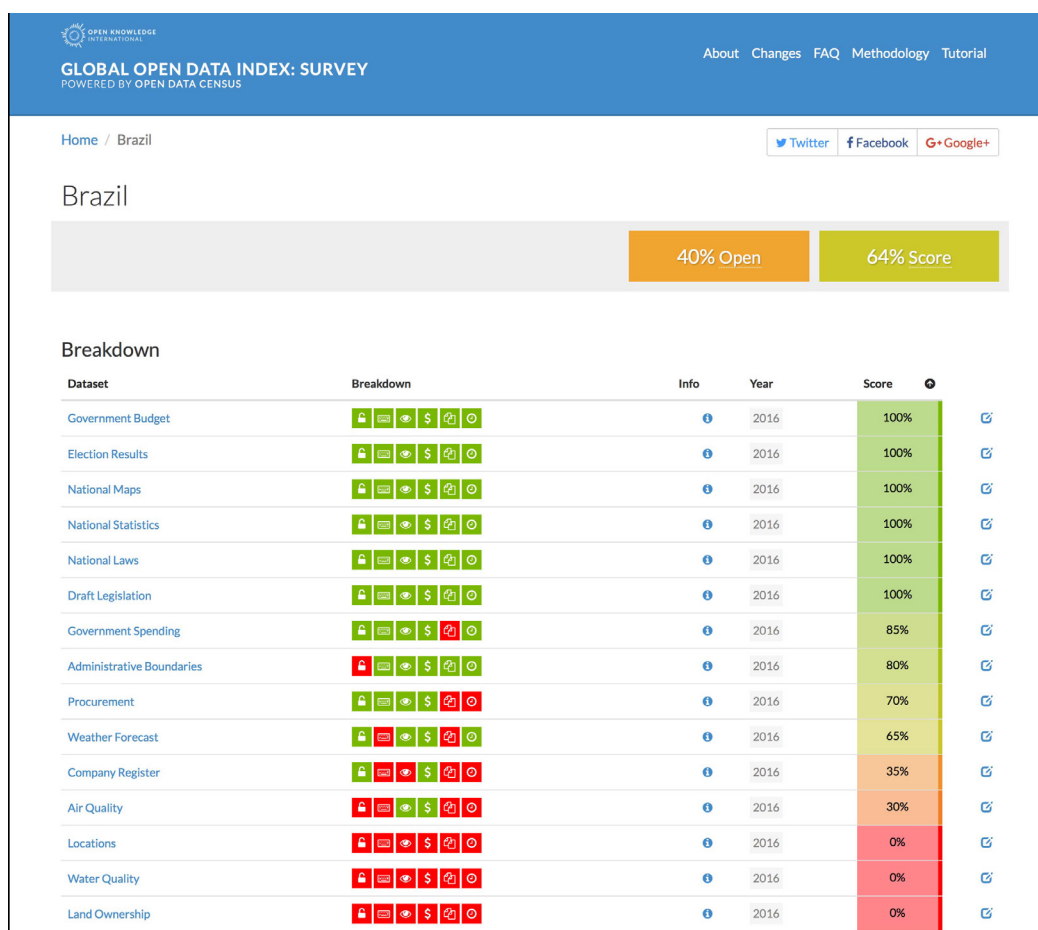
Este índice já existe para os anos de 2013, 2014 e 2015 e agora está sendo construído para o ano de 2016, usando uma metodologia já aprimorada pela experiência acumulada ao longo desse período. A FGV/DAPP, por meio da parceria firmada com a Open Knowledge do Brasil (OKBr), foi a entidade brasileira responsável pelo levantamento das informações que compõem o índice nacional nesta edição.

¹ Veja a “Open Definition” para mais detalhes em: <http://opendefinition.org/>

RESULTADOS

O resultado da avaliação no Brasil indicou um escore de **64%**. O país se posiciona em **8º** no ranking, empatado com os Estados Unidos e a Letônia, e o **primeiro** dentre os países da **América Latina**. No total, foram **15 dimensões** analisadas de acordo com **nove critérios**². A figura a seguir mostra como foram avaliadas cada uma das dimensões. Do conjunto analisado, seis — ou 40% — receberam a pontuação máxima: orçamento público, resultados eleitorais, mapas nacionais, estatísticas socioeconômicas, leis em vigor e atividade legislativa.

FIGURA 1 - AVALIAÇÃO DAS DIMENSÕES DE DADOS ABERTOS NO BRASIL³



2 Se a base de dados existe e está completa de acordo com as características requeridas (1). Se é facilmente encontrada (2) e está disponível online (3) para qualquer cidadão de forma gratuita (4) e sem necessidade de identificação (5), se está atualizada (6), se está apresenta uma boa usabilidade (7), se disponibiliza download em formatos abertos (8) e se possui licença de dados abertos (9).

3 Os nomes das dimensões do índice foram traduzidos para o português pelos autores deste documento, levando em consideração a capacidade explicativa que o nome tem em relação ao conteúdo da dimensão. Ou seja, não são traduções literais em grande parte dos casos. Seguem os nomes adotados em português na ordem apresentada na figura 1: Orçamento Público, Resultados Eleitorais, Mapas Nacionais, Estatísticas Socioeconômicas, Leis em vigor, Atividade Legislativa, Gastos Públicos, Limites Administrativos, Compras Públicas, Previsões do Tempo, Registro de Empresas, Qualidade do Ar, Localizações, Qualidade da Água, Propriedade da Terra.

As dimensões relacionadas a **Localizações, Qualidade da Água e Propriedade da Terra** receberam 0% porque não foram encontrados os dados referentes a elas. Como nestes três casos nenhum dos demais critérios puderam ser avaliados, as próximas análises deste relatório vão considerar não 15, mas 12 *datasets*.

Para analisar o estado dos dados abertos do Brasil com ano-base em 2016, adotamos uma metodologia de contagem dos problemas encontrados em cada dimensão e categorizamos os problemas em dois tipos: usabilidade e processo. Na primeira categoria, estão listados problemas relacionados à dificuldade de uso do dado em si e à capacidade de transformação deste em informação. Como questões de processo, identificamos: dificuldades de se acessar o dado e omissão/inadequação da licença. A tabela a seguir apresenta a síntese de problemas em cada dimensão, ordenadas da mais para a menos urgente.

FIGURA 2 - CATEGORIZAÇÃO DOS PROBLEMAS EM USABILIDADE E PROCESSO

USABILIDADE	PROCESSO
Dataset incompleto	Restrição de acesso
Dataset desatualizado	Dificuldade de localizar dados
Indisponibilidade de formato aberto	Download da base completa indisponível
Dificuldade de trabalhar dados	Licença não transparente

Em geral, encontramos mais problemas de usabilidade que de processo. A frequência de problemas de usabilidade é maior nas dimensões de Qualidade do Ar (aparece quatro vezes), **Previsão do Tempo** e **Registro de Empresas** (três vezes cada). Os problemas de processo aparecem mais vezes em **Qualidade do Ar** e **Registro de Empresas**, três vezes em cada um deles.

A tabela da Figura 3 apresenta a frequência de problemas identificados por dimensão do índice. Dentre os 12 *datasets* avaliados, três não apresentaram qualquer gargalo, seja em relação à usabilidade, seja em relação ao seu processo de disponibilização: **Estatísticas Socioeconômicas**, **Mapas Nacionais** e **Orçamento Público**. No extremo oposto, a dimensão **Qualidade do Ar** apresenta problemas em quase todos os critérios da avaliação⁴.

Todas as bases são providenciadas por órgãos de nível federal, e indicamos na última coluna o poder do órgão responsável pela base avaliada. A maior parte dos provedores de dados enquadra-se como órgãos da esfera do Poder Executivo — oito das 12 bases analisadas estão associadas a este poder. Três bases de dados são de responsabilidade de órgãos do Poder Legislativo e apenas uma base relacionada ao Poder Judiciário.

FIGURA 3 - RANKING DE DIMENSÕES DE DADOS ABERTOS NO BRASIL

RANKING	DIMENSÕES DO ÍNDICE	GARGALOS ENCONTRADOS	PODER RESPONSÁVEL
1	Estatísticas Socioeconômicas	0	Executivo
	Mapas Nacionais	0	Executivo
	Orçamento Público	0	Executivo
2	Atividade Legislativa	1	Legislativo
	Leis em Vigor	1	Legislativo
	Resultados Eleitorais	1	Judiciário
3	Gastos Públicos	2	Legislativo
4	Compras Públicas	3	Executivo
5	Limites Administrativos	4	Executivo
	Previsão do Tempo	4	Executivo
6	Registro de Empresas	5	Executivo
7	Qualidade do Ar	7	Executivo

As bases de dados estudadas nos dão exemplos de **boas práticas** que podem ser utilizadas como referências por outros provedores de dados. É o caso da dimensão **Orçamento Público** em relação à disponibilização de um manual técnico atualizado sobre o assunto. A Secretaria de Orçamento Federal **apresenta anualmente um manual com explicações e informações** sobre o orçamento da federação. Este manual auxilia o cidadão a lidar com o tema explicando as etapas de organização e processo do orçamento.

Outra prática notável foi identificada na dimensão **Gastos Públicos** em relação à riqueza dos dados fornecidos, apesar de não atingir 100% no índice. O Senado Federal disponibiliza todas as classificações oficiais das finanças públicas de um determinado gasto. Desta forma, o órgão apresenta bem mais informação sobre os gastos públicos do que o considerado essencial segundo o índice de dados abertos. Isso é positivo porque oferece uma maior capacidade de análise para cidadão, ampliando, assim, as possibilidades de utilização dos dados.

Sob este ponto de vista, as duas boas práticas tornam os dados mais transparentes porque fortalecem o poder de atuação do cidadão por meio de duas formas diferentes: o manual instrui e informa, enquanto que a disponibilidade de informações em mais de uma classificação fortalece a capacidade de análise do cidadão.

O problema mais comum, identificado em nove das 12 bases analisadas, é a **má usabilidade dos dados**, ou a dificuldade de se trabalhar com eles. O segundo problema, identificado em cinco bases, é a indisponibilidade de opção de download da base completa. A figura 4 indica o número de vezes que nos deparamos com cada gargalo.

FIGURA 4 - FREQUÊNCIA DE PROBLEMAS ENCONTRADOS NO BRASIL

GARGALOS MAPEADOS	FREQUÊNCIA DE IDENTIFICAÇÃO DOS GARGALOS
Dificuldade de trabalhar dados	9
Download da base completa indisponível	5
Dificuldade de localizar dados	3
Indisponibilidade de formato aberto	3
Dataset desatualizado	3
Licença não transparente	2
Dataset incompleto	2
Restrição de Acesso	1

Os três bancos de dados não encontrados para avaliação são considerados os mais críticos, ou seja, os que requerem maior atenção dentro da política de dados abertos, a saber: **Localizações**, **Qualidade da Água** e **Propriedade da Terra**.

Além disso, dois problemas são particularmente críticos para bancos avaliados, quando observados, justamente por serem aspectos fundamentais da transparência: a **inexistência da informação** e a **restrição do acesso** a ela. Esses itens devem ser considerados prioritários em uma estratégia de aprimoramento da situação dos dados abertos nacionais. No caso brasileiro, esses dois problemas são os menos frequentes da lista. A questão do dataset incompleto aparece duas vezes, em **Previsão do Tempo** e **Qualidade do Ar**. A restrição de acesso aparece uma vez, em **Registro de Empresas**. Vale notar que essas três dimensões possuem bases de dados providenciadas por órgãos do Poder Executivo federal.

SOBRE AS BASES DE DADOS

Esta seção apresenta comentários a respeito dos conjuntos de dados analisados a fim de dar profundidade ao índice, explicando as questões particulares de cada dimensão analisada.



Orçamento Público (Nota: 100%)

Não foram encontrados problemas neste dataset

No Brasil, existem várias fontes diferentes de dados de **Orçamento Público** (Ministério da Fazenda, Planejamento, Senado e Câmara dos Deputados), que estão listadas neste documento na seção final. Para a avaliação, foram considerados os dados abertos do **Sistema Integrado de Orçamento e Planejamento (SIOP)** da Secretaria do Orçamento Federal (SOF) do Ministério do Planejamento, que possui a atribuição de coordenar o processo orçamentário e consolidar informações sobre receita esperada e metas fiscais. A base de dados recebeu nota máxima na avaliação e nenhum gargalo foi detectado.

No entanto, um fato ressaltado ao longo do processo de revisão deve ser mencionado aqui. Os dados, ainda que disponíveis e acompanhados de documentação (Manual Técnico do Orçamento), são difíceis de serem trabalhados por duas razões: o uso eficiente da base de dados depende de um bom conhecimento sobre orçamento público, que é um tema árido e difícil; e a necessidade de ter habilidades com manipulação de dados em ecossistema RDF, o que reduz o universo de pessoas capazes de lidar com essa base. Apesar desses aspectos, um usuário que queira trabalhar com esses dados, mesmo que tenha um custo de entrada, possui acesso completo e documentação adequada para aprender a lidar com a base.

Resultados Eleitorais (Nota: 100%)



Problemas encontrados - Usabilidade

Dificuldades de trabalhar os dados

As informações a respeito dos resultados eleitorais no Brasil são consolidadas em nível federal pelo **Tribunal Superior Eleitoral (TSE)**. Ainda que tribunais regionais levantem e disponibilizem tais informações em alguns casos, o repositório de dados eleitorais do TSE é uma fonte consolidada e confiável de informação. Neste caso, os dados do TSE foram avaliados a respeito das eleições gerais (presidente, governador de estado, senador, deputado federal e deputado estadual) para todo o território brasileiro.

No que diz respeito à metodologia do índice e ao preenchimento do *survey*, todos os pré-requisitos foram atendidos. No entanto, um problema foi levantado, que é uma certa dificuldade de trabalhar a partir dos dados, uma vez que são codificados. A documentação facilita o estudo, mas é necessário tempo para se familiarizar com a natureza da informação e fazer algumas transformações necessárias para análise. Além disso, resalta-se que não há, no website acessado, informações sobre licença ou termos de uso, mas os dados estão amparados pela legislação federal de acesso à informação⁵, como informado pelos técnicos do TSE. Ou seja, ainda que os dados tenham licença pública, é importante que ela esteja disponibilizada de forma mais clara juntamente aos dados.



Mapas Nacionais (Nota: 100%)

Não foram encontrados problemas neste dataset

A principal referência de dados cartográficos no Brasil é o **Instituto Brasileiro de Geografia e Estatística (IBGE)**, que fornece as informações requisitadas pelo índice nesta dimensão. Esta base recebeu nota máxima e não apresentou problemas em nenhum quesito. De todo modo, vale ressaltar dois aspectos: os dados em formato *shapefile* não são facilmente encontrados no site; e, no que diz respeito à licença, foi encontrada apenas uma menção ao caráter público dos dados na missão do instituto, o que sugere uma necessidade de maior clareza da licença de dados abertos.

⁵ Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/l12527.htm>. Acesso em 20 abr 2017.



Estatísticas Socioeconômicas (Nota: 100%)

Não foram encontrados problemas neste dataset

As estatísticas solicitadas pela metodologia do índice – PIB, população e desemprego – são disponibilizadas em nível federal pelo **Instituto Brasileiro de Geografia e Estatística (IBGE)**. Neste caso, foram avaliadas as estatísticas no agregado do país por meio das contas nacionais trimestrais, Pesquisa Nacional por Amostra de Domicílios (PNAD) contínua e estimativas populacionais. Ainda que cumpram todos os requisitos da metodologia, os dados não são encontrados com muita facilidade por alguém que não tenha conhecimento prévio sobre esse tipo de informação, já que é possível obtê-los com diferentes recortes e metodologias em diferentes locais dentro do site do IBGE. No que diz respeito à licença, foi encontrada apenas uma menção ao caráter público dos dados na missão do instituto, o que sugere uma necessidade de maior clareza da licença de dados abertos.



Leis em Vigor (Nota: 100%)

Problemas encontrados - Usabilidade

Dificuldades de trabalhar os dados

Existem várias fontes possíveis para informações sobre leis federais em vigor no Brasil. A principal base de dados utilizada para avaliação foi o projeto *Lexml*, parte da iniciativa Interlegis, que contém não só uma página para consultas como também acesso à sua API, cuja documentação consta na página de dados abertos do **Senado Federal**. A Presidência da República também possui um portal de consultas, mas não permite download da base completa.

No processo de avaliação, ainda que tenha obtido nota máxima, a base de dados foi diagnosticada com um problema de usabilidade, uma vez que sua API é de difícil utilização e com poucos metadados, o que torna difícil a vida de usuários que queiram utilizar a base, mesmo que tenham habilidades avançadas de programação. Uma documentação mais adequada e maior clareza nos metadados pode facilitar o uso desta base.



Atividade Legislativa (Nota: 100%)

Problemas encontrados - Usabilidade

Dificuldades de trabalhar os dados

Pela própria natureza bicameral do Congresso Nacional brasileiro, esse tipo de dado é fornecido separadamente pela **Câmara dos Deputados** e pelo **Senado Federal**. Ambas bases receberam nota máxima na avaliação. Ambos órgãos apresentam não só portais de busca como também acesso à suas APIs, que foram julgadas como sendo de difícil usabilidade. Na questão da licença de dados abertos, o Senado apresentou uma licença compatível com a *Open Definition*, mas a Câmara apresenta apenas um FAQ (*Frequently Asked Questions*) com menção ao caráter público dos dados, sem divulgar uma licença específica. O download em bloco é também disponibilizado em formato .csv no caso da Câmara, mas não é claramente disponível no caso do Senado, ainda que seja possível acessar a base completa por meio da API.

Gastos Públicos (Nota: 83%)



Problemas encontrados - Usabilidade

Dificuldades de trabalhar os dados

Problemas encontrados - Processo

Download da base completa indisponível

As informações requeridas nesta dimensão são geradas pelo Tesouro Nacional mas disponibilizadas com acesso aberto por três diferentes órgãos brasileiros: O Ministério da Transparência, Fiscalização e Controladoria Geral da União, o Senado Federal e o Ministério do Planejamento. Para a avaliação, consideramos a base de dados do **Senado Federal**. A fonte possui como ponto positivo o fato de disponibilizar muitas informações diferentes sobre os gastos, através de diversas classificações oficiais das finanças públicas. Além disso, a aplicação permite buscas bastante flexíveis de dados. Por outro lado, notamos que além de não existir um acesso à base diretamente via API, ao construir uma busca ao banco com muitas variáveis, é possível que haja problemas de download, por ser uma base pesada. Por conta disso, considerou-se que não há opção de download da base completa nesta dimensão.

Limites Administrativos (Nota: 78%)



Problemas encontrados - Usabilidade

Dificuldades de trabalhar os dados

Problemas encontrados - Processo

Dificuldade para localizar informações

Licença não transparente

A base de dados analisada para esta dimensão é fornecida pelo **Instituto Brasileiro de Geografia e Estatística (IBGE)**, entidade vinculada ao Ministério do Planejamento, Desenvolvimento e Gestão, e portanto, um órgão do Poder Executivo federal. Assim como no caso da dimensão de **Mapas Nacionais**, cuja fonte é a mesma que a desta dimensão, ressaltamos a dificuldade de se encontrar os dados no site e a falta de transparência na licença específica da base (consideramos a declaração de publicização geral, realizada pelo instituto)⁶.

Previsão do Tempo (Nota: 72%)

Problemas encontrados - Usabilidade



Ausência de informações requeridas pela metodologia: temperatura média

Dificuldades de trabalhar os dados

Indisponibilidade de dados em formato aberto

Problemas encontrados - Processo

Download da base completa indisponível

Para esta dimensão, duas bases de dados diferentes foram consideradas, uma do **Instituto Nacional de Pesquisas Espaciais (INPE)** e outra do **Instituto Nacional de Meteorologia (INMET)**. A natureza metodológica e a amplitude das informações encontradas é diferente entre as duas bases. A informação de velocidade do vento, por exemplo, só está disponível na base do INPE. Tendo isso em vista, considera-se que a consolidação dessas informações em um portal único pode ser um aprimoramento possível para esta dimensão.

A usabilidade do dado é prejudicada pela indisponibilidade de download, restringindo-se à navegação web, o que impacta também a questão do formato, que não é aberto. Além disso, não foi possível encontrar a temperatura média em quaisquer bases. Na dimensão de processos, além de a base não estar disponibilizada de forma consolidada, ressalta-se que as licenças específicas não foram encontradas, mas apenas termos e condições de uso e replicação (INMET) e um rodapé dizendo que os usuários devem sempre mencionar a fonte do dado (INPE)..

⁶ Existe uma divergência na avaliação da licença dos dados do IBGE. Em outras dimensões a licença transparente considerada foi a declaração de publicidade do instituto. No caso da dimensão de Limites Administrativos, no entanto, o processo de revisão avaliou que a licença não era transparente porque não havia uma declaração específica sobre a base de dados. Os resultados divulgados neste relatório estão em conformidade com a avaliação oficial em sua versão atual (24/04/2017). Em uma próxima revisão, é possível que haja alteração neste ponto para garantir a consistência e coerência interna quanto a licença dos dados do IBGE.

Compras Públicas (Nota: 67%)

Problemas encontrados - Usabilidade

Dificuldades de trabalhar os dados
Dados desatualizados

Problemas encontrados - Processo

Download da base completa indisponível

As informações sobre compras públicas são consolidadas e disponibilizadas pelo **Ministério do Planejamento, Desenvolvimento e Gestão**. Todos os dados listados na metodologia estão disponíveis. No entanto, eles não podem ser baixados juntamente em uma única base. Para consolidá-las, é necessário fazer download de diferentes bases e manipular os dados em algum outro programa. Vale ressaltar que existe um acesso via API, mas que apresenta o mesmo problema de desagregação da informação. Portanto, a não disponibilização da base completa, nem para download, nem para consulta, gera também um problema de usabilidade na dimensão. Além disso, os dados não estavam atualizados no período em que o site foi avaliado, sendo a base disponibilizada para download atualizada até setembro de 2016 e os dados disponíveis via API até 2014.

Registro de Empresas (Nota: 39%)

Problemas encontrados - Usabilidade

Indisponibilidade de dados em formato aberto
Dados desatualizados
Dificuldades de trabalhar os dados

Problemas encontrados - Processo

Download da base completa indisponível
Dificuldade para localizar informações
Restrição de acesso

Existem informações coletadas e disponibilizadas sobre registro de empresas pelo IBGE, pelo Ministério do Planejamento, Desenvolvimento e Gestão, pelo Ministério da Transparência, Fiscalização e Controladoria Geral da União e por juntas comerciais estaduais, mas a única entidade que possui a completude e a granularidade dos dados conforme requisitado – e, portanto, que poderia qualificar para a avaliação do índice – é a **Receita Federal** (secretaria vinculada ao Ministério da Fazenda), que cataloga essas informações para fins de tributação.

Ainda assim, essa base de dados possui uma série de problemas. Para acessar informações sobre cada empresa individualmente, é necessário possuir, de antemão, o identificador único da empresa (CNPJ). Ou seja, não é possível encontrar informações

apenas pelo nome. Também é necessário preencher um *captcha* para acessar as informações (ou seja, há uma restrição de acesso), que só são disponibilizadas em uma página da web (sem download disponível). Além disso, a data de atualização do site não é conhecida. Não foi encontrada uma licença específica, apenas uma menção de que se trata de informação de domínio público.

Qualidade do Ar (Nota: 33%)

Problemas encontrados - Usabilidade

Ausência de informações requeridas pela metodologia: óxidos de nitrogênio

Indisponibilidade de dados em formato aberto

Dados desatualizados

Dificuldades de trabalhar os dados

Problemas encontrados - Processo

Download da base completa indisponível

Dificuldade para localizar informações

Falta de clareza sobre a licença pública de uso de dados

Existem diversas fontes para dados de qualidade do ar no Brasil, no entanto nenhuma delas responde aos critérios do índice de dados abertos. As duas principais fontes em nível nacional utilizadas foram o **Instituto Nacional de Pesquisas Espaciais (INPE)** e o **Ministério do Meio Ambiente (MMA)**, por serem as únicas que oferecem informações sobre todo o território. Esses dados possuem melhor qualidade no nível subnacional, no entanto, não há uma consolidação eficaz dessas informações para todo o território.

Dentre os problemas de usabilidade, destaca-se a ausência de informações sobre óxidos de nitrogênio no ar, a indisponibilidade do download de dados em formato aberto (os dados estão disponibilizados em PDF ou em um mapa plotado na web), a ausência de dados atualizados (no caso do INPE, ainda que os dados sejam em tempo real, eles estão apenas plotados, não sendo visíveis) e a consequente dificuldade em retirar informações e análises dos dados fornecidos. Na dimensão de processos, além da dificuldade em localizar informações razoáveis e a indisponibilidade de download da base completa, destaca-se a ausência de referências diretas à licença de dados ou ao caráter público das informações.



Localizações (Nota: 0%)

Este dataset não foi encontrado.

De acordo com a *Universal Postal Union*, organização internacional de serviço postal, a entidade governamental brasileira responsável pelos códigos postais no Brasil é o Ministério das Comunicações, sendo a **Empresa Brasileira de Correios e Telégrafos** o operador designado. Os Correios disponibilizam as informações demandadas pelo índice em um serviço de busca online, mas não disponibiliza a base para download. O IBGE possui dados úteis sobre localizações, mas que não são compatíveis com a definição do índice de dados abertos. Neste caso, é necessário avaliar, de acordo com os critérios da metodologia, uma forma de disponibilizar uma base nacional sobre o assunto em formato aberto.



Qualidade da Água (Nota: 0%)

Este dataset não foi encontrado.

Por mais que existam algumas fontes que coletam informações sobre qualidade de água no Brasil (Agência Nacional das Águas, Fundação Oswaldo Cruz, Ministério da Saúde) – sendo que a submissão considerou a base do Sistema Nacional de Informações sobre Saneamento (SNIS) do **Ministério das Cidades** – o processo de revisão do índice considerou que o Brasil não possui uma base de dados que qualifique para os critérios nesta dimensão. Neste caso, é necessário avaliar, de acordo com os critérios da metodologia, uma forma de disponibilizar uma base nacional sobre o assunto em formato aberto.



Propriedade da terra (Nota: 0%)

Este dataset não foi encontrado.

Os dados de propriedade da terra demandados pela metodologia do índice não são disponibilizados por entidades governamentais no Brasil. O registro de propriedade é realizado nos cartórios e os dados consolidados não são disponíveis publicamente. O mais próximo de uma base sobre o assunto encontrada foi para certificação de cadastro rural realizada pelo **Instituto Nacional de Colonização e Reforma Agrária (INCRA)**. No entanto, a base não é compatível com os critérios da avaliação.

CONCLUSÃO

O Brasil teve um bom desempenho no índice quando comparado aos demais países. No entanto, a pontuação de 65% indica que ainda há muito espaço para aprimoramentos. Apenas 40% das dimensões do índice receberam a nota total e três delas nem mesmo existem publicamente. A maior parte dos problemas identificados está mais relacionada à usabilidade (completude, atualização e formato) do que a processo (meios de acesso e licença de uso dos dados).

O problema mais recorrente é a dificuldade de se trabalhar os dados e, em segundo lugar, a não disponibilização da base de dados completa para download. Isso indica que os órgãos responsáveis pelo provimento das informações talvez não tenham a compreensão de que a facilitação do processo de edição e manipulação dos dados é também sua responsabilidade e parte importante da transparência.

Tais resultados podem auxiliar na abertura de dados no âmbito nacional, direcionando a atenção para dimensões prioritárias e problemas mais comuns de acordo com a definição de dados abertos. Em outras palavras, este documento traz algumas diretrizes para uma política de dados abertos do país.

Essa iniciativa pode ter uma série de benefícios, tanto internamente quanto externamente à administração pública. Internamente, porque o próprio processo de abertura de dados cria um mecanismo de cobrança interna para resultados das diferentes unidades, o que contribui para uma **maior eficiência da gestão pública**. Além disso, a abertura de dados permite que os cidadãos possam contribuir de diversas formas, seja com análises, controle social, criação de aplicativos, replicação da informação, entre outras, o que acaba por tornar a gestão mais eficiente ao congregar esforços por meio da participação da sociedade. Por outro lado, há também benefícios externos: a administração passa a ser vista como mais accountable aos seus cidadãos. Além dos ganhos de imagem, a possibilidade de se tornar referência internacional em dados abertos contribui positivamente para o advocacy da questão da transparência em outros países.

Com isso, a FGV/DAPP e a OKBr sinalizam algumas recomendações gerais de atuação para o aprimoramento do desempenho do país. Em primeiro lugar, identificamos a necessidade de sincronização do modo de provimento dos dados em todas as dimensões, de forma a parametrizar a qualidade do serviço de disponibilização dos dados entre os diferentes órgãos. Para isso, é importante determinar estratégias de atuação para cada unidade de acordo com um manual de orientação geral que pode ser gerado com insumos deste mesmo relatório. Recomendamos, adicionalmente, a criação de formas inovadoras de apresentação das informações, por meio da criação de uma metodologia de visualização de dados. Esta ação deve contribuir significativamente para melhora da usabilidade das bases no escopo nacional, além de reforçar a posição de liderança no Brasil no que diz respeito ao avanço da política de transparência.

Aqui são elencadas as fontes utilizadas (incluindo os links acessados) para a avaliação por dimensão do índice. São destacados os órgãos provedores da informação e o tipo de administração.

É importante observar que a avaliação é feita para o estado dos dados abertos para o ano de 2016. Estas fontes foram consultadas no período do levantamento que foi de 01 a 30 de novembro de 2016. O processo de revisão foi realizado entre janeiro e fevereiro de 2017.

Orçamento Público

Secretaria de Orçamento Federal - Ministério do Planejamento, Orçamento e Gestão

Órgão do Poder Executivo Federal

Link principal utilizado para avaliação:

Sistema integrado de planejamento e orçamento (SIOP)

https://www1.siop.planejamento.gov.br/QvAJAZfc/pendoc.htm?document=IAS%2FExecucao_Orcamentaria.qvw&host=QVS%40pqlk04&anonymous=true

Links alternativos:

Secretaria de Orçamento Federal (SOF)

http://www.orcamentofederal.gov.br/clientes/portalsof/portalsof/orcamentos-anuais/orcamento-2016/orcamentos_anuais_view?anoOrc=2016

Ministério do Planejamento, Desenvolvimento e Gestão (MP)

<http://www.planejamento.gov.br/assuntos/planeja/orcamento-federal/orcamentos-anuais>

Sistema SIGA Brasil - Senado Federal

<https://www12.senado.leg.br/orcamento/sigabrasil>

Lei Orçamentária Anual - Câmara dos Deputados

<http://www2.camara.leg.br/atividade-legislativa/orcamentobrasil/loa>

Portal de dados abertos do Governo Federal - Secretaria de Tecnologia da Informação do Ministério do Planejamento, Desenvolvimento e Gestão (MP)

<http://dados.gov.br/dataset/orcamento-federal>

Resultados Eleitorais

TSE - Tribunal Superior Eleitoral

Órgão do Poder Judiciário Federal

Repositório de dados eleitorais

<http://www.tse.jus.br/eleicoes/estatisticas/repositorio-de-dados-eleitorais>

Mapas Nacionais

IBGE - Instituto Brasileiro de Geografia e Estatística

Entidade vinculada ao Ministério do Planejamento, Desenvolvimento e Gestão, órgão do Poder Executivo Federal

Portal de Downloads em Geociências -

Base completa

http://downloads.ibge.gov.br/downloads_geociencias.htm

Portal de Mapas - Ferramenta de busca

<http://portaldemapas.ibge.gov.br/portal.php#mapa1106>

Estatísticas Socioeconômicas

IBGE - Instituto Brasileiro de Geografia e Estatística

Entidade vinculada ao Ministério do Planejamento, Desenvolvimento e Gestão, órgão do Poder Executivo Federal

Produto Interno Bruto - Contas Nacionais Trimestrais
<http://www.ibge.gov.br/home/estatistica/indicadores/pib/defaultcnt.shtm>

Taxas de desemprego trimestrais a partir da Pesquisa Nacional por Amostra de Domicílios (PNAD) Contínua
http://www.ibge.gov.br/home/estatistica/indicadores/trabalhoerendimento/pnad_continua_mensal/default.shtm

Estimativas populacionais
<http://www.ibge.gov.br/home/estatistica/populacao/estimativa2016/default.shtm>

Portal SIDRA - Sistema IBGE de Recuperação Automática
<https://sidra.ibge.gov.br/home/pms/brasil>

Leis em Vigor

Senado Federal

Órgão do Poder Legislativo Federal

Links principais utilizados para avaliação:
Projeto Lexml, parte de iniciativa Interlegis
<http://projeto.lexml.gov.br/>

Portal de dados abertos do Senado Federal
Documentação
<http://dadosabertos.senado.gov.br/dataset/api-search-and-retrieve-via-url-sru>
http://legis.senado.leg.br/dadosabertos/docs/resource_LegislacaoService.html#resource_LegislacaoService_listaNormasXml_GET

Portal de dados abertos do Senado Federal - Busca
<http://dadosabertos.senado.gov.br/group/legislacao2>

Links alternativos:

Portal da Presidência da República para consultas
<http://www4.planalto.gov.br/legislacao>

<https://legislacao.planalto.gov.br/legisla/legislacao.nsf/fraWeb?OpenFrameSet&Frame=frmWeb2&Src=/legisla/legislacao.nsf%2FFrmConsultaWeb1%3FOpenForm%26AutoFramed>

Atividade Legislativa

Câmara dos Deputados e Senado Federal

Órgãos do Poder Legislativo Federal

Portal de dados abertos da Câmara dos Deputados
<http://www2.camara.leg.br/transparencia/dados-abertos/dados-abertos-legislativo>

Portal de dados abertos do Senado
<http://dadosabertos.senado.gov.br/>

Portal de busca da Câmara dos Deputados
<http://www.camara.leg.br/buscaProposicoesWeb/pesquisaSimplificada>

Portal de busca do Senado

<http://www25.senado.leg.br/web/atividade>

Gastos Públicos

Senado Federal

Órgão do Poder Legislativo Federal

Link principal utilizado para avaliação:
Sistema SIGA Brasil - Senado Federal
<https://www12.senado.leg.br/orcamento/sigabrasil>

Links alternativos:

Sistema integrado de planejamento e orçamento (SIOP)

https://www1.siop.planejamento.gov.br/QvAJAXZfc/opepdoc.htm?document=IAS%2FExecucao_Orcamentaria.qvw&host=QVS%40pqlk04&anonymous=true

Portal da Transparência - Ministério da Transparência, Fiscalização e Controladoria-Geral da União

<http://transparencia.gov.br/>

Limites Administrativos

IBGE - Instituto Brasileiro de Geografia e Estatística

Entidade vinculada ao Ministério do Planejamento, Desenvolvimento e Gestão, órgão do Poder Executivo Federal

Portal de Downloads em Geociências

Base completa

http://downloads.ibge.gov.br/downloads_geociencias.htm

Previsão do Tempo

INPE - Instituto Nacional de Pesquisas Espaciais

Entidade vinculada ao Ministério da Ciência, Tecnologia, Inovações e Comunicações, órgão do Poder Executivo Federal

INMET - Instituto Nacional de Meteorologia

Entidade vinculada ao Ministério da Agricultura, Pecuária e Abastecimento, órgão do Poder Executivo Federal

Portal do CPTEC - Centro de Previsão de Tempo e Estudos Climáticos - do INPE

<http://tempo.cptec.inpe.br/>

Portal do INMET

<http://www.inmet.gov.br/portal/>

Compras Públicas

Ministério do Planejamento, Desenvolvimento e Gestão (MP)

Órgão do Poder Executivo Federal

Links principais utilizados para avaliação:

API de Compras Públicas - Portal de dados abertos do Governo Federal da Secretaria de Tecnologia da Informação do Ministério do Planejamento, Desenvolvimento e Gestão (MP)

<http://dados.gov.br/dataset/compras-publicas-do-governo-federal>

Documentação da API

<http://compras.dados.gov.br/docs/home.html>

Links alternativos:

Portal da Transparência - Ministério da Transparência, Fiscalização e Controladoria-Geral da União

<http://transparencia.gov.br/>

Seção de Downloads do Portal da Transparência - Ministério da Transparência, Fiscalização e Controladoria-Geral da União

<http://transparencia.gov.br/downloads/>

Registro de Empresas

Secretaria Receita Federal do Brasil - Ministério da Fazenda

Órgão do Poder Executivo Federal

Ferramenta de busca de situação cadastral de empresas via CNPJ - Receita Federal

http://www.receita.fazenda.gov.br/PessoaJuridica/CNPJ/cnpjreva/Cnpjreva_Solicitacao2.asp

Qualidade do Ar

INPE - Instituto Nacional de Pesquisas Espaciais

Entidade vinculada ao Ministério da Ciência, Tecnologia, Inovações e Comunicações, órgão do Poder Executivo Federal

Ministério do Meio Ambiente (MMA)

Órgão do Poder Executivo Federal

Links principais utilizados para avaliação:

Mapa com informações em tempo real sobre qualidade do ar e emissão de poluentes do CPTEC - Centro de Previsão de Tempo e Estudos Climáticos - do INPE
<http://meioambiente.cptec.inpe.br/index.php?lang=pt>

Relatório de monitoramento da qualidade do ar por estação de monitoramento em todo o território - Ministério do Meio Ambiente (MMA)
http://www.mma.gov.br/images/arquivo/80060/Diagnostico_Nete_de_Monitoramento_da_Qualidade_do_Ar.pdf

Links alternativos:

Dados para download sobre emissão de poluentes na região amazônica - INPE
http://inpe-em.ccst.inpe.br/conteudo_pt/downloads.html

Séries estatísticas de emissões de poluentes por região metropolitana - IBGE

http://seriesestatisticas.ibge.gov.br/series.aspx?v_codigo=IU5&t=concentracao-poluente-ar-areas-urbanas-1

Localizações

Dataset não encontrado

Qualidade da Água

Dataset não encontrado

Propriedade da terra

Dataset não encontrado

Inovação e Políticas Públicas

