

Parceria para Governo Aberto (OGP)
4º PLANO DE AÇÃO

MODELO DE REFERÊNCIA DE ABERTURA DE DADOS

Documento de referência do Marco 5 do Compromisso 2:
Ecossistema de Dados Abertos

Em parceria com

CONTROLADORIA-GERAL DA UNIÃO GOVERNO FEDERAL

ceweb.br nic.br



MINISTÉRIO DA
**CIÊNCIA, TECNOLOGIA
E INOVAÇÕES**

Dados internacionais de Catalogação na Publicação
Ficha catalográfica elaborada pelo Sistema de Bibliotecas/FGV

Modelo de referência de abertura de dados [recurso eletrônico] : documento de referência do Marco 5 do Compromisso 2 : ecossistema de dados abertos / Coordenação Wagner Oliveira; pesquisadores Ana Eliza Duarte...[et al.]. - Rio de Janeiro : FGV DAPP, 2020.

1 recurso online (96 p.) : ePub

Dados eletrônicos.
Inclui bibliografia.
ISBN: 978-65-86845-02-0

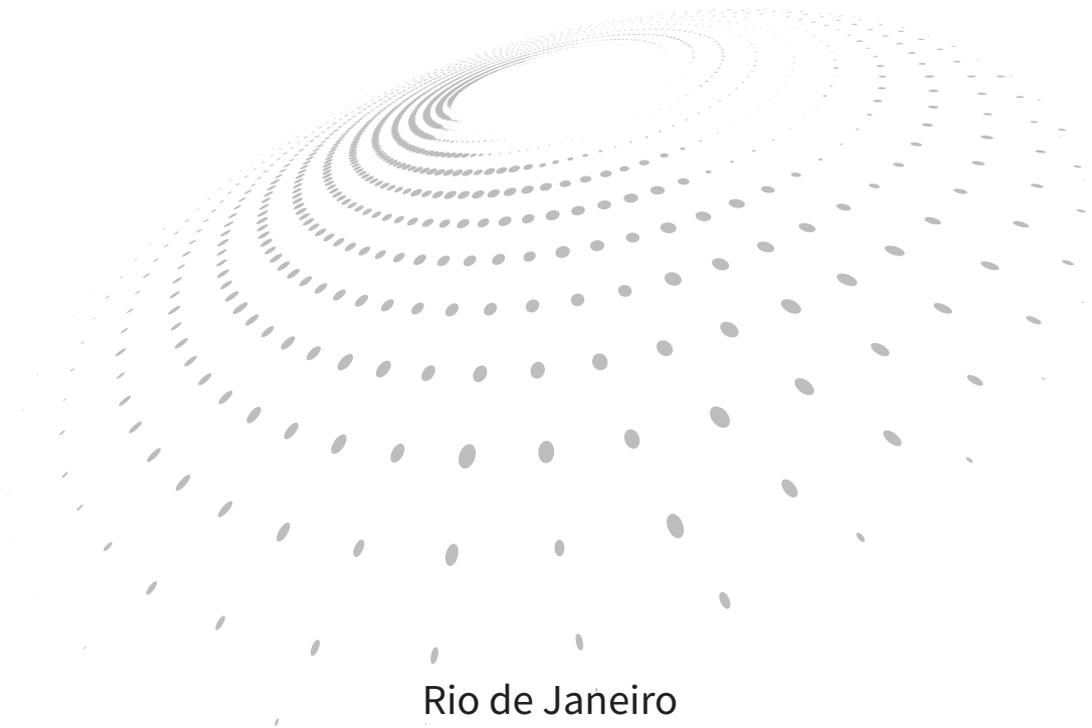
1. Abertura de informação ao público. 2. Serviços de informação on-line. 3. Transparência na administração pública. 4. Políticas públicas. I. Oliveira, Wagner. II. Duarte, Ana Eliza. III. Fundação Getulio Vargas. Diretoria de Análise de Políticas Públicas.

CDD – 351

Diretoria de Análise de Políticas Públicas
da Fundação Getulio Vargas

MODELO DE REFERÊNCIA DE ABERTURA DE DADOS

Documento de referência do Marco 5 do Compromisso 2:
Ecossistema de Dados Abertos



Rio de Janeiro
FGV DAPP
2020

Expediente



Fundada em 1944, a Fundação Getúlio Vargas nasceu com o objetivo de promover o desenvolvimento socioeconômico do Brasil por meio da formação de administradores qualificados, nas áreas pública e privada. Ao longo do tempo, a FGV ampliou sua atuação para outras áreas do conhecimento, como Ciências Sociais, Direito, Economia, História, Matemática Aplicada e Relações Internacionais, sendo referência em qualidade e excelência, com suas dez escolas.

ESCRITÓRIO

Edifício Luiz Simões Lopes (Sede)
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 E FUNDADOR

Luiz Simões Lopes

PRESIDENTE

Carlos Ivan Simonsen Leal

VICE-PRESIDENTES

Francisco Oswaldo Neves Dornelles (licenciado)
Marcos Cintra Cavalcanti de Albuquerque (licenciado)
Sergio Franklin Quintella



DIRETOR

Marco Aurelio Ruediger

Diretoria de Análise de Políticas Públicas (DAPP)

(21) 3799-6208

www.dapp.fgv.br | dapp@fgv.br

EQUIPE DE EXECUÇÃO

Coordenação geral

Wagner Oliveira

Pesquisadores

Beatriz Meirelles

Leonor Jungstedt

Paula Audibert

Revisão de texto

Dalby Dienstbach

Projeto gráfico

Daniel Almada

Parceiros

CONTROLADORIA-GERAL
DA UNIÃO
GOVERNO
FEDERAL

A Controladoria-Geral da União atua para o aperfeiçoamento da gestão pública e o combate à corrupção. Entre seus objetivos estratégicos está ampliar e aperfeiçoar as políticas de promoção de transparência, ética e de participação social para melhoria do Estado.

Controladoria-Geral da União (CGU)

www.gov.br/cgu

Marcelo de Brito Vidal

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

Guilherme Guimarães Fortuna

Coordenador-Geral de Governo Aberto e Transparência Substituto
guilherme.fortuna@cgu.gov.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.

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

ceweb.br

Caroline Burle

Relações Institucionais e Internacionais
cburle@nic.br

Ana Eliza Duarte

Analista de Projetos Web
anaeliza@nic.br



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.

Open Knowledge Brasil (OKBr)

www.ok.org.br

Camille Moura

Advocacy e Pesquisa
camille@ok.org.br



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.

Instituto de Estudos Socioeconômicos (INESC)

www.inesc.org.br

Carmela Zigoni

Assessora Política

carmela@inesc.org.br

MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E INOVAÇÕES



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

Ministério da Ciência, Tecnologia e Inovações (MCTI)

Centro de Tecnologia da Informação Renato Archer (CTI)

www.cti.gov.br/pt-br

Jarbas Lopes Cardoso Junior

Tecnologista Senior

jarbas.cardoso@cti.gov.br

Sumário

Apresentação	14
1. Introdução	16
1.1 O que são dados abertos	16
1.2 Benefícios da publicação de dados abertos	20
1.3 Política de Dados Abertos do Poder Executivo Federal	21
2. Publicação de Dados	24
2.1 Catalogação dos Dados	24
2.1.1 Plataformas de gerenciamento de dados abertos mais utilizadas.....	25
2.1.2 Definição de metadados obrigatórios e facultativos.....	28
2.1.3 Definição da licença dos dados publicados.....	29
2.1.4 Informações sobre a qualidade dos dados.....	30
2.1.5 Informações de procedência de dados.....	30
2.1.6 Indicador da versão dos dados e histórico das versões.....	31
2.1.7 Definição de identificadores de dados.....	32
2.1.8 Dados disponibilizados em diversos formatos.....	33
2.1.9 Vocabulários de dados.....	34
2.1.9.1 Vocabulário do Catálogo de Dados (DCAT).....	35
2.1.10 Acesso aos dados.....	36
2.1.11 Atualização dos dados.....	37
2.1.12 Utilização de APIs.....	38
2.1.13 Federação de dados.....	39

2.1.14 Proteção e anonimização de dados pessoais/individualizados.....	39
2.2 Usuário	40
2.2.1 Ferramentas de colaboração e compartilhamento.....	40
2.2.2 Feedback dos consumidores de dados e compartilhamento dos feedbacks.....	41
2.2.3 Interação com o publicador de dados.....	42
2.2.4 Linguagem clara e acessível ao cidadão.....	42
2.2.5 Design e navegação simples e intuitiva.....	43
2.2.6 Visualização dos dados.....	43
2.2.7 Seleções, filtros ou extração de dados.....	44
2.2.8 Mecanismos de mapeamento da demanda por dados abertos.....	45
2.2.9 Tutoriais e manuais para utilização adequada dos dados.....	45
3. Reutilização de dados	46
3.1 Aplicativos e soluções desenvolvidas com o uso de dados abertos.....	47
3.1.1 Citymapper	47
3.1.2 Influence Explorer	48
3.1.3 Prescribing Analytics.....	48
3.1.4 QEdu	49
3.1.5 Radar Parlamentar	50
3.1.6 Appelou	50
3.1.7 Colaboradados.....	50
3.1.8 Brasil.io	51
3.1.9 CoronApp.....	51
3.1.10 Coronavírus SUS.....	51
3.1.11 Coronavírus no Mundo, no Brasil e em Pernambuco.....	52

3.1.12 Ciência Aberta é Vida - Diretório de fontes de informação científica de livre acesso sobre o coronavírus	52
3.1.13 Información Legal	52
3.1.14 Observatório de Oncologia	53
3.1.15 Serenata do Amor	53
3.2 Boas práticas de utilização de dados abertos	53
3.2.1 Enriquecimento de Dados	54
3.2.2 Republicação de Dados	55
4. Diretrizes de abertura para dados públicos estaduais e municipais	56
4.1 Metodologia da pesquisa com especialistas	56
4.2 Dados abertos essenciais por área	58
4.2.1 Legislativo	58
4.2.2 Dados geográficos	59
4.2.3 Educação	60
4.2.4 Meio ambiente e saneamento	62
4.2.5 Orçamento público	64
4.2.6 Infraestrutura e mobilidade urbana	65
4.2.7 Contratos e Licitações	67
4.2.8 Indicadores econômicos	67
4.2.9 Propriedade e uso da terra	69
4.2.10 Saúde	70
4.2.11 Indicadores sociais	71
4.2.12 Judiciário	72
4.2.13 Segurança	73

4.2.14 Outros	75
4.3 Considerações da seção	76
5. Diretrizes de abertura de dados definidas por temas nacionais estratégicos.....	77
5.1 Educação	77
5.2 Saúde	78
5.3 Meio ambiente e saneamento	78
5.4 Infraestrutura e mobilidade urbana.....	79
5.5 Segurança	79
5.6 Indicadores sociais.....	79
5.7 Orçamento público	80
5.8 Propriedade e uso da terra	80
5.9 Judiciário	80
5.10 Indicadores econômicos	80
5.11 Dados geográficos	81
5.12 Ciência, Tecnologia e Inovação	81
5.13 Contratos e Licitações	81
5.14 Legislativo	81
6. Governança dos dados abertos	82
6.1 Fundamentação	82
6.1.1 Participação multilateral.....	84
6.1.2 Diversidade	85
6.1.3 Processo democrático	85
6.1.4 Inovação	85

6.1.5 Colaboração	85
6.1.6 Transparência	86
6.1.7 Integridade	86
6.2 Legislação associada	86
Referências	87
Anexo	95

Apresentação

As discussões sobre dados abertos vêm ganhando espaço e importância nos setores público e privado brasileiros nos últimos anos. Podemos citar, como marco desse processo, a publicação da Lei de Acesso à Informação (Lei n.º 12.527/2011), que dispõe sobre o acesso a dados e informações produzidas ou custodiadas pelo governo. Além de regulamentar o fornecimento de informações sob demanda do cidadão, o normativo também determina que órgãos e entidades devem se antecipar aos pedidos e publicar seus dados e informações na internet (em transparência ativa), bem como exige que os dados sejam publicados em formatos abertos e não-proprietários.

No contexto da Parceria para Governo Aberto (OGP)¹, o tema faz parte dos compromissos assumidos pelo Brasil desde o 1º Plano de Ação Nacional², em 2011, quando o país se tornou membro cofundador da OGP e assumiu internacionalmente um compromisso de estruturação dos primeiros passos para uma política de dados abertos.

Mais recentemente, o Brasil realizou oficinas de cocriação, entre maio e agosto de 2018, para elaborar o 4º Plano de Ação Nacional de Governo Aberto³. O resultado dessas rodadas de reuniões integradas por membros do governo e da sociedade civil foi a definição de onze compromissos brasileiros, dentre os quais se destaca o compromisso n.º 2, sobre o ecossistema de dados abertos.

Esse compromisso objetiva estabelecer, de forma colaborativa, um 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. Todos os detalhes e informações sobre a realização dessas oficinas e o conteúdo completo

¹ O Brasil é membro da Parceria para Governo Aberto, uma iniciativa multilateral que visa garantir compromissos concretos de governos nacionais e subnacionais para promover um governo aberto, capacitar cidadãos, combater a corrupção e aproveitar novas tecnologias para fortalecer a governança. Este documento é um dos resultados desse esforço. Disponível em: <https://www.opengovpartnership.org/>. Acesso em: 28 ago. 2020.

² Disponível em: <https://www.gov.br/cgu/pt-br/governo-aberto/a-ogp/planos-de-acao/1o-plano-de-acao>. Acesso em: 10 set. 2020.

³ Disponível em: <https://www.gov.br/cgu/pt-br/governo-aberto/a-ogp/planos-de-acao/4o-plano-de-acao-brasileiro>. Acesso em: 10 set. 2020.

do compromisso, tais como sua estrutura e relatórios de execução das ações, podem ser visualizados no Portal da Parceria para Governo Aberto⁴.

Dentre os marcos/ações estabelecidos para possibilitar a execução do compromisso sobre dados abertos, a elaboração do “Modelo de referência para abertura de dados” se configura como o mais importante, visto que pretende, com ampla divulgação nas três esferas de governo, incentivar e padronizar a abertura de dados públicos no âmbito dos entes federativos.

A finalidade do documento consiste em integrar, capacitar e sensibilizar a sociedade e as três esferas de governo no contexto da abertura de dados públicos. Portanto, o presente modelo busca orientar gestores, agentes públicos e sociedade quanto à importância, ao uso, à publicação, à sustentação e ao monitoramento de dados abertos nos seus respectivos campos de atuação.

Sobre a estrutura do documento, a introdução apresenta aspectos gerais sobre o tema de dados abertos, como conceitos, princípios, benefícios gerados pela reutilização dos dados e informações sobre o funcionamento da Política de Dados Abertos do Poder Executivo Federal.

O segundo eixo aborda o tema da publicação dos dados, desmembrando-o nos subtemas: catalogação dos dados e usuário de dados. O modelo de abertura de dados proposto indica a importância da utilização de certos padrões para disponibilização de dados, assim como também busca permitir a usabilidade e a interatividade do cidadão consumidor desses dados.

Já o terceiro eixo busca esclarecer a importância de se desenvolverem mecanismos que impulsionem a reutilização dos dados pelos cidadãos, de modo que seja agregado valor aos dados publicados pelos órgãos públicos.

Os quarto (“Diretrizes de abertura para dados públicos estaduais e municipais”) e quinto eixos (“Diretrizes de abertura de dados definidos por temas nacionais estratégicos”) do documento expõem alguns padrões a serem seguidos na disponibilização de dados com a finalidade de propiciar maior uniformidade na publicação de dados abertos pelos diversos entes federativos. Por fim, o sexto eixo (“Governança dos dados abertos”) aborda uma estrutura de governança de dados para gestão dos processos de abertura, acompanhamento e monitoramento dos dados publicados, bem como dos produtos desenvolvidos pela sociedade a partir desses dados.

⁴ Disponível em: <https://www.gov.br/cgu/pt-br/governo-aberto/a-ogp/planos-de-acao/4o-plano-de-acao-brasileiro/compromisso-2-docs/ecossistema-de-dados-abertos-monitoramento-e-execucao>. Acesso em: 10 set. 2020.

1. Introdução

1.1 O que são dados abertos

De acordo com a Open Knowledge Foundation⁵, dados abertos são dados que podem ser livremente acessados, utilizados, modificados e compartilhados por qualquer pessoa, estando sujeitos a, no máximo, exigências que visem preservar sua proveniência e sua abertura (por exemplo, citação da fonte ou divulgação com a mesma licença).

A partir do conceito apresentado, cabe destacar as principais características⁶ dos dados abertos: disponibilidade, ou seja, dados acessíveis de forma completa, preferencialmente, que possam ser encontrados na internet; reutilização, ou seja, sem restrições para que sejam reutilizados e compartilhados; e participação universal, ou seja, sem restrições para utilização por certos grupos e indivíduos, quaisquer que sejam seus campos de atuação.

Para melhor compreender o conceito, destacamos as Três Leis dos Dados Abertos definidas por David Eaves⁷, professor da Harvard Kennedy School:

1. Se os dados não podem ser encontrados ou indexados na Web, eles não existem.
2. Se os dados não estão disponíveis em um formato aberto e legível por máquina, eles não podem ser reutilizados.
3. Se dispositivos legais não permitem que os dados sejam partilhados, eles não são úteis.

A definição dos princípios de dados abertos foi estabelecida em 2007⁸, em Sebastopol, Califórnia, durante um encontro que contou com a participação de pesquisadores e representantes de organizações da sociedade civil, entre outros. Os principais objetivos foram debater

⁵ Disponível em: <https://br.okfn.org/dados-abertos/>. Acesso em: 10 set. 2020.

⁶ Disponível em: <http://ceweb.br/guias/dados-abertos/capitulo-2/>. Acesso em: 10 set. 2020.

⁷ Disponível em: <https://eaves.ca/2009/09/30/three-law-of-open-government-data/>. Acesso em: 10 set. 2020.

⁸ Disponível em: <https://opengovdata.org/>. Acesso em 10 set. 2020.

os aspectos principais sobre abertura dos dados governamentais e, ao final do encontro, chegar a um consenso sobre as diretrizes a serem seguidas para publicação de dados⁹.

A seguir, apresentam-se os oito princípios de dados abertos¹⁰:

- 1. Completos:** Todos os dados públicos devem ser disponibilizados. Ou seja, se o dado não está sujeito a limitações de privacidade, segurança ou controle de acesso, ele deve estar disponível aos cidadãos;
- 2. Primários:** Os dados devem ser publicados em seu formato original, sem modificações e com o maior detalhamento possível;
- 3. Atuais:** Os dados devem ser disponibilizados tempestivamente, pois somente geram valor se ainda forem relevantes para os que os utilizam;
- 4. Acessíveis:** Os dados devem ser disponibilizados para o público mais amplo possível e para os propósitos mais variados possíveis;
- 5. Processáveis por máquina:** Os dados devem ser razoavelmente estruturados para possibilitarem o seu processamento automatizado;
- 6. Acesso não discriminatório:** Os dados devem estar disponíveis a todos, sem que sejam necessários identificação ou registro;
- 7. Formatos não-proprietários:** Os dados devem estar disponíveis em um formato sobre o qual nenhum ente tenha controle exclusivo;
- 8. Livres de licenças:** Os dados não devem estar sujeitos a regulações de direitos autorais, marcas, patentes ou segredo industrial. Restrições razoáveis de privacidade, segurança e controle de acesso podem ser permitidas na forma regulada por estatutos.

Nesse mesmo sentido, a Carta Internacional de Dados Abertos (Open Data Charter)¹¹, desenvolvida em 2015, por governos nacionais, sociedade civil e especialistas na área de dados abertos, apresenta seis princípios que representam o consenso, em nível internacional, para nortear a publicação de dados em formato aberto. São eles:

⁹ Disponível em: <https://br.okfn.org/2015/04/14/quando-surgiram-os-oito-principios-dos-dados-abertos/>. Acesso em: 10 set. 2020.

¹⁰ Disponível em: <https://opengovdata.org/>. Acesso em: 10 set. 2020.

¹¹ Disponível em: <https://opendatacharter.net/principles/>. Acesso em: 10 set. 2020.

1. **“Open by default”**: presunção de publicação de todos os dados públicos, salvo se protegidos por sigilo;
2. **Oportunos e compreensíveis**: o dado aberto somente é útil se ainda for relevante. Sempre que possível, o dado deve ser publicado em seu formato original, sem modificações;
3. **Acessível e utilizável**: deve-se assegurar que o dado possa ser processado por máquinas e que seja de fácil localização;
4. **Comparável e interoperável**: quanto mais os datasets possuírem alta qualidade e puderem interagir entre si, maior potencial poderá ser extraído pelos usuários;
5. **Para melhorar a governança e a participação social**: dados abertos permitem que os cidadãos e os próprios representantes dos governos saibam como atuam os políticos e os agentes públicos. A transparência é capaz de promover a melhoria de serviços públicos e a responsabilização de governantes;
6. **Para desenvolvimento inclusivo e inovação**: dados abertos podem ajudar a desenvolver uma economia inclusiva. Também podem ser utilizados por empresas no desenvolvimento de seus negócios.

Assim, é importante que os dados publicados em formato aberto se adequem às diretrizes expostas nesses princípios para que possibilitem a geração de benefícios concretos aos cidadãos.

Por fim, é importante abordar a representação das cinco estrelas para publicação de dados abertos¹², elaborada pelo fundador da Web, Tim Bernes-Lee. O raciocínio esquematizado

¹² Disponível em: <http://5stardata.info/en/>. Acesso em: 10 set. 2020.

de Bernes-Lee mostra que dados publicados se tornam mais úteis e de fáceis aproveitamento e reutilização cada vez que alcançam uma nova estrela. Seguem as explicações¹³:

★ **(1 estrela):** Neste estágio, basta que os dados estejam disponíveis na Web, independentemente do formato, porém, com licença aberta (ex: PDF sob licença aberta);

★★ **(2 estrelas):** Além da condição anterior, os dados estão publicados de forma estruturada e legível por máquina (ex: arquivo em Excel);

★★★ **(3 estrelas):** Compreende as condições anteriores, mas, agora, utilizando formato não-proprietário (ex: arquivo em CSV em substituição do arquivo em Excel);

★★★★ **(4 estrelas):** Nessa etapa, o dado deve estar em formato aberto e, adicionalmente, deve ser publicado na Web com URI que descreva cada um;

★★★★★ **(5 estrelas):** Compreende todas as condições anteriores, acrescentando-se a condição de que os dados possam ser conectados a outras bases.

Conforme o esquema, o dado é considerado efetivamente aberto quando alcança o nível de três estrelas. As demais estrelas fornecem mais possibilidades para os que utilizam esses dados, uma vez que podem navegar mais facilmente por eles e encontrar novas informações e relações.

¹³ Disponível em: <https://br.okfn.org/2013/01/17/maturidade-em-dados-abertos-entenda-as-5-estrelas/>. Acesso em: 10 set. 2020.

1.2 Benefícios da publicação de dados abertos

A publicação de dados governamentais em formato aberto permite que a sociedade se aproprie desses dados para, então, gerar valor à coletividade. Assim, a abertura de dados se apresenta como uma ferramenta essencial para a consolidação de um Estado transparente e para o fortalecimento dos processos democráticos¹⁴.

Muitas organizações e indivíduos se beneficiam da disponibilidade desses dados, incluindo o próprio governo¹⁵. Certas análises ou produtos gerados pela sociedade, pelos órgãos públicos ou por pesquisas científicas a partir de dados abertos fornecem subsídios para a melhoria de política públicas e ações governamentais.

Abaixo, elencam-se alguns benefícios propiciados pela utilização dos dados abertos:

- Incremento do controle social diante das ações do governo;
- Geração de novos negócios;
- Transparência na prestação de contas dos resultados e das ações da gestão pública;
- Economia de tempo para atendimento a pedidos por dados e informações governamentais;
- Possibilidade de desenvolvimento de pesquisas científicas a partir dos dados disponibilizados;
- Promoção de benefícios socioambientais e desenvolvimento sustentável;
- Promoção do desenvolvimento científico, econômico e social do país.

¹⁴ Disponível em: <http://ceweb.br/guias/dados-abertos/capitulo-4/>. Acesso em: 10 set. 2020.

¹⁵ *Manual dos dados abertos: governo*. Disponível em: https://www.w3c.br/pub/Materiais/PublicacoesW3C/Manual_Dados_Abertos_WEB.pdf. Acesso em: 10 set. 2020.

1.3 Política de Dados Abertos do Poder Executivo Federal¹⁶

A Política de Dados Abertos do Poder Executivo Federal tem o objetivo de aprimorar a cultura de transparência pública por meio da disponibilização de dados em formato aberto, uma vez que estabelece regras para publicação de dados produzidos ou acumulados pelos órgãos do Poder Executivo Federal. O órgão responsável pela gestão e pelo monitoramento da Política de Dados Abertos é a Controladoria-Geral da União (CGU), por meio da Infraestrutura Nacional de Dados Abertos (INDA)¹⁷.

A obrigatoriedade do 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.

O Plano de Dados Abertos (PDA) é o instrumento que operacionaliza a Política de Dados Abertos, pois organiza o planejamento das ações de implementação e promoção da abertura de dados dos órgãos. Adicionalmente, com o objetivo de adequar a oferta de dados às demandas sociais por esses dados, os órgãos devem utilizar mecanismos de participação social para priorizar as bases de dados que serão abertas. Outras obrigações se encontram disponíveis no Decreto nº 8.777 de 2016 e na Resolução nº 03 do Comitê Gestor da INDA (CGINDA)¹⁸.

Segue o passo a passo recomendado aos órgãos para a construção de seus PDAs, conforme Ary, Carvalho e Wosgrau (2019):

1. Realize **discussões** com as áreas finalísticas do órgão para compartilhar o contexto da Política de Dados Abertos e definir responsáveis pela elaboração e cumprimento do PDA. Um Grupo de Trabalho - GT pode ser criado, se o órgão desejar, para melhor conduzir os trabalhos;
2. Liste todas as bases de dados do órgão/entidade, por secretaria/departamento, para a elaboração de um **inventário único de dados**;

¹⁶ Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2016/decreto/d8777.htm. Acesso em: 10 set. 2020.

¹⁷ Disponível em: <https://www.gov.br/governodigital/pt-br/legislacao/InstrucaoNormativaINDA42012.pdf>. Acesso em: 10 set. 2020.

¹⁸ Disponível em: <http://wiki.dados.gov.br/Comite-Gestor-da-INDA.ashx>. Acesso em: 10 set. 2020.

3. Adote um ou mais mecanismos de **participação social** para auxiliar na definição e na priorização de abertura de bases de dados, conforme a demanda e desejo da sociedade. Sugestões de mecanismos: consulta pública no site do órgão ou em sites especializados;

4. Priorize a abertura de bases do órgão utilizado, minimamente, os critérios listados no ART.1º da Resolução nº3 do CGINDA:

- I) Grau de relevância para o cidadão (descoberto via consulta pública*);
- II) Estímulo ao controle social;
- III) Obrigatoriedade legal ou compromisso assumido de disponibilização daquele dado;
- IV) Dado se referir a projetos estratégicos do governo;
- V) Dado demonstrar resultados diretos e efetivos dos serviços públicos disponibilizados ao cidadão pelo Estado;
- VI) Sua capacidade de fomento ao desenvolvimento sustentável;
- VII) Possibilidade de fomento a negócios na sociedade;
- VIII) Dados mais solicitados em transparência passiva desde o início da vigência da Lei nº 12.527, de 18 de novembro de 2011 - Lei de Acesso à Informação.

*Obs: Acesse a matriz de priorização sugerida no **Manual de Elaboração de Planos de Dados Abertos**.

5. Defina a **relação final das bases** que serão abertas durante o período de vigência do PDA e elabore um cronograma para abertura dessas bases;

Nome da base de dados	Descrição	Unidade e contato do responsável pela base	Meta/prazo para abertura

6. Defina **estratégias e cronogramas** para a **publicação, sustentação, catalogação e difusão** dessas bases;

7. Registre, em linguagem simples e objetiva, todas as ações estratégicas definidas nos passos anteriores em um documento, denominado **Plano de Dados Abertos/nome do órgão/vigência do PDA**, seguindo a seguinte estrutura:

- I) Introdução;
- II) Cenário institucional;

- III) Objetivos;
- IV) Construção e execução do PDA;
- V) Dados selecionados para abertura;
- VI) O processo de catalogação;
- VII) Sustentação e Governança;
- VIII) Monitoramento e controle;
- IX) Plano de Ação;
- X) Anexos.

8. Publique o PDA no Portal do órgão na seção "Acesso à Informação" > "Dados Abertos".

Os principais instrumentos que regulam a Política são:

- Lei n.º 12.527/2011 – Lei de Acesso à Informação, em especial, o art. 8º; e Decreto n.º 7.724/2012, que a regulamenta no Poder Executivo Federal;
- Instrução Normativa SLTI/MP n.º 4/2012, que institui a Infraestrutura Nacional de Dados Abertos (INDA);
- Decreto n.º 8.777/2016, que institui a Política de Dados Abertos do Poder Executivo Federal;
- Decreto n.º 9.903/2019, que transfere a gestão da Política de Dados Abertos do Poder Executivo Federal para Controladoria-Geral da União (CGU);
- Resolução n.º 3/2017, do Comitê Gestor da Infraestrutura Nacional de Dados Abertos (CGINDA).

Por fim, cabe destacar que o modelo utilizado pelo Poder Executivo Federal para incentivar a abertura de dados pode ser replicado no âmbito dos entes federativos e demais poderes públicos. O Distrito Federal, por exemplo, estabeleceu sua política de dados abertos (Decreto n.º 38.354/2017¹⁹) nos moldes da política desenvolvida no âmbito federal.

¹⁹ Disponível em: http://www.sinj.df.gov.br/sinj/Norma/2a90db6875624a65936a47e18e1c337b/Decreto_38354_24_07_2017.html. Acesso em: 10 set. 2020.

2. Publicação de Dados

Este tópico irá dispor sobre aspectos relativos à publicação de dados abertos, abordando os parâmetros recomendados para catalogação, bem como critérios relevantes para propiciar maior interatividade aos consumidores de dados. O grande volume de dados governamentais disponíveis atinge o status de realmente abertos quando ocorre uma comunicação social bidirecional entre a administração pública e a população. Um dos pontos-chave para tornar essa comunicação efetiva é a disponibilização de dados por meio de ferramentas que possibilitem compreender seu significado e sua reutilização.

2.1 Catalogação dos Dados

A catalogação de dados é importante para o acesso à informação por parte dos usuários ou das máquinas. Para essa catalogação ocorrer de forma efetiva, é necessário que a informação esteja clara, objetiva e estruturada, com metadados completos, garantindo a qualidade e confiabilidade dos dados.

Existem diferentes formas e técnicas para catalogar e publicar dados na Web. O livro *Fundamentos para publicação de dados na Web*²⁰ apresenta algumas dessas técnicas:

- Utilizar Web APIs (Interface de Programação de Aplicações). Atualmente, existe um novo tipo de API chamado RESTful service. “É possível dizer que os dados disponíveis por meio de APIs são semelhantes aos dados disponíveis na Deep Web²¹, ou seja, não podem ser facilmente encontrados e indexados”²²;

²⁰ Disponível em: <https://ceweb.br/publicacao/livro-fundamentos-dados-web/>. Acesso em: 10 set. 2020.

²¹ Deep Web é o nome que se dá à camada da Internet que não pode ser acessada através de mecanismos de busca, pois não há indexação.

²² Disponível em: <https://ceweb.br/publicacao/livro-fundamentos-dados-web/>. Acesso em: 10 set. 2020.

- Fazer a inclusão dos dados nas páginas HTML, com o uso do RDFa²³ e, também, da linguagem JSON-LD²⁴;
- Usar plataformas para catalogação de dados, como CKAN²⁵, Socrata²⁶ e ArcGIS Open Data²⁷.

2.1.1 Plataformas de gerenciamento de dados abertos mais utilizadas²⁸

A catalogação de dados é importante para o acesso à informação por parte dos usuários ou das máquinas. Para essa catalogação ocorrer de forma efetiva, é necessário que a informação esteja clara, objetiva e estruturada, com metadados completos, garantindo a qualidade e confiabilidade dos dados.

Plataforma de Dados Abertos, do inglês *Open Data Platform* (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)²⁹.

²³ Rich Structured Data Markup for Web Documents é um padrão do Consórcio World Wide Web (W3C) que permite a especificação de atributos para descrição de dados estruturados em qualquer linguagem de marcação, em particular, XHTML e HTML. Disponível em: <https://www.w3.org/TR/rdfa-primer/>. Acesso em: 10 set. 2020.

²⁴ JSON for Linked Data é um padrão do W3C baseado no formato JSON, que permite o uso de vocabulários e ontologias para a descrição dos dados. Disponível em: <https://www.w3.org/TR/json-ld/>. Acesso em: 10 set. 2020.

²⁵ Comprehensive Knowledge Archive Network. Disponível em: <https://ckan.org/>. Acesso em: 10 set. 2020.

²⁶ Socrata. Disponível em: <https://dev.socrata.com/>. Acesso em: 10 set. 2020.

²⁷ ArcGIS é um software de mapeamento e análise de dados. Disponível em: <https://www.esri.com/en-us/arcgis/about-arcgis/overview>. Acesso em: 10 set. 2020.

²⁸ Texto reproduzido do Documento Modelos de abertura de dados: potencialidades e limites. Disponível em: <http://governoaberto.cgu.gov.br/a-ogp/planos-de-acao/4o-plano-de-acao-brasileiro/compromisso-2-docs/marcos-2-e-3-modelos-de-abertura-de-dados.pdf>. Acesso em: 10 set. 2020.

²⁹ É importante mencionar que o acrônimo ODP também é utilizado em, pelo menos, duas outras situações ou significados distintos: a) como *Open Data Platform*, uma iniciativa homônima do setor de Tecnologia da Informação (TI) que fornece um conjunto comum de ferramentas e tecnologias para a plataforma *Apache Hadoop* (Disponível em: <https://hadoop.apache.org/>. Acesso em: 10 set. 2020). 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; e b) como *Open Data Portal*, denominação utilizada pelo portal de dados abertos da União Europeia (UE) (Disponível em: <https://data.europa.eu/euodp/en/home>. Acesso em: 10 set. 2020). O portal contém conjuntos de dados recolhidos e publicados pela UE.

As ODPs devem apresentar as seguintes características:

1. As plataformas devem ser construídas especificamente para o gerenciamento de dados abertos, ter alguma base instalada e ter a funcionalidade de pesquisa/busca de datasets por campos de metadados. Não basta proverem uma estrutura comum de portal na Web. Como exemplo de boa prática, o site *Data Portals*³⁰ 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.

São descritas, aqui, as plataformas de gerenciamento de dados mais utilizadas.

Comprehensive Knowledge Archive Network (CKAN) 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 Foundation³¹ e disponibilizada como uma estrutura de código aberto (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, no uso de formatos abertos e na sua alta flexibilidade, visto que a plataforma é disponibilizada em código aberto. Por outro lado, a CKAN não foi pensada para permitir seleções das bases de dados; ela não permite filtros e não gera visualizações por default. Isso se deve à orientação da plataforma, focada nos conjuntos de dados. Por isso, pode não ser amigável para um usuário que não tenha interesse em acessar os dados crus ou deseje uma informação de consumo mais rápido. No entanto, por ser de código aberto, diversas modificações podem ser implementadas quando o modelo é aplicado.

Lançada em 2007, a plataforma Socrata foi considerada a principal solução para o gerenciamento de dados abertos antes do crescimento da CKAN. Com avaliação similar à da CKAN, de acordo com Osagie *et al.* (2015), a Socrata também apresenta onze das doze

³⁰ *Data Portals*. Disponível em: <http://dataportals.org/>. Acesso em: 10 set. 2020.

³¹ Disponível em: <https://okfn.org/>. Acesso em: 10 set. 2020.

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 a visualização de dados 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.

A ArcGIS Open Data, plataforma de mapeamento e análise de dados desenvolvida pela ESRI³², 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 de mapas (CORREA; ZANDER; SILVA, 2018).

O Frictionless Data³³ é um conjunto de ferramentas e estratégias focado em facilitar a interoperabilidade de dados por meio da definição de padrões e critérios técnicos e com o objetivo de otimizar o armazenamento e os usos de dados. Nesse sentido, o projeto oferece uma série de bibliotecas, aplicações e instruções em diferentes linguagens para aplicar os padrões Frictionless, facilitando a gestão de metadados, a limpeza, a organização, a extração, o compartilhamento, entre outros recursos.

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 OpenML³⁴, 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, acessíveis, colaborativos e abertos, 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.

³² Environmental Systems Research Institute. Disponível em: <https://www.esri.com/en-us/home>. Acesso em: 10 set.2020.

³³ Disponível em: <https://frictionlessdata.io/>. Acesso em: 10 set. 2020.

³⁴ Disponível em: <https://www.openml.org/>. Acesso em: 10 set. 2020.

2.1.2 Definição de metadados obrigatórios e facultativos³⁵

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: XML³⁶, CSV³⁷, JSON³⁸, XLS³⁹, PDF⁴⁰ e HTML⁴¹. Os formatos legíveis por máquina mais comuns são: CSV, XML, JSON, GeoJSON, XLS. O formato RDF⁴² permite que os dados possam ser consultados via SPARQL⁴³.

³⁵ Textos reproduzidos da seção sobre metadados da *Recomendação Boas Práticas para Dados na Web (DWBP)*. Disponível em: <https://w3c.br/traducoes/DWBP-pt-br/#metadata>. Acesso em: 10 set. 2020.

³⁶ Extensible Markup Language. É um formato baseado em texto para representar informações estruturadas de documentos, dados, configuração, livros, transações, faturas e outros documentos. É uma recomendação do W3C. Disponível em: <https://www.w3.org/XML/Schema>. Acesso em: 10 set. 2020.

³⁷ Comma-Separated Values. Representa arquivos de texto de formato aberto regulamentado pelo RFC 4180, que faz uma ordenação de bytes ou um formato de terminador de linha, separando valores com vírgulas. Usado em software de planilha e em pacotes de análise de dados. Disponível em: <https://tools.ietf.org/html/rfc4180>. Acesso em: 10 set. 2020.

³⁸ JavaScript Object Notation. É um padrão aberto para troca de dados entre sistemas. Disponível em: <https://www.json.org/>. Acesso em: 10 set. 2020.

³⁹ XLS. Formato proprietário de arquivo utilizado pelo software Excel/Microsoft em versões anteriores a 2012. Nas versões mais recentes, o formato é o XLSX. Disponível em: <https://www.reviversoft.com/pt/file-extensions/xls>. Acesso em: 10 set. 2020.

⁴⁰ Portable Document Format. É um padrão gratuito para visualizar, imprimir, compartilhar e comentar documentos. Disponível em: <https://acrobat.adobe.com/br/pt/acrobat/pdf-reader.html>. Acesso em: 10 set. 2020.

⁴¹ HyperText Markup Language. É uma linguagem de marcação utilizada na construção de páginas na Web. Documentos em HTML podem ser interpretados por navegadores. Disponível em: <https://www.w3.org/standards/webdesign/htmlcss>. Acesso em: 10 set. 2020.

⁴² Resource Description Framework. É o modelo padrão para intercâmbio de dados na Web; padrão mantido pelo W3C. Disponível em: <https://www.w3.org/RDF/>. Acesso em: 10 set. 2020.

⁴³ SPARQL. É uma linguagem de consulta semântica para bancos de dados que permite recuperar e manipular dados armazenados no formato RDF; padrão mantido pelo W3C. Disponível em: <https://www.w3.org/TR/sparql-11-overview/>. Acesso em: 10 set. 2020.

Para catalogar os dados abertos, sugere-se definir quais metadados são obrigatórios e quais são facultativos. Com base na *Recomendação Boas Práticas para Dados na Web* (DWBP)⁴⁴, sugerimos que os metadados a seguir sejam obrigatórios:

- Título;
- URI;
- Palavras-chave;
- Data da publicação;
- Publicador;
- Autoria;
- Ponto de contato (e-mail e/ou telefone);
- Período da publicação;
- Frequência de atualização;
- Categoria (tema);
- Idioma;
- Formato do recurso publicado;
- Versão (status);
- Licença;
- Vocabulário utilizado (URI do dicionário).

2.1.3 Definição da licença dos dados publicados⁴⁵

Disponibilizar informações sobre a licença⁴⁶ é fundamental para que os consumidores de dados avaliem a utilidade dos mesmos. Os agentes de software podem utilizar a presença ou a ausência das informações de licença como um gatilho para a inclusão ou a exclusão de dados apresentados a um consumidor em potencial.

A Boa Prática 4 do documento DWBP define que é preciso fornecer informações sobre a licença de dados por meio de um link ou de uma cópia dos termos da licença que controla a utilização dos dados. Dessa forma, os consumidores serão capazes de compreender a licença de dados que descreve eventuais restrições impostas à utilização de certos dados, e as má-

⁴⁴ Data on the Web Best Practices, traduzida para o português como *Boas Práticas para Dados na Web*. Disponível em: <https://w3c.br/traducoes/DWBP-pt-br/>. Acesso em: 10 set. 2020.

⁴⁵ Textos reproduzidos da seção sobre licenças de dados da *Recomendação Boas Práticas para Dados na Web* (DWBP). Disponível em: <https://w3c.br/traducoes/DWBP-pt-br/#licenses>. Acesso em: 10 set. 2020.

⁴⁶ Disponível em: <https://w3c.br/traducoes/DWBP-pt-br/#license>. Acesso em: 10 set. 2020.

quinas serão capazes de detectar automaticamente a licença de dados de uma distribuição. No contexto dos dados na Web, a licença de um conjunto de dados pode ser especificada dentro ou fora dos metadados ou em um documento separado ao qual esteja conectado.

2.1.4 Informações sobre a qualidade dos dados⁴⁷

Documentar a qualidade dos dados facilita significativamente o processo de seleção de conjuntos de dados, aumentando as chances de reutilização dos mesmos. Independentemente das peculiaridades específicas do domínio, a qualidade dos dados deve ser documentada, e os problemas de qualidade que são conhecidos devem ser declarados de forma explícita nos metadados.

A Boa Prática 6 do documento DWBP descreve a importância de prover informações sobre a qualidade de dados e a adequação para fins específicos. Dessa forma, usuários e máquinas (agentes de software) serão capazes de avaliar a qualidade e, portanto, a adequação de um conjunto de dados para a sua aplicação.

2.1.5 Informações de procedência de dados⁴⁸

Saber qual é a procedência dos dados é uma das formas que os consumidores dispõem para julgar a qualidade de um conjunto de dados. Entender sua origem e sua história auxilia a determinar se o dado é confiável e fornece um contexto importante para interpretação dos dados.

A Boa Prática 5 do documento DWBP estabelece a importância de fornecer informações de procedência dos dados. As informações devem ser completas sobre as origens dos dados e de quaisquer alterações que tenham sido feitas. Dessa forma, usuários saberão a origem e

⁴⁷ Textos reproduzidos da seção sobre qualidade dos dados da *Recomendação Boas Práticas para Dados na Web* (DWBP). Disponível em: <https://w3c.br/traducoes/DWBP-pt-br/#quality>. Acesso em: 10 set. 2020.

⁴⁸ Textos reproduzidos da seção sobre procedência de dados da *Recomendação Boas Práticas para Dados na Web* (DWBP). Disponível em: <https://w3c.br/traducoes/DWBP-pt-br/#provenance>. Acesso em: 10 set. 2020.

o histórico do conjunto de dados, e as máquinas (agentes de software) poderão processar automaticamente a informação de procedência.

2.1.6 Indicador da versão dos dados e histórico das versões⁴⁹

As informações sobre a versão dos dados tornam a revisão de um conjunto de dados identificável de forma única. Essa unicidade pode ser utilizada pelos consumidores de dados para determinar especificamente com qual versão de um conjunto de dados estão trabalhando. O bom versionamento dos dados possibilita aos consumidores entender se uma nova versão de um conjunto de dados está disponível. O versionamento explícito leva em conta a repetitividade na pesquisa, permite comparações e evita confusão. A utilização de números únicos de versão com uma abordagem padronizada também pode contribuir na expectativas do consumidor acerca da diferença entre versões.

Ao criar aplicações que utilizam dados, é útil compreender a variação desses ao longo do tempo. A interpretação de dados também é potencializada pela compreensão de sua dinâmica. Determinar como diversas versões de um conjunto de dados se diferenciam umas das outras é geralmente muito trabalhoso, a não ser que um sumário dessas diferenças seja fornecido.

As Boas Práticas 7 e 8 do documento DWBP estabelecem que é preciso fornecer indicador de versão ao publicar dados na Web. Assim, deve se designar e se indicar um número de versão ou data para cada conjunto de dados. Dessa forma, pessoas e máquinas (agentes de software) poderão facilmente identificar com qual versão do conjunto de dados estão trabalhando.

Além disso, o publicador deve fornecer o histórico de versão completo e explicar as alterações feitas em cada versão. Com isso, usuários e máquinas (agentes de software) serão capazes de entender como o conjunto de dados se altera tipicamente de versão a versão, assim como poderão especificar como duas versões em particular se diferenciam uma da outra.

⁴⁹ Textos reproduzidos da seção sobre versionamento de dados da *Recomendação Boas Práticas para Dados na Web* (DWBP). Disponível em: <https://w3c.br/traducoes/DWBP-pt-br/#dataVersioning>. Acesso em: 10 set. 2020.

2.1.7 Definição de identificadores de dados⁵⁰

Adotar um sistema de identificação comum permite a identificação básica dos dados e a comparação dos processos por qualquer um dos atores envolvidos de forma confiável. Essas identificação e comparação são pré-condições essenciais para o gerenciamento e reutilização dos dados de forma adequada.

Os desenvolvedores podem construir URIs dentro de seus códigos, e, para isso, é importante que tais URIs sejam persistentes e que sejam referenciadas para o mesmo recurso ao longo do tempo, sem a necessidade de intervenção humana.

As Boas Práticas 9, 10 e 11 do documento DWBP explicam que é necessário usar URIs persistentes como identificadores de conjuntos de dados. Dessa maneira, os conjuntos de dados ou as informações sobre os conjuntos de dados serão encontrados e citados com facilidade em qualquer momento, independentemente do status, da disponibilidade ou do formato dos dados.

Também é preciso usar URIs persistentes como identificadores dentro de conjuntos de dados e reutilizar URIs de terceiros como identificadores dentro dos conjuntos de dados quando possível. Isso significa utilizar os mesmos identificadores por meio de diversos conjuntos de dados e garantir que os seus identificadores possam ser referidos por outros conjuntos de dados. Quando tais identificadores são URIs HTTP, eles podem ser consultados, e mais dados podem ser descobertos; pois os itens dos dados serão relacionados por meio da Web, criando um espaço de informação global, acessível tanto para pessoas quanto para máquinas.

Além disso, deve se atribuir URIs para as versões dos conjuntos de dados e séries, pois usuários e máquinas (agentes de software) serão capazes de recorrer a versões específicas de um conjunto de dados e a conceitos, tais como “séries de conjuntos de dados” e “a versão mais recente”.

⁵⁰ Textos reproduzidos da seção sobre identificadores de dados da *Recomendação Boas Práticas para Dados na Web* (DWBP). Disponível em: <https://w3c.br/traducoes/DWBP-pt-br/#DataIdentifiers>. Acesso em: 10 set. 2020.

2.1.8 Dados disponibilizados em diversos formatos⁵¹

Disponibilizar dados em um formato que não seja legível por máquinas impõe severas limitações para a utilidade dos dados. Os dados se tornam úteis ao serem processados e transformados em informação. Há uma importante distinção entre formatos que podem ser lidos e editados por pessoas utilizando um computador e formatos que são legíveis por máquinas. O último termo implica que os dados são prontamente extraídos, transformados e processados por um computador.

Utilizar formatos de dados não padronizados é caro e ineficiente, e os dados podem perder significado conforme são transformados. Em contrapartida, formatos de dados padronizados viabilizam tanto a interoperabilidade quanto usos futuros, tais como a remixagem ou a visualização, muitos dos quais não podem ser antecipados quando os dados são publicados pela primeira vez. Também é importante observar que a maior parte dos formatos padronizados que são legíveis por máquinas também é neutra no que se refere à localidade.

Dados cujos conteúdos são legíveis por máquinas e não são específicos de nenhum idioma ou cultura em particular são mais duráveis e menos suscetíveis a interpretações errôneas do que dados que usam uma das muitas possíveis interpretações culturais em seus conteúdos. Coisas como datas, moedas e números podem soar parecidas, mas têm significados diferentes em locais distintos. Por exemplo, a data 4/7 pode ser lida tanto como 04 de julho quanto como 07 de abril, dependendo de onde os dados foram criados. Da mesma forma, €2,000 pode significar tanto dois mil Euros quanto uma representação excessivamente precisa de dois Euros. Utilizar um formato de localidade neutra evita a necessidade de estabelecer regras intercambiáveis específicas que variam de acordo com o idioma ou com a localização do usuário. Quando os dados já estão em um formato de localidade específica, explicitar o local e a linguagem por meio do fornecimento de parâmetros de localidade permite aos usuários determinar se poderão trabalhar com os dados e habilitar serviços de tradução automatizados prontamente ou não.

Além disso, disponibilizar dados em mais de um formato reduz os custos decorrentes da transformação de dados e minimiza a possibilidade de erros no processo de transformação.

⁵¹ Textos reproduzidos da seção sobre formatos de dados da *Recomendação Boas Práticas para Dados na Web* (DWBP). Disponível em: <https://w3c.br/traducoes/DWBP-pt-br/#dataFormats>. Acesso em: 10 set. 2020.

Caso muitos usuários precisem transformar os dados em um formato específico, publicá-los nesse formato desde o início poupa tempo, dinheiro e evita erros de uma forma mais eficiente. Por último, disponibilizar dados em diversos formatos aumenta o número de ferramentas e aplicações que podem processar os dados.

As Boas Práticas 12, 13 e 14 do documento DWBP especificam que devemos usar formatos de dados padronizados e legíveis por máquinas, de forma que as máquinas serão capazes de ler e processar facilmente os dados publicados na Web e as pessoas poderão utilizar as ferramentas computacionais usualmente disponíveis para trabalhar com dados.

Elas também dizem que é preciso usar representações de dados que sejam independentes de localidade (locale neutral). Quando isso não for possível, deve se fornecer metadados sobre os parâmetros regionais dos conteúdos de dados, pois, assim, os usuários e as máquinas (agentes de software) serão capazes de interpretar o significado de strings que representam datas, horários, moedas, números etc. de forma precisa.

Finalmente, mas não menos importante, é necessário fornecer dados em múltiplos formatos. Assim, o maior número possível de usuários será capaz de utilizar os dados sem ter que, primeiramente, transformá-los para seu formato de preferência.

2.1.9 Vocabulários de dados⁵²

Vocabulários⁵³ definem os conceitos e as relações (também conhecidas como ‘termos’ ou ‘atributos’) utilizados para descrever e representar uma área de interesse.

A utilização de vocabulários já em uso por outras pessoas estimula e facilita o consenso em comunidades. Aumenta a interoperabilidade e reduz as redundâncias, incentivando, assim, a reutilização de seus próprios dados. Particularmente, a aplicação de vocabulários compartilhados para metadados – especialmente os metadados estruturais, de procedência, de qualidade e de versionamento – auxilia o processo de comparação e o processamento automático

⁵² Textos reproduzidos da seção sobre vocabulários de dados da *Recomendação Boas Práticas para Dados na Web* (DWBP). Disponível em: <https://w3c.br/traducoes/DWBP-pt-br/#dataVocabularies>. Acesso em: 10 set. 2020.

⁵³ Definição em inglês disponível em: <https://www.w3.org/standards/semanticweb/ontology>. Acesso em: 10 set. 2020.

tanto dos dados quanto dos metadados. Além disso, a referência a códigos e termos padronizados ajuda a evitar ambiguidade e conflitos entre elementos e valores similares.

Escolher um vocabulário muito simples é sempre atrativo, mas há um perigo: o desejo de manter a simplicidade pode induzir o publicador a omitir alguns dados que fornecem informações importantes, tais como a localização geográfica dos pontos de ônibus, o que impediria que fossem exibidos em um mapa. Portanto, o equilíbrio deve ser atingido tendo em mente que o objetivo não é simplesmente compartilhar seus dados, mas que outros possam reutilizá-los.

As Boas Práticas 15 e 16 do documento DWBP explicam que é necessário reutilizar vocabulários compartilhados, dando preferência aos padronizados, para codificar dados e metadados. Dessa forma, melhorará a interoperabilidade e o consenso entre os publicadores e consumidores de dados.

Escolher o nível de formalização adequado também é importante para o vocabulário se ajustar tanto aos dados quanto às aplicações mais prováveis de serem utilizadas. Assim, as aplicações mais prováveis serão suportadas sem um grau de complexidade maior do que o necessário.

2.1.9.1 Vocabulário do Catálogo de Dados (DCAT)

Criado pelo W3C, o Vocabulário do Catálogo de Dados – em inglês, Data Catalog Vocabulary (DCAT) – é um vocabulário controlado que fornece às pessoas e às máquinas uma abordagem específica e independente de domínio para criar catálogos que expressam os principais elementos de uma descrição de conjunto de dados de maneira padronizada e adequada para publicação na Web. Permite a interoperabilidade entre domínios, sendo usada por conta própria ou paralelamente, como complemento de outros padrões de catálogo de dados. O DCAT é projetado, portanto, para facilitar a interoperabilidade entre catálogos de dados publicados na Web.

O DCAT atende a, pelo menos, sete das recomendações de Boas Práticas de Dados na Web (1, 2, 3, 6, 12, 15 e 16), sendo uma opção válida de vocabulário, que prioriza a reutilização e a redução de redundâncias dos dados.

A versão 2⁵⁴ está disponível, e o Grupo de Trabalho do W3C para a padronização de Interação de Conjuntos de Dados – em inglês, Dataset Exchange Working Group (DXWG)⁵⁵ – está trabalhando na terceira versão do DCAT.

2.1.10 Acesso aos dados

As Boas Práticas 17, 18, 19, 20, 21 e 22 do documento DWBP tratam sobre o acesso aos dados⁵⁶.

Acessar os dados em massa pode ser útil quando dados na Web estiverem distribuídos através de diferentes URIs e puderem ser organizados logicamente como um pacote. O acesso em massa garante uma forma consistente de tratar os dados como um conjunto de dados. Acessar dados individualmente, ao longo de muitas consultas, pode ser difícil, e, caso forem usados para remontar o conjunto de dados completo, isso pode levar a uma manipulação inconsistente dos dados. Assim, o publicador pode fornecer download em massa (bulk download), permitindo que os consumidores acessem o conjunto de dados completo em uma única solicitação. Dessa maneira, transferências de arquivos grandes, que exigiriam mais tempo do que um usuário típico consideraria razoável, serão possíveis por meio de protocolos de transferência de arquivos, do inglês file-transfer protocols (FTP).

Pode ser difícil mover conjuntos de dados muito extensos de um lugar para outro. Também pode ser inconveniente, para os usuários, armazenar ou analisar um conjunto de dados grande. Os usuários não deveriam ter que fazer download de um conjunto de dados completo, caso apenas necessitem de um subconjunto do mesmo. Além disso, aplicações Web que acessam conjuntos de dados grandes operam com melhor desempenho, caso seus desenvolvedores consigam se apropriar das vantagens do carregamento lento, pois é possível trabalhar com porções menores de um todo e acessar novas porções apenas quando necessário. A possibilidade de trabalhar com subconjuntos de dados também

⁵⁴ Disponível em: <https://www.w3.org/TR/vocab-dcat-2/>. Acesso em: 10 set. 2020.

⁵⁵ Disponível em: https://www.w3.org/2017/dxwg/wiki/Main_Page#Dataset_Exchange_Working_Group_28DXWG.29. Acesso em: 10 set. 2020.

⁵⁶ Disponível em: <https://w3c.br/traducoes/DWBP-pt-br/#dataAccess>. Acesso em: 10 set. 2020.

permite que o processamento offline funcione com mais eficiência. Em particular, as aplicações em tempo real se beneficiam bastante dessa possibilidade, pois podem ser atualizadas mais rapidamente.

Dessa forma, quem publica deve fornecer subconjuntos para conjuntos de dados extensos. Com isso, pessoas e aplicações serão capazes de acessar subconjuntos necessários de um conjunto de dados em vez de desnecessariamente acessar o conjunto como um todo. Conjuntos de dados estáticos que os usuários do domínio considerem grandes demais poderão ser descarregados em porções menores. As APIs tornarão subconjuntos ou partes filtradas de dados disponíveis com granularidade, dependendo das necessidades do domínio e das demandas de desempenho da aplicação Web.

Também é possível disponibilizar dados em uma página HTML, com dados legíveis por pessoas mesclados a dados legíveis por máquina usando RDFa, por exemplo. No entanto, como a arquitetura da Web (WEBARCH) e o vocabulário DCAT (VOCAB-DCAT) deixam claro, um recurso como um conjunto de dados, por exemplo, pode ter muitas representações. Os mesmos dados podem estar disponíveis em JSON, XML, RDF, CSV e HTML. Essas representações múltiplas podem ser disponibilizadas por meio de uma API; no entanto, devem ser disponibilizadas a partir do mesmo URL, utilizando-se a negociação de conteúdo (em inglês) para retornar a representação apropriada – o que o DCAT denomina distribuição. URIs específicos podem ser usados para identificar diretamente representações individuais dos dados, ignorando a negociação de conteúdo. Sugere-se, portanto, usar negociação de conteúdo para disponibilizar dados em múltiplos formatos. A negociação de conteúdo permitirá habilitar recursos diversos ou representações diferentes do mesmo recurso a ser disponibilizado de acordo com a solicitação feita pelo cliente.

2.1.11 Atualização dos dados

A disponibilidade dos dados na Web deve corresponder rigorosamente à data de criação ou de coleta dos dados, eventualmente após esses terem sido processados ou alterados. Sincronizar cuidadosamente a publicação dos dados com a frequência de atualização incentiva a confiança do consumidor e o reuso dos dados.

A Boa Prática 21 do documento DWBP sugere que os dados fornecidos estejam sempre atualizados e que a frequência de atualização seja explícita, de tal maneira que a maior parte dos dados mais recentes disponibilizados online reflita os dados mais recentes divulgados por qualquer outro canal. Quando novos dados estiverem disponíveis, serão publicados na Web o mais rapidamente possível.

A presença de dados em tempo real na Web viabiliza o acesso a dados cronologicamente sensíveis e incentiva o desenvolvimento de aplicações Web em tempo real. O acesso em tempo real depende de os produtores de dados em tempo real disponibilizarem seus dados prontamente ao publicador de dados. A necessidade de fornecer acesso em tempo real para uma determinada aplicação precisará ser avaliada caso a caso, considerando-se as taxas de atualização, a latência introduzida pelos passos de pós-processamento de dados, a disponibilidade de infraestrutura e os dados de que os consumidores necessitam. Além de tornar os dados acessíveis, os publicadores de dados podem fornecer informações adicionais, descrevendo lacunas, erros e anomalias de dados, bem como atrasos de publicações.

A Boa Prática 20 do documento DWBP explica a importância de fornecer acesso em tempo real para que as aplicações sejam capazes de acessar dados temporalmente críticos em tempo real ou quase em tempo real – em que tempo real significa um intervalo de milissegundos até poucos segundos após a criação dos dados.

2.1.12 Utilização de APIs⁵⁷

Uma API (Interface de Programação de Aplicações ou, em inglês, Application Programming Interface) oferece aos consumidores de seus dados maiores flexibilidade e facilidade de processamento. Ela pode habilitar o uso de dados em tempo real, realizar filtragens a partir de solicitações, além de permitir trabalhar com os dados em um nível atômico. Caso o seu conjunto de dados seja grande, frequentemente atualizado ou altamente complexo, é provável que uma API seja a melhor opção para publicar seus dados.

⁵⁷ Textos reproduzidos da seção APIs de acesso a dados da *Recomendação Boas Práticas para Dados na Web* (DWBP). Disponível em: <https://w3c.br/traducoes/DWBP-pt-br/#accessAPIs>. Acesso em: 10 set. 2020.

As Boas Práticas 23, 24, 25 e 26 do documento DWBP tratam sobre o acesso aos dados por meio de uma API. É importante fornecer informações completas na Web sobre a API e atualizar a documentação conforme características adicionadas ou modificações realizadas.

2.1.13 Federação de dados

A federação de dados 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 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 dados⁵⁸.

2.1.14 Proteção e anonimização de dados pessoais/individualizados

A segurança, a sensibilidade comercial e, acima de tudo, a privacidade dos indivíduos devem ser levadas em consideração. Cabe aos publicadores de dados determinar as políticas por meio das quais os dados devem ser compartilhados, bem como sob quais circunstâncias. Espera-se que as políticas de compartilhamento de dados avaliem o risco de exposição e estabeleçam as medidas de segurança apropriadas para a proteção de dados sensíveis, tais como a autenticação segura e a autorização.

No Brasil, a Lei n.º 13.709, de 14 de agosto de 2018, intitulada “Lei Geral de Proteção de Dados Pessoais” (LGPD)⁵⁹ dispõe sobre o tratamento de dados pessoais, inclusive nos meios digitais, por pessoa natural ou por pessoa jurídica de direito público ou privado, com o objetivo de proteger os direitos fundamentais de liberdade e de privacidade e o livre desenvolvimento da personalidade da pessoa natural.

⁵⁸ Texto reproduzido do documento de referência dos *Modelos de abertura de dados: potencialidades e limites*. Disponível em: <http://governoaberto.cgu.gov.br/a-ogp/planos-de-acao/4o-plano-de-acao-brasileiro/compromisso-2-docs/marcos-2-e-3-modelos-de-abertura-de-dados.pdf>. Acesso em: 10 set. 2020.

⁵⁹ Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/L13709.htm. Acesso em: 10 set. 2020.

2.2 Usuário

Ao publicar dados em formato aberto, é necessário atentar não somente aos aspectos relativos à usabilidade do usuário, mas também às formas de interação entre os provedores e os consumidores desses dados. Nesse sentido, pretende-se apontar aspectos relevantes para facilitar a navegação e o consumo de dados pelos usuários, com a finalidade de potencializar, por exemplo, a reutilização dos dados abertos pelos cidadãos. Outros impactos positivos, sob os pontos de vista social e econômico, são a melhoria do acesso à informação, da participação e do controle social; o fomento a novos negócios; o incremento da produção científica; a simetria de informação.

2.2.1 Ferramentas de colaboração e compartilhamento

As ferramentas colaborativas possibilitam o compartilhamento de experiências entre os membros da comunidade de dados abertos por meio de cruzamentos, visualizações, softwares, estudos desenvolvidos. Os principais objetivos são reforçar a colaboração entre os atores e criar um ambiente dinâmico e interativo.

Para promover maior interatividade, é fundamental permitir a criação de comunidades sobre determinados conjuntos de dados. Tal organização facilita o diálogo entre os públicos provedor e consumidor, devendo ser complementada por mecanismos como um sistema de notificações (ou sistema push), para que o usuário possa seguir e acompanhar todas as futuras atualizações das bases; além disso, essa organização possibilita integração com as redes sociais etc.

Outra iniciativa interessante seria a criação de canais em que o usuário de dados possa compartilhar, quando achar conveniente, o uso que fez dos dados e as aplicações ou produtos que gerou. Isso possibilitaria maior divulgação dos produtos desenvolvidos a partir dos dados públicos, permitindo que a população tenha acesso às ferramentas no âmbito do controle social, da transparência da gestão pública, da geração de novos negócios, entre outros.

Como exemplo, destaca-se o Portal de Dados Abertos dos Estados Unidos⁶⁰, que permite o compartilhamento das bases de dados publicadas e dos produtos das reutilizações dos dados.

2.2.2 Feedback dos consumidores de dados e compartilhamento dos feedbacks

O feedback ajuda os provedores a melhor entenderem as necessidades dos consumidores, contribuindo também para a garantia de qualidade e integridade dos dados publicados. Consiste na opinião do público consumidor sobre diversos aspectos relacionados aos conjuntos de dados disponibilizados pelos órgãos ou entidades públicas, podendo ser elogios, questionamentos ou reclamações. O objetivo é viabilizar para os principais destinatários da ação espaços para exporem suas opiniões e experiências de uso, bem como indicarem melhorias a serem adotadas.

Assim, recomenda-se fornecer um ou mais mecanismos de feedback, incluindo formulário de contato, botões para classificar a qualidade dos dados, caixa de comentários, enquetes, entre outros.

A possibilidade de compartilhar os feedbacks também é relevante para dinamizar as interações entre os usuários e os produtores de dados ou, até mesmo, entre os próprios usuários de dados. É uma forma de os produtores demonstrarem aos usuários que suas preocupações estão sendo tratadas, o que, de certa maneira, contribui para o aumento da confiança. Também auxilia os consumidores a entenderem eventuais problemas que podem afetar suas capacidades de utilizar os dados, promovendo um senso de comunidade entre os públicos.

⁶⁰ Data Catalog. Disponível em: <https://catalog.data.gov/dataset/hud-low-and-moderate-income-areas>. Acesso em: 10 set. 2020.

2.2.3 Interação com o publicador de dados

A publicação de dados abertos é destinada à sua divulgação e à utilização por outras pessoas e aplicações. Entretanto, somente publicar conjuntos de dados coloca toda a carga de interpretação de um único lado, ou seja, sobre os consumidores dos dados. A interação ocorre com a existência de comunicação direta entre o consumidor e o provedor da base de dados. Isso permite que os feedbacks gerados, explicados no tópico anterior, cheguem ao conhecimento do detentor da base de dados, para reportarem algum problema, solicitarem alterações e atualizações, dentre outros.

A organização por meio de comunidades potencializa a criação de canais de comunicação mais efetivos e dinâmicos entre o produtor e o usuário. Essa questão é fundamental para facilitar a reutilização das bases de dados disponibilizadas, beneficiando vários grupos e setores da sociedade. Também é importante que existam canais de comunicação entre os próprios usuários de dados, com o objetivo de permitir maior interação entre eles.

Como exemplo, o Portal de Dados Abertos da França⁶¹ possibilita que os usuários entrem em contato com o provedor da base disponibilizada.

2.2.4 Linguagem clara e acessível ao cidadão

A utilização de uma linguagem clara e acessível é fundamental para permitir o entendimento das informações por parte dos interessados. Além disso, trata-se de determinação constante na própria Lei de Acesso à Informação (Lei n.º 12.527/2011), que estabelece que o Estado deve garantir o direito de acesso à informação mediante procedimentos objetivos e ágeis, de forma transparente, clara e em linguagem de fácil compreensão.

O objetivo de se utilizar uma linguagem mais clara e de fácil compreensão é auxiliar os cidadãos no trabalho de localizar as informações desejadas de maneira rápida e objetiva.

⁶¹ Statistiques nationales du commerce extérieur. Disponível em: <https://www.data.gouv.fr/en/datasets/statistiques-nationales-du-commerce-exterieur/>. Acesso em: 10 set. 2020.

Assim, é essencial que conteúdos voltados à transparência pública sejam publicados em formatos e linguagens mais acessíveis aos cidadãos.

O entendimento de informações é outro requisito de qualidade essencial para transparência, além de necessário às bases de uma sociedade engajada, participativa e democrática. Os órgãos governamentais determinam o uso de tecnologias da informação, como a Internet, para que a transparência – ativa e passiva – seja habilitada e garantida às partes interessadas, como o cidadão, o maior alcance possível dos dados públicos e dos dados abertos.

Por exemplo, destacam-se os conteúdos textuais e audiovisuais explicativos – disponibilizados de forma direta e objetiva – sobre o portal ou sobre as bases de dados.

2.2.5 Design e navegação simples e intuitiva

Para proporcionar uma experiência positiva na visita a um portal ou website, um dos aspectos fundamentais é a facilidade com que se percebe sua estrutura e sua navegação. Esses aspectos são determinantes para que qualquer usuário possa encontrar, de um modo fácil e rápido, a informação ou a base de dados desejada. A estrutura e a arquitetura da informação disponível e a solução gráfica utilizada facilitam a usabilidade do usuário e permitem uma navegação intuitiva.

2.2.6 Visualização dos dados

A visualização de dados consiste na representação gráfica de informações e dados. O objetivo é permitir que os cidadãos possam compreender certos conjuntos de dados por meio de elementos visuais, como diagramas, gráficos, infográficos, mapas, entre outras formas de visualização exclusivamente não verbais. Assim, a visualização de dados é uma forma acessível de ver e entender tendências e padrões nos dados.

Utilizando técnicas de visualização de informação, por meio de ferramentas específicas, é possível converter dados numéricos e alfanuméricos e apresentá-los visualmente ao usuário por meio de imagens ou infografias, de forma que possam ser melhor compreendidos.

No portal de dados abertos dos Estados Unidos⁶² são disponibilizados milhares de conjuntos de dados brutos, assim como no portal de dados abertos da Inglaterra⁶³. Os dados são distribuídos de acordo com o órgão governamental correspondente, e ambos apresentam um eficiente mecanismo de busca. Ambos disponibilizam centenas de aplicativos, elaborados por desenvolvedores, que permitem cruzar as informações disponíveis. Isso possibilita, por exemplo, acesso a mapas de ruas com estatísticas e, até mesmo, ao horário dos transportes públicos que circulam em Londres, além da visualização instantânea do próximo ônibus, de acordo com a linha ou a parada em que uma pessoa se encontra.

2.2.7 Seleções, filtros ou extração de dados

Esse recurso permite que os usuários encontrem os dados que necessitam com maior facilidade, mediante ferramentas de busca (filtros, publicador, palavras-chave, nome da base, sigla da base, tags, UF, tópicos). A disponibilização desses mecanismos de busca garante maior facilidade e rapidez na busca por conjuntos de dados disponíveis. Além da ferramenta de busca, é recomendável viabilizar a consulta de dados por criação, atualização, cobertura, frequência e origem do conjunto de dados, consulta da quantidade de conjuntos de dados disponíveis etc.

Uma prática interessante para tornar os dados abertos mais fáceis de serem encontrados é criar catálogos de dados. Eles funcionam como um espaço de acesso a todos os conjuntos de dados disponibilizados por um governo e por seus departamentos, e ainda servem para convidar a sociedade a reutilizar as informações. É por meio desses catálogos, como o norte-americano ou o inglês, que governos mostram os formatos em que os conjuntos de dados estão disponíveis, fornecem links para feeds ou APIs e permitem que os usuários indiquem os conjuntos de dados mais relevantes para as comunidades, em um processo de gestão colaborativa da informação pública.

⁶² Disponível em: <https://www.data.gov/>. Acesso em: 10 set. 2020.

⁶³ Disponível em: <https://data.gov.uk/>. Acesso em: 10 set. 2020.

2.2.8 Mecanismos de mapeamento da demanda por dados abertos

Com o objetivo de adequar a oferta de dados com a demanda dos cidadãos por eles, é recomendável utilizar mecanismos de participação social para priorizar as bases de dados que serão abertas. O objetivo principal da disponibilização de dados abertos é que sejam utilizados, reutilizados e compartilhados pela sociedade. Assim, sua efetividade está associada à utilização dos dados para geração de valor, razão pela qual a participação social é fundamental.

Na priorização de bases a serem abertas, é relevante considerar a adoção de metodologias para identificar bases estruturantes; coletar contribuições para medir o interesse da sociedade; analisar os pedidos de acesso à informação ou manifestações de ouvidoria que solicitam a abertura de bases; estabelecer critérios mínimos para definir a prioridade das bases (utilidade, relevância); realizar consultas públicas específicas para aferir o interesse dos usuários nas bases que integram o inventário do publicador etc.

A aferição do interesse da população quanto aos dados que gostaria que fossem disponibilizados é fundamental para otimizar o trabalho do detentor da base de dados, pois permite que ele empreenda esforços para publicação daqueles dados que têm maior potencial de serem reutilizados. Dessa forma, aumenta-se a perspectiva de que a abertura de dados públicos gere transformações efetivas na vida das pessoas.

2.2.9 Tutoriais e manuais para utilização adequada dos dados

Mostra-se importante disponibilizar materiais didáticos para informar sobre formatos abertos e processos de publicação, definição do padrão de metadados, possibilidades de estruturação de catálogos de dados abertos setoriais, entre outras informações.

3. Reutilização de dados

É importante distinguir entre o uso e o reuso. O uso acontece quando um indivíduo ou uma organização extrai dados, respostas ou, até mesmo, conhecimento a partir de perguntas específicas – quer seja para um projeto de pesquisa específico, quer seja para ações administrativas de governo. Quando esse conjunto de dados é disponibilizado em um repositório, de forma que possa ser recuperado por outra pessoa ou organização para ser utilizado em outro projeto ou iniciativa, ele é considerado reuso. Em síntese, reuso geralmente implica o uso de um conjunto de dados por alguém ou organização que não seja quem originou os dados⁶⁴.

A capacidade de reutilizar dados, quer sejam eles oriundos de pesquisa científica ou de governo, é reconhecidamente um benefício fundamental para comunidades ou para a sociedade como um todo. Isso é mais evidente quando é preciso resolver problemas complexos que exigem colaboração entre diferentes atores – quer seja na busca por respostas a problemas globais, como, por exemplo, no caso de pandemias ou das mudanças climáticas, quer seja na busca por problemas locais ou, até mesmo, no desenvolvimento e na oferta de serviços inovadores para cidadãos.

É no reuso dos dados abertos que se abrem oportunidades inovadoras – não apenas para verificação dos dados ou das informações por eles trazidas, mas também para economia de recursos, para tempo, para produção de novos conhecimentos, para fortalecimento de vínculos entre fornecedores dos dados e usuários, para cruzamento de dados, para transparência e responsabilidade (accountability) e controle social das ações governamentais. O reuso de dados traz oportunidades, sobretudo, para o compartilhamento e o fortalecimento de ações colaborativas como soluções de problemas globais⁶⁵.

Há diversas formas de promover o uso de dados abertos. A Seção 2.2 “Usuário” deste documento lista itens importantes a serem abordados. Além disso, o *Guia de incentivo ao*

⁶⁴ Duas referências importantes, fundamentais e recentes são Pasquetto, Randles e Borgman (2017) e van de Sandt et al. (2019).

⁶⁵ Ver, por exemplo, Tenopir et al. (2015).

(re)uso de dados abertos⁶⁶, do NIC.br, traz exemplos de como os empreendedores de dados abertos governamentais podem ser motivados por benefícios econômicos diretos ou indiretos ou, ainda, atuar sem visar ao lucro direto com essa atividade. Outra fonte importante sobre reuso de dados é o estudo realizado por Caggemini Invent para o Portal Europeu de dados abertos, que resultou no documento “Re-using open data: a study on companies transforming open data into economic & societal value”⁶⁷.

Soluções e aplicativos desenvolvidos podem ser gratuitos aos usuários, gerando renda por meio de publicidade, visibilidade de marca, reputação, subsídios cruzados; ou, ainda, podem ser viabilizados por meio de recursos públicos ou de filantropia, assim como por meio de modelos comunitários de participação voluntária.

3.1 Aplicativos e soluções desenvolvidas com o uso de dados abertos

O *Guia de Dados Abertos*⁶⁸ do NIC.br traz alguns exemplos de aplicativos e soluções desenvolvidos com o uso de dados abertos.

3.1.1 Citymapper⁶⁹

Citymapper é um dos bons exemplos ao redor do mundo que mostram o potencial dos dados abertos para fomentar a geração de novos negócios. É um aplicativo gratuito para

⁶⁶ Disponível em: <https://ceweb.br/publicacao/guia-de-incentivo-ao-reuso-de-dados-abertos/>. Acesso em: 10 set. 2020.

⁶⁷ Disponível em: https://www.europeandataportal.eu/sites/default/files/re-using_open_data.pdf. Acesso em: 10 set. 2020

⁶⁸ Disponível em: <https://ceweb.br/publicacao/guia-de-dados-abertos/>. Acesso em: 10 set. 2020.

⁶⁹ Texto retirado do Guia de Dados Abertos do NIC.br. Disponível em: <https://ceweb.br/guias/dados-abertos/capitulo-5/>. Acesso em: 10 set. 2020.

smartphones, criado por um ex-funcionário do Google, em 2011, em Londres, com o objetivo de melhorar a experiência de milhões de pessoas na cidade, todos os dias, que dependem do transporte público.

Os desenvolvedores do Citymapper tiraram vantagem dos dados sobre ônibus, trens e metrô, publicados em tempo real pela autoridade de trânsito londrina, para formular um algoritmo que mostra sempre o melhor tempo de viagem entre um ponto e outro; quando será a próxima viagem; e quantas calorias se gastou caminhando. As informações são atualizadas em tempo real e mostram as condições de trânsito e do clima e problemas técnicos nos veículos.

A receita deu certo, e a empresa expandiu seus serviços para outras cidades. De 2011 a 2015, o aplicativo chegou a sete países e treze cidades: Londres, Nova York, Paris, Berlim, Washington, DC, Madri, Boston, Barcelona, São Francisco, Chicago, Milão, Roma e Cidade do México. O aplicativo pode ser conectado em qualquer cidade que ofereça dados sobre serviços de transporte público em tempo real.

3.1.2 Influence Explorer⁷⁰

Site da Web que rastreia as doações políticas em níveis estadual e federal, permitindo que qualquer pessoa possa acompanhar o grau de influência por legislador, empresa ou indivíduo. A ferramenta oferece uma visão geral sobre financiamentos de campanha, lobby, verbas parlamentares, irregularidades em contratações e gastos públicos federais.

3.1.3 Prescribing Analytics⁷¹

O Prescribing Analytics mostra como agentes do governo e profissionais independentes podem trabalhar juntos para promover formas de economizar dinheiro público. Trata-se

⁷⁰ Disponível em: <http://www.influenceexplorer.com/>. Acesso em: 10 set. 2020.

⁷¹ Disponível em: <http://www.prescribinganalytics.com/>. Acesso em: 10 set. 2020.

de uma ferramenta de análise de dados sobre os gastos do governo britânico com um grupo específico de remédios, a estatina. Esse medicamento ajuda no combate às altas taxas de colesterol em pacientes com problemas de saúde.

O serviço foi criado por um grupo de desenvolvedores independentes e médicos que trabalham no setor público britânico. O sistema de saúde britânico recebeu a missão de economizar 20 bilhões de libras até 2015, sendo que metade disso é gasto com remédios. Uma das substâncias mais prescritas pelos médicos do sistema público de saúde do Reino Unido é justamente a estatina. A ferramenta acessa a base de dados aberta sobre receitas médicas no país e analisa quais tipos de estatina foram prescritas pelos médicos no sistema público: as variantes mais caras ou as “genéricas”, normalmente mais baratas.

Os dados mostram que o governo britânico poderia ter economizado, entre 2011 e 2012, 27 milhões de libras por mês se os médicos tivessem receitado variantes genéricas da estatina, que possuem a mesma eficácia das mais caras.

3.1.4 QEdu⁷²

O QEdu é um dos bons exemplos brasileiros que demonstra os benefícios do trabalho em conjunto do terceiro setor, da academia e do governo para melhor entender a educação básica no Brasil e fornecer subsídios para políticas públicas. Trata-se de um portal gratuito, desenvolvido por uma parceria entre a Fundação Lemann e a Meritt, que permite a qualquer pessoa obter informações sobre a qualidade do aprendizado nas escolas brasileiras, com dados sobre escolas públicas e particulares. As visualizações são geradas a partir de bases do governo brasileiro, como dados do Exame Nacional do Ensino Médio (Enem), da Prova Brasil, do Censo Escolar e de indicadores especiais do Instituto Nacional de Estudos e Pesquisas (INEP).

A ferramenta mostra como foi o aprendizado de alunos dos 5º e 9º anos em matemática e português; perfil dos alunos, professores e diretores das escolas que realizaram a Prova Brasil; matrículas para cada etapa escolar; taxas de aprovação, abandono e reprovação; nível socioeconômico; infraestrutura escolar; distorções de idade em relação às séries; e muito mais. É possível fazer buscas e comparações a partir de escolas, cidades ou estados.

⁷² Disponível em: <http://www.qedu.org.br/>. Acesso em: 10 set. 2020.

O portal também permite que qualquer pessoa possa filtrar e baixar os dados que precisar em formato aberto.

Outros exemplos de aplicativos e soluções brasileiras são apresentados a seguir.

3.1.5 Radar Parlamentar⁷³

O Radar Parlamentar é um aplicativo que ilustra as semelhanças entre partidos políticos com base na análise matemática dos dados de votações que ocorrem na casa legislativa. As semelhanças são apresentadas em um gráfico bidimensional, em que círculos representam partidos ou parlamentares e a distância entre esses círculos representa o quão parecido os mesmos votam.

3.1.6 Appelou⁷⁴

Appelou é um aplicativo baseado nas informações coletadas a partir da nota fiscal do consumidor eletrônica (NFC-e), que são armazenadas e utilizadas no sistema do Economiza Alagoas⁷⁵. O Appelou permite pesquisar preços baratos e favoritar os locais em que o usuário faz compras com mais frequência.

3.1.7 Colaborados⁷⁶

Colaborados é uma iniciativa colaborativa que busca reunir, demonstrar, investigar e monitorar o acesso à informação no Brasil.

⁷³ Disponível em: <http://radarparlamentar.polignu.org/>. Acesso em: 10 set. 2020.

⁷⁴ Disponível em: <https://www.facebook.com/appeloubr>. Acesso em: 10 set. 2020.

⁷⁵ Disponível em: <https://economizaalagoas.sefaz.al.gov.br/>. Acesso em: 10 set. 2020.

⁷⁶ Disponível em: <http://colaborados.com.br/>. Acesso em: 10 set. 2020.

3.1.8 Brasil.io⁷⁷

Brasil.io é um repositório de dados públicos, disponibilizados em formato acessível.

3.1.9 CoronApp⁷⁸

O CoronApp é um aplicativo que utiliza geolocalização para informar casos de contaminação pelo novo coronavírus ao redor do celular em que está instalado. Também é possível obter informações sobre a localização dos hospitais, farmácias, supermercados e mercearias mais próximos. Possui plataforma crowdsourcing, em que usuários podem inserir novas informações a respeito do comércio local e também sobre recursos que eles podem doar. Além disso, o aplicativo fornece informações a respeito do novo coronavírus extraídas do Twitter.

3.1.10 Coronavírus SUS⁷⁹

O aplicativo Coronavírus SUS, criado pelo Ministério da Saúde, tem o objetivo de conscientizar a população com informações sobre o novo coronavírus. Também informa sobre as unidades de saúde próximas à localização do usuário. Na aba “Situação”, podemos obter informações sobre os casos confirmados, óbitos e letalidade do vírus em cada estado do Brasil. Na aba “Notícias”, temos o espelho do perfil do Ministério da Saúde no Twitter.

⁷⁷ Disponível em: <https://brasil.io/home> . Acesso em: 10 set. 2020.

⁷⁸ Disponível em: <https://devpost.com/software/coronapp>. Acesso em: 10 set. 2020.

⁷⁹ Disponível em: https://play.google.com/store/apps/details?id=br.gov.datasus.guardioes&hl=pt_BR. Acesso em: 10 set. 2020.

3.1.11 Coronavírus no mundo, no Brasil e em Pernambuco⁸⁰

A Secretaria de Planejamento de Gestão de Pernambuco (Seplag) criou um mapa virtual que mostra a localização das contaminações pela COVID-19 (respeitando uma margem de 15 metros da localização real, para não ferir a privacidade das pessoas infectadas). O mapa permite algumas modalidades de visualização, como o número de contaminados por rua ou por bairro. Além do mapa, é possível obter outras informações relacionadas ao vírus, como faixa etária, letalidade e número de pacientes recuperados em Pernambuco. O mapa é atualizado diariamente pela Seplag.

3.1.12 Ciência Aberta é Vida – Diretório de fontes de informação científica de livre acesso sobre o coronavírus⁸¹

O Diretório, mantido pelo Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), reúne diversas fontes de informação científica de acesso aberto sobre o novo coronavírus. Além de artigos já publicados, os usuários também têm acesso a trabalhos no prelo.

3.1.13 Información Legal⁸²

Información Legal é uma campanha pública argentina que visa disseminar informação sobre a COVID-19 para grupos vulneráveis, como crianças, pessoas com deficiência e moradores de favelas.

⁸⁰ Disponível em: https://dados.seplag.pe.gov.br/apps/corona.html?fbclid=IwAR3pbSiShX9DzO1K7T33hnbjdzl-QKr9L0sqpduni12MUjp_Qc9oSZ-poc4#mapas. Acesso em: 28 ago. 2020.

⁸¹ Disponível em: <http://diretoriodefontes.ibict.br/coronavirus/>. Acesso em: 28 ago. 2020.

⁸² Disponível em: <https://acij.org.ar/covid19yderechos/>. Acesso em: 28 ago. 2020.

3.1.14 Observatório de Oncologia⁸³

Segundo informações extraídas do portal, o Observatório de Oncologia é uma plataforma de monitoramento de dados abertos e compartilhamento de informações relevantes da área de oncologia no Brasil. As principais bases de dados utilizadas são as do Ministério da Saúde (DataSUS), do Instituto Nacional do Câncer (INCA) e do Instituto Brasileiro de Geografia e Estatística (IBGE). Os dados abrangem quatro dimensões: demográfica, epidemiológica, procedimentos de assistência à saúde e estrutura da rede assistencial.

3.1.15 Serenata do Amor⁸⁴

A iniciativa Operação Serenata de Amor utiliza tecnologias baseadas em inteligência artificial para ofertar serviços de fiscalização dos gastos públicos e compartilhar informações com a sociedade. Em especial, o site faz análises dos gastos reembolsados pela Cota para Exercício da Atividade Parlamentar (CEAP)⁸⁵ de deputados federais e senadores, feitos em exercício de sua função, identificando suspeitas e incentivando a população a questioná-las.

3.2 Boas práticas de utilização de dados abertos

A recomendação do Boas Práticas para Dados na Web (DWBP) traz sugestões sobre o reúso dos dados abertos.

⁸³ Disponível em: <https://observatoriodeoncologia.com.br/#sobre>. Acesso em: 28 ago. 2020.

⁸⁴ Disponível em: <https://serenata.ai/>. Acesso em: 28 ago. 2020.

⁸⁵ Disponível em: <https://www.camara.leg.br/cota-parlamentar/>. Acesso em: 28 ago. 2020.

3.2.1 Enriquecimento de dados⁸⁶

Enriquecimento de dados se refere a um conjunto de processos que pode ser utilizado para aperfeiçoar, aprimorar ou melhorar de outra forma dados brutos ou dados previamente processados.

O enriquecimento pode acentuar consideravelmente a facilidade de processamento, especialmente no caso de dados não estruturados. Em algumas circunstâncias, dados faltantes podem ser adicionados, e novas atribuições e mensurações podem ser acrescentadas a partir de dados brutos pré-existentes. Conjuntos de dados também podem ser enriquecidos por meio de coletas adicionais, da mesma forma como os dados originais foram coletados ou combinando dados originais com outros conjuntos de dados. Publicar conjuntos de dados mais completos pode aumentar a confiança se feito de maneira adequada e ética. Derivar valores adicionais que sejam úteis economiza tempo para o usuário e incentiva mais tipos de reusos. Há muitas técnicas inteligentes que podem ser utilizadas para enriquecer dados, tornando o conjunto de dados um bem ainda mais valioso.

Dados publicados online se destinam a passar informações para terceiros sobre seu tema. Porém, apenas publicar conjuntos de dados para download ou através de API delega ao consumidor o esforço de interpretá-los. A Web oferece oportunidades sem paralelos para a apresentação de dados, de forma a permitir que os usuários aprendam e explorem, sem ter que criar suas próprias ferramentas.

As Boas Práticas 31 e 32 do documento DWBP sugerem enriquecer dados, primeiro, por meio da geração de novos dados, pois, ao fazê-lo, aumentará o valor dos mesmos. Assim, conjuntos de dados com dados faltantes serão melhorados por meio do preenchimento desses valores. A estrutura do conjunto de dados será verificada, e a utilidade será potencializada, caso providências ou atributos relevantes sejam adicionados; porém, somente se esse incremento não distorcer os resultados analíticos, a significância ou a potência estatística. Segundo, ao fornecer visualizações complementares, apresentando os dados em formas complementares e diretamente informativas, tais como visualizações, tabelas, aplicações

⁸⁶ Textos reproduzidos da seção sobre enriquecimento de dados da *Recomendação de Boas Práticas para Dados na Web* (DWBP). Disponível em: <https://w3c.br/traducoes/DWBP-pt-br/#enrichment>. Acesso em: 28 ago. 2020.

Web ou resumos. Visualizações complementares dos dados permitirão que consumidores tenham uma percepção imediata dos mesmos por meio de visualizações de fácil entendimento.

3.2.2 Republicação de dados⁸⁷

Reutilizar dados é outra forma de publicar dados; é simplesmente republicá-los. Pode tomar a forma de uma combinação de dados existentes com outros conjuntos de dados; de uma criação de aplicações ou visualizações Web; ou de um reempacotamento de dados em um novo formato, como uma tradução.

As Boas Práticas 33, 34 e 35 do documento DWBP explicam sobre os requisitos para republicar dados na Web.

É importante fornecer feedback para o publicador original. É importante informá-lo, caso encontre um erro ou tenha sugestões ou elogios a tecer. Os publicadores geralmente querem saber se os dados que publicam têm sido úteis.

Também é necessário obedecer os termos de licença dos dados a serem utilizados. É necessário encontrar e seguir os requisitos de licença indicados pelo publicador original do conjunto de dados.

Finalmente, é importante citar a publicação original do conjunto de dados, ou seja, reconhecer a fonte de seus dados nos metadados. Caso forneça uma interface de usuário, deve se incluir a citação na interface de maneira clara.

⁸⁷ Disponível em: <https://w3c.br/traducoes/DWBP-pt-br/#Reuse>. Acesso em 28 ago. 2020.

4. Diretrizes de abertura para dados públicos estaduais e municipais

Tendo em vista o objetivo deste documento de apresentar um modelo de referência para abertura de dados que seja replicável em todas as esferas administrativas, esta seção aborda especificamente sugestões de temas para orientar o processo de abertura de dados públicos nas esferas estaduais e municipais. A ideia é apresentar diretrizes que possam auxiliar estados e municípios a disponibilizar e a selecionar, de forma aberta, um conteúdo mínimo de informações e temas de sua competência, com base em pesquisas realizadas com especialistas e em boas experiências nacionais.

Para estabelecer esse conteúdo mínimo, utilizaram-se duas abordagens distintas. A primeira, a ser detalhada a seguir, buscou ouvir especialistas em diversas áreas para levantar demandas de cada setor com relação à abertura de dados nos níveis estadual e municipal. A segunda abordagem levou em consideração agendas mais amplas de discussão na arena pública, que demandam uma série de indicadores de estados e municípios, a saber: dados associados a dimensões de competência específica de cada ente subnacional; dados associados aos 17 Objetivos do Desenvolvimento Sustentável (ODS), das Nações Unidas; dados que possam subsidiar soluções para Cidades Inteligentes (Smart Cities); dados que permitam o aprimoramento de serviços públicos; e dados que permitam o fomento à inovação e aos negócios. Evidentemente, diversas informações atendem a mais de um critério.

A seção está subdividida em uma apresentação rápida sobre a metodologia empregada na primeira abordagem e, em seguida, nos resultados por tema, que, por sua vez, estão subdivididos em dados considerados para o modelo de referência dentro de cada tema. O resultado é baseado nas duas abordagens mencionadas acima.

4.1 Metodologia da pesquisa com especialistas

Com relação à primeira abordagem, a ideia é que, para apresentar um modelo que corresponda a um conteúdo mínimo necessário de informações de competência estadual e

municipal que precisam estar abertas, é importante mapear quais informações são mais essenciais ou mais prioritárias para abertura em cada área de expertise para a qual o governo produz dados.

Para isso, foi conduzida uma pesquisa – via websurvey, usando a plataforma *Google Forms* – com especialistas em dados abertos no Brasil para levantar informações sobre o conteúdo mínimo de temas necessários para abertura de dados em cada área para estados e municípios. Os participantes da execução do Compromisso 2 elaboraram, em conjunto, uma lista de pontos focais para o primeiro disparo do questionário, e, em seguida, essa lista foi retroalimentada com contatos fornecidos pelos próprios respondentes. Ou seja, foi utilizada uma metodologia do tipo snowball (bola de neve) para buscar especialistas em cada área. Vale destacar que o questionário foi também divulgado em grupos públicos de dados abertos, como o grupo do Telegram Dados Abertos BR⁸⁸, a Comunidade Dados Abertos BR⁸⁹ e o grupo público Dados Abertos no Facebook⁹⁰.

A pesquisa ficou aberta ao longo de um período de 24 dias, desde 13 de fevereiro até 08 de março de 2020. O questionário foi enviado para um total de 186 contatos diretos (144 na primeira leva e mais 42 através da bola de neve) e recebeu 60 respostas diferentes de 57 respondentes (três responderam para duas áreas diferentes). Os entrevistados escolheram uma dentre as 13 áreas disponíveis ou optaram por sugerir uma área diferente dessas 13 opções. A partir dessa escolha de área, o questionário basicamente solicitou uma listagem de dados considerados essenciais ou prioritários para abertura em níveis estadual e municipal, separadamente. Ao final, abriu-se um campo para que o entrevistado sugerisse contatos de outras pessoas para participarem da pesquisa. Mais detalhes podem ser consultados no **Anexo**.

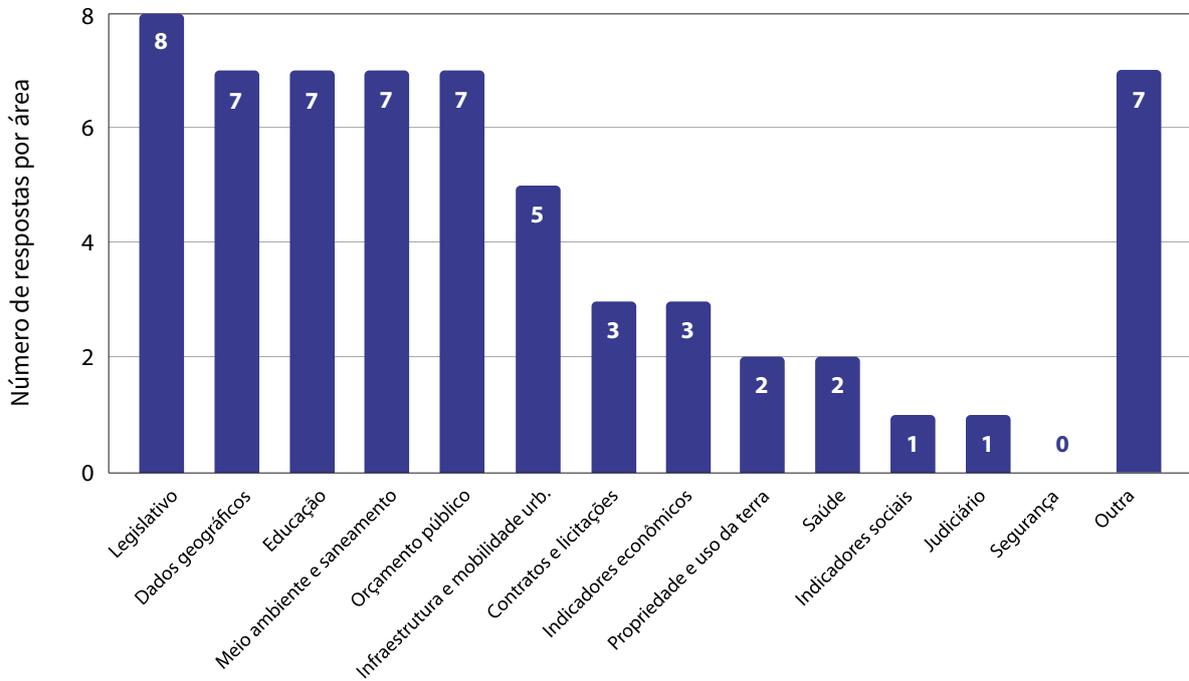
Foi possível obter pelo menos um respondente de cada área mapeada, como mostra a **Figura 1**. A única área que não recebeu qualquer resposta foi a de segurança. No entanto, a frequência foi maior para a área de dados legislativos (em primeiro lugar, com oito respostas), em seguida, para dados geográficos, de educação, de meio ambiente e saneamento e de orçamento público, cada uma com sete respostas.

⁸⁸ Disponível em: <https://t.me/dadosabertos/>. Acesso em: 11 set. 2020.

⁸⁹ Disponível em: <https://dadosabertos.social/>. Acesso em: 28 ago. 2020.

⁹⁰ Disponível em: <https://www.facebook.com/groups/dadosabertos/>. Acesso em: 28 ago. 2020.

Figura 1 – Número de respostas por área



Elaboração: FGV DAPP.

Entre os que escolheram uma área diferente das 13 inicialmente mapeadas, observa-se uma concentração em dois grupos não necessariamente excludentes: informações científicas⁹¹ e dados sobre uso de tecnologia⁹². A seguir, serão apresentados os resultados por área.

4.2 Dados abertos essenciais por área

4.2.1 Legislativo

Na área de dados legislativos, as opiniões dos especialistas foram muito semelhantes

⁹¹ As áreas sugeridas nesse sentido foram “dados científicos”, “dados de ciência e tecnologia” e “informação científica”

⁹² As áreas sugeridas nesse sentido foram “pesquisa e desenvolvimento em inteligência artificial”, “tecnologia” e “uso de TIC em governo”.

entre dados essenciais nas esferas estadual e municipal. A análise das respostas permite observar a importância da divulgação das seguintes informações:

- Dados de atividade legislativa, como pauta do dia, transcrições, atas, resultados de votações e presença dos parlamentares;
- Dados sobre execução orçamentária associada ao funcionamento do legislativo e das verbas parlamentares, com especial destaque para licitações, despesas sem licitação, gastos com pessoal, gastos com publicidade e comunicação e verbas indenizatórias;
- Agenda pública do legislativo e dos parlamentares;
- Folha de pagamento e relação de servidores do legislativo, com destaque para comissionados;
- Dados de todo o processo legislativo, proposições e leis sancionadas;
- Dados do perfil dos parlamentares;
- Relação de comissões e bancadas com respectivos membros; e
- Canais de comunicação com a população.

Além desses indicadores, vale destacar a importância da representatividade de gênero no parlamento, algo que é colocado explicitamente no ODS 5 (Igualdade de Gênero). Sendo assim, um dado importante seria a proporção de assentos na Câmara de Vereadores ou Assembleia Legislativa ocupados por mulheres.

4.2.2 Dados geográficos

No que se refere aos dados geográficos, também foi possível observar, em geral, uma semelhança entre os dados considerados prioritários nas esferas estadual e municipal. No entanto, há algumas particularidades nesse caso. Em relação aos dados relevantes para ambas as dimensões, destacam-se:

- Geolocalização de estabelecimentos e equipamentos públicos em áreas como ensino (p. ex., escolas), saúde (p. ex., postos de saúde) e infraestrutura (p. ex., obras), além de sedes de distritos/subprefeituras, estabelecimentos de turismo etc;
- Dados relativos ao sistema de transporte, rodovias, ferrovias etc.;
- Dados sobre hidrografia;
- Bases cartográficas sistemáticas (em escalas grandes, como 25 mil ou 50 mil, no caso de estados, e 1 mil ou 5 mil, no caso de municípios);
- Limites político-administrativos;
- Áreas ambientalmente protegidas;
- Limites de propriedades públicas;
- Imagens de satélite, aerolevantamentos convencionais ou com Aeronaves Remotamente Pilotadas (ARP).

Especificamente na dimensão estadual, destacam-se ainda:

- Bases de dados de levantamentos cartográficos contratados por estados ou Distrito Federal (DF);
- Dados de Zoneamento Ecológico Econômico atualizado, com todas as camadas disponíveis.

E com relação aos municípios, destacam-se:

- Base cadastral de logradouros;
- Mapas de uso e cobertura do solo, incluindo seu histórico;
- Plano diretor.

4.2.3 Educação

Quanto aos dados de educação, há bastante sobreposição entre dados considerados essenciais pelos especialistas nas esferas estaduais e municipais. As respostas ressaltam a

necessidade de divulgar as seguintes informações:

- Dados sobre o local de moradia dos alunos por escola, sem permitir a identificação individual (p. ex., bairro ou setor censitário);
- Dados de participação de alunos em programas de transferência de renda;
- Números de alunos aprovados, reprovados, transferidos, além de dados de evasão escolar por unidade;
- Dados detalhados das matrículas ativas;
- Dados sobre destino dos egressos;
- Dados de orçamento destinado à educação e a processos administrativos associados à educação, incluindo notas fiscais detalhadas de compras e contratações;
- Dados sobre formação de professores;
- Dados sobre acessibilidade;
- Listagem de escolas administradas pelo estado (ou município) e conveniadas, incluindo endereço, site, horário de funcionamento, quantidade de salas, turmas por série, professores e alunos por turma, calendário escolar e indicadores de desempenho.

Indicadores específicos na área de educação podem ser úteis para o desenho de políticas em nível local, tal como sugerido pelo ODS 4 (Educação de Qualidade). Enumeram-se, a seguir, alguns desses indicadores que poderiam ser abertos nos níveis estadual e municipal:

- Proporção de crianças e jovens em cada nível de ensino e em cada grupo etário que atingiram um nível mínimo de proficiência em leitura e matemática;
- Taxa de participação de jovens e adultos na educação formal e não formal;
- Proporção de escolas com acesso a eletricidade, internet para fins pedagógicos, computadores para fins pedagógicos, infraestrutura e materiais adaptados para alunos com deficiência, água potável, instalações sanitárias separadas por sexo e instalações básicas para lavagem das mãos;
- Proporção de professores em cada nível de ensino que receberam pelo menos a formação mínima (p. ex., formação pedagógica), antes ou durante o exercício da profissão, requerida para lecionar nesse nível.

Outros indicadores não contemplados acima, mas que também surgem em outras diretrizes específicas de cidades (tal como o Programa Cidades Sustentáveis), são:

- Escolas da rede municipal que abrem aos finais de semana para a comunidade;
- Alunos com acesso à quadra esportiva (coberta e descoberta);
- Alunos com deficiência na rede regular de ensino;
- Analfabetismo funcional da população;
- Demanda atendida de vagas em todos os níveis de ensino;
- Proporção de alunos matriculados em educação em tempo integral em todos os níveis;
- Registros de escolas com dependências adequadas a pessoas com deficiência.

4.2.4 Meio ambiente e saneamento

Na dimensão de meio ambiente e saneamento, boa parte dos dados considerados essenciais é similar, independentemente do ente federativo e a depender de quem é o gestor dessa área localmente. Além disso, há dados específicos da dimensão estadual. No que se refere aos dados aplicáveis independentemente da entidade federativa, destacam-se:

- Dados de acesso aos serviços de saneamento oferecidos direta ou indiretamente pelo estado ou município, incluindo a infraestrutura existente, com destaque para gestão de resíduos sólidos;
- Dados orçamentários, financeiros e contratuais na área de saneamento;
- Dados sobre sistemas coletores e adutores, estações de tratamento de água e esgoto e pontos de lançamento de efluentes da indústria;
- Dados geográficos sobre aspectos físicos (geologia, geomorfologia, clima, hidrografia etc.) do estado ou município;
- Dados de outorgas de uso de recursos hídricos e dados relativos a cobrança pelo uso da água, contratos de programa (assinados com as companhias estaduais de saneamento básico) e contratos de concessão de prestação dos serviços de saneamento;

- Dados de resultados das análises de qualidade de água (bruta e potável) e balneabilidade;
- Dados resultantes do monitoramento da qualidade dos serviços e das metas contratuais.

Especificamente na dimensão estadual, destacam-se, ainda:

- Dados de autorizações de desmatamento, licenças ambientais, autos de infração, embargos, emissão de títulos (fundiários e áreas públicas) e de cadastro ambiental rural, entre outras licenças, permissões, cadastros e autorizações ambientais de competência estadual;
- Dados sobre áreas de preservação ambiental.

Os ODS podem oferecer um bom guia para a disponibilização de indicadores de meio ambiente e de saneamento nos níveis estadual e municipal. Por exemplo, no ODS 6 (Água Potável e Saneamento), constam os indicadores:

- Proporção da população que utiliza serviços de água potável e serviços de saneamento gerenciados de forma segura;
- Proporção de corpos hídricos com boa qualidade ambiental;
- Nível de estresse hídrico: proporção das retiradas de água doce em relação ao total dos recursos de água doce disponíveis.

Há também outros indicadores úteis nos ODS 7 (Energia Acessível e Limpa), 11 (Cidades e Comunidades Sustentáveis) e 15 (Vida Terrestre), como, por exemplo:

- Percentagem da população com acesso a eletricidade;
- Participação das energias renováveis na oferta de energia;
- Proporção de resíduos sólidos urbanos regularmente coletados e com destino final adequado em relação ao total de resíduos sólidos urbanos gerados;
- Área florestal como proporção da área total do território.

Outros indicadores não contemplados acima, mas que também surgem em outras diretrizes específicas de cidades (tal como o Programa Cidades Sustentáveis), são:

- Área desmatada e restaurada de florestas degradadas;
- Áreas de mananciais protegidas e destinadas às Unidades de Conservação;
- Cobertura vegetal nativa remanescente;
- Indicador de perda de água tratada;
- Registros de queimadas e incêndios florestais;
- Área de zona costeira em área de conservação e interesse científico;
- Consumo de água per capita;
- Consumo de eletricidade per capita;
- Consumo municipal anual de agrotóxicos e afins por área plantada;
- Presença de reciclagem e compostagem na destinação de resíduos;
- Domicílios em áreas de risco;
- Temperatura média mensal.

4.2.5 Orçamento público

No caso dos dados de orçamento público, os especialistas apresentaram conjuntos de dados praticamente idênticos como prioritários para estados ou municípios, dentre os quais, destacam-se:

- Dados de execução orçamentária que permitem o acompanhamento completo do processo orçamentário (autorizado, empenhado, liquidado, pago, restos a pagar) para todas as despesas financeiras e não financeiras no nível institucional das unidades gestoras (e subníveis mais agregados);
- Dados de execução orçamentária de programas e ações, bem como acompanhamento dos projetos prioritários do Plano Plurianual (PPA) – descrição, valores, metas e indicadores;
- Dados de benefícios tributários e fiscais;

- Dados detalhados de receita prevista e arrecadada;
- Estrutura de pessoal (quantidade de servidores, remunerações, alíquotas de contribuições, despesas totais) de todas as esferas (Judiciário, Executivo, Legislativo, Ministério Público e Defensoria);
- Dados comparativos entre despesas setoriais (segurança, saneamento etc.) e indicadores de qualidade do gasto público;
- Dados de contratos, convênios e parcerias público-privadas.

Apesar de a lista acima ser bastante abrangente para o universo do orçamento público estadual e municipal, alguns indicadores específicos de monitoramento dos ODS (em especial, o ODS 1, de Erradicação da Pobreza) explicitam necessidades de divulgação de dados que revelem os esforços governamentais em áreas sociais, tais como:

- Proporção de recursos gerados domesticamente e alocados pelo governo diretamente a programas de redução de pobreza;
- Proporção do total das despesas públicas com serviços essenciais (educação, saúde e proteção social);
- Soma das subvenções totais e das entradas que não geram dívidas diretamente alocadas a programas de redução da pobreza, como proporção do Produto Interno Bruto (PIB).

Além disso, o ODS 17 (Parcerias e Meios de Implementação) especifica alguns indicadores úteis relacionados à capacidade fiscal de estados e municípios, a saber:

- Total das receitas do governo em percentagem do PIB por fonte;
- Percentagem do orçamento do governo financiada por impostos cobrados internamente.

4.2.6 Infraestrutura e mobilidade urbana

Na dimensão de infraestrutura e mobilidade urbana, parte significativa dos dados considerados essenciais é similar, independentemente do ente federativo e a depender de quem é

o gestor dessa área localmente. Além disso, há dados específicos das dimensões estadual e municipal. No que se refere aos dados aplicáveis independentemente da entidade federativa, destacam-se:

- Dados detalhados de projetos para infraestrutura de mobilidade urbana (pistas, viadutos, pontes, ciclovias, calçadas), em particular, dados detalhados e espacializados dos investimentos previstos pelo Plano Plurianual para transporte e mobilidade;
- Dados relativos aos padrões de demanda dos diversos modais de transporte, motorizados e não motorizados;
- Dados operacionais sobre transporte intermunicipal de passageiros;
- Malhas viárias (com as respectivas velocidades permitidas), cicloviárias, ferroviárias em Shapefile;
- Linhas de transporte público (ônibus, metrô, veículo leve sobre trilhos [VLT] etc.) e os respectivos percursos em Shapefile;
- Dados da operação das linhas de transporte público em tempo real, informando a quantidade de veículos em operação para cada linha e em que momento cada veículo entrou ou saiu de operação;
- Quantitativos de passageiros para cada linha de transporte público;
- Local (via sistema de posicionamento global [GPS]) e hora de registro da passagem (catraca/cartão eletrônico), para auxiliar pesquisas de demanda, independentemente do modal de transporte público;
- Registros de ocorrências de trânsito, detalhando os tipos de veículos envolvidos (carros, motos, bicicletas etc.) e as vítimas;
- Registros de infrações e crimes de trânsito;
- Dados de consumo de energia por transporte;
- Dados de emissões atmosféricas por transporte.

Especificamente na dimensão estadual, destacam-se, ainda:

- Dados dos repasses estaduais para os municípios referentes a impostos relacionados a transporte, como o Imposto sobre Propriedade de Veículos Automotores (IPVA) e multas.

E, com relação aos municípios, se destacam:

- Dados dos municípios referentes a receitas oriundas de multas de trânsito.

4.2.7 Contratos e licitações

No caso dos dados sobre contratos e licitações, os especialistas convergiram para um mesmo conjunto de dados prioritários para estados ou municípios, dentre os quais, destacam-se:

- Dados para acompanhamento das licitações: preço de referência, datas (início da fase interna, abertura de edital, homologação etc.), valores de referência, íntegra da documentação das licitações, objeto da licitação (incluindo descrição dos itens), dados sobre os licitantes, modalidade, órgão vinculado, dotação orçamentária;
- Dados detalhados de acompanhamento da execução contratual e das compras governamentais: órgão, modalidade, identificador do contrato, íntegra dos contratos, objeto, favorecido, geolocalização, valor do contrato (incluindo ajustes, aditivos etc.), vinculação com a dotação orçamentária, datas importantes (publicação, assinatura, vigência etc.);
- Vinculação dos empenhos e pagamentos (execução orçamentária) com licitações/contratos no nível de item ou de serviço adquirido.

4.2.8 Indicadores econômicos

Em relação aos indicadores econômicos, os respondentes também sugeriram conjuntos prioritários de dados idênticos para as esferas estadual e municipal. No entanto, parte significativa das respostas se enquadra em outras categorias, como orçamento público, saúde, segurança e educação. O aspecto a se ressaltar, dentre as respostas específicas sobre indicadores econômicos, é a necessidade de disponibilizar séries temporais diversas, como arrecadação de cada imposto (estadual ou municipal), separando a arrecadação direta de multas e juros de dívida ativa, além da sua recuperação, e dados de desonerações fiscais abertas por tipo (crédito presumido, isenções, redução de base, Sistema Integrado de Pagamento de Impostos e Contribuições das Microempresas e Empresas de Pequeno Porte [SIMPLES]) e por produto/setor/município.

Nessa dimensão, a análise dos indicadores de cumprimento dos ODS pode ser muito útil, em especial, dos ODS 2 (Fome Zero e Agricultura Sustentável), 8 (Trabalho Decente e Crescimento Econômico) e 9 (Indústria, Inovação e Infraestrutura). Elencamos, a seguir, alguns desses indicadores:

- Renda média dos pequenos produtores de alimentos;
- Proporção da área agrícola sob agricultura produtiva e sustentável;
- Proporção de pessoas que possuem telefone celular móvel;
- Taxa de crescimento real do PIB per capita;
- Taxa de variação anual do PIB real por pessoa ocupada (produtividade);
- Proporção de trabalhadores ocupados em atividades não agrícolas informais;
- Salário médio por hora de empregados;
- Taxa de desocupação;
- Percentagem de jovens (15-24) que não estão na força de trabalho (ocupados e não ocupados), não são estudantes e nem estão em treinamento para o trabalho (nem-nem);
- Proporção e número de crianças de 5-17 anos envolvidos em trabalho infantil;
- Turismo em percentagem do PIB e por taxa de variação;
- Número de agências bancárias e de postos de multibanco (ATM) por 100 mil adultos;
- Proporção de adultos (15 anos ou mais) com uma conta em um banco ou em outra instituição financeira ou com um serviço móvel de dinheiro;
- Valor adicionado da indústria em proporção do PIB e per capita;
- Emprego na indústria em proporção do emprego total;
- Proporção do valor adicionado nas indústrias de média e alta intensidade tecnológica em relação ao valor adicionado total;
- Proporção da população coberta por rede móvel;
- Proporção de indivíduos que utilizam a internet.

Outros indicadores não contemplados acima, mas que também surgem em outras diretrizes específicas de cidades (tal como o Programa Cidades Sustentáveis), são:

- Acesso ao crédito para pequenos empresários;
- Aprendizes contratados no município;
- Oferta de capacitação profissional e incentivo ao primeiro emprego;
- Lista suja do trabalho análogo à escravidão.

4.2.9 Propriedade e uso da terra

Sobre propriedade e uso da terra, os dados considerados prioritários, independentemente da esfera (estadual e municipal), pelos respondentes foram:

- Coordenadas geográficas das propriedades;
- Tipo de cobertura vegetal utilizada (p. ex., no caso de agricultura, o tipo de plantação, no caso de pecuária, o tipo de capim etc.);
- Planejamento futuro para a utilização da terra.

Especificamente na dimensão estadual, destacam-se:

- Guias de Trânsito Animal (GTA);
- Cadastro Ambiental Rural (incluindo dados do proprietário);
- Dados de autorizações e licenças (especialmente a Autorização de Desmatamento ou Autorização de Supressão Vegetal);
- Autuações e embargos emitidos pela Secretaria de Meio Ambiente.

Por fim, na dimensão municipal, foi destacada a importância de divulgar a tabela de referência de impostos sobre a propriedade, como o IPTU e o Imposto Territorial Rural (ITR). Adiciona-se, ainda, um indicador previsto no ODS 11 (Cidades e Comunidades Sustentáveis),

que é a proporção de população urbana vivendo em assentamentos precários, assentamentos informais ou domicílios inadequados.

Outro indicador não contemplado acima, mas que também surge em outras diretrizes específicas de cidades (tal como o Programa Cidades Sustentáveis), é:

- Registro de demarcação de terras indígenas e quilombolas.

4.2.10 Saúde

No caso dos dados de saúde, tanto para estados quanto para municípios, os especialistas chamaram a atenção para dados no nível macro (como mortalidade infantil, expectativa de vida etc.) e no nível mais micro (nascimentos, óbitos, casos de doenças infecciosas, e fila de atendimento para consultas e exames por especialidade). Chamou-se a atenção para a importância de se terem dados no nível submunicipal (bairros, distritos etc.).

Nessa dimensão, a análise dos indicadores de cumprimento dos ODS também pode ser muito útil, em especial, do ODS 3 (Saúde e Bem-Estar). Elencamos, a seguir, alguns desses indicadores:

- Mortalidade materna;
- Mortalidade neonatal e em menores de cinco anos;
- Incidência de tuberculose, malária, hepatite e HIV/AIDS;
- Taxa de mortalidade por doenças do aparelho circulatório, tumores malignos, diabetes mellitus e doenças crônicas respiratórias;
- Taxa de mortalidade por suicídio;
- Cobertura das intervenções (farmacológicas, psicossociais, de reabilitação e de pós-tratamento) para o tratamento do abuso de substâncias;
- Taxa de mortalidade por acidentes de trânsito;
- Cobertura da Atenção Primária à Saúde;

- Taxa de mortalidade atribuída a fontes de água inseguras, saneamento inseguro e falta de higiene;
- Prevalência de fumantes na população de quinze anos ou mais;
- Taxa de cobertura vacinal da população em relação às vacinas incluídas no Programa Nacional de Vacinação;
- Proporção de estabelecimentos de saúde que dispõem de um conjunto básico de medicamentos essenciais e relevantes disponíveis e a custo acessível em uma base sustentável;
- Número de profissionais de saúde por habitante.

Outros indicadores não contemplados acima, mas que também surgem em outras diretrizes específicas de cidades (tal como o Programa Cidades Sustentáveis), são:

- Crianças com baixo peso ao nascer;
- Estabelecimentos que realizam interrupção voluntária de gravidez segundo a lei;
- Excesso de peso em adultos e crianças;
- Gravidez na adolescência;
- Leitos hospitalares;
- Casos e óbitos por doenças sazonais e arboviroses;
- Cobertura de pré-natal;
- Estabelecimentos de saúde (clínicas da família, unidades básicas, postos de saúde etc.).

4.2.11 Indicadores sociais

A dimensão de indicadores sociais recebeu apenas uma contribuição na pesquisa. O especialista destacou a necessidade de ter dados sobre cobertura, investimento, recursos humanos e equipamentos de políticas públicas estaduais e municipais.

Nessa dimensão, por outro lado, o arcabouço dos ODS oferece um bom guia para indicadores essenciais para abertura de dados em estados e municípios. Abaixo, seguem alguns exemplos para os ODS 1 (Erradicação da Pobreza) e 2 (Fome Zero e Agricultura Sustentável):

- Proporção da população vivendo abaixo da linha da pobreza (internacional ou nacional);
- Proporção da população abrangida por regimes de proteção social;
- Proporção da população que vive em domicílios com acesso a serviços básicos;
- Número de mortes, de pessoas desaparecidas e de pessoas diretamente afetadas por desastres por 100 mil habitantes;
- Prevalência de insegurança alimentar moderada ou grave, baseado na Escala de Experiência de Insegurança Alimentar (FIES);
- Prevalência de subnutrição nas crianças com menos de cinco anos de idade.

Além desses, indicadores básicos de demografia devem estar disponíveis nos níveis estadual e municipal, uma vez que servem de balizadores para diversos outros indicadores e também para a formulação de políticas públicas. As estimativas populacionais e projeções por idade e sexo são essenciais nesse sentido.

Outros indicadores não contemplados acima, mas que também surgem em outras diretrizes específicas de cidades (tal como o Programa Cidades Sustentáveis), são:

- População em situação de rua ou habitando abrigos/albergues públicos;
- População inscrita no Cadastro Único.

4.2.12 Judiciário

A dimensão de dados do judiciário também recebeu apenas uma contribuição na pesquisa. O especialista destacou, na dimensão estadual, a necessidade de abertura de:

- Bases de dados dos processos, andamentos e respectivos detalhes em nível estadual para uso em pesquisas;
- Dados referentes a salários e verbas indenizatórias.

No caso da dimensão dos municípios, embora não exista judiciário no nível municipal, é fundamental que os municípios sejam transparentes em relação aos casos que levam ao judiciário. Também é importante abrir ao público os valores de honorários destinados aos procuradores pela atuação em casos perante o judiciário.

4.2.13 Segurança

Apesar de não ter havido nenhum respondente que tenha escolhido a categoria segurança, algumas respostas mencionaram a importância de se terem dados públicos nessa área, sem maiores detalhamentos, além da necessidade de harmonização de metodologias entre diferentes entes da federação.

A análise dos ODS permite identificar alguns indicadores relevantes para estados e municípios na área da segurança, em especial, os ODS 5 (Igualdade de Gênero) e 16 (Paz, Justiça e Instituições Eficazes), tais como:

- Proporção de mulheres de quinze anos de idade ou mais que sofreram violência física, sexual ou psicológica;
- Número de vítimas de homicídio intencional por 100 mil habitantes;
- Óbitos relacionados a conflitos por 100 mil habitantes;
- Proporção de armas apreendidas, encontradas ou entregues, cuja origem ou contexto ilícito tenham sido detectados ou estabelecidos por uma autoridade competente, em linha com instrumentos internacionais;
- Proporção de presos sem sentença em relação à população prisional em geral.

Além disso, algumas fontes nacionais, que consolidam dados na área da segurança, dão diretrizes sobre indicadores que devem ser publicados nos níveis de estados e municípios. Segundo o Atlas da Violência⁹³ e o Anuário Brasileiro de Segurança Pública⁹⁴, tem-se os seguintes registros aplicáveis a estados e municípios:

- Mortes violentas intencionais;
- Roubo e furto de veículos;
- Estupros.

Por se tratar de uma competência estadual, no entanto, a maior parte dos indicadores utilizados por essas publicações deve estar disponível em nível de estado. Os registros utilizados e não anteriormente mencionados são:

- Mortes a esclarecer;
- Pessoas desaparecidas e pessoas localizadas;
- Crimes violentos não letais contra o patrimônio;
- Suicídios;
- Crimes violentos não letais intencionais contra a pessoa;
- Tráfico de entorpecentes e posse e uso de entorpecentes;
- Registros de injúria racial;
- Número de delegacias especializadas e registros de crimes contra população de lésbicas, gays, bissexuais, travestis, transexuais e transgêneros (LGBT), incluindo número total de casos de violência contra essa população e número de denúncias;
- Homicídios de mulheres e feminicídios;
- Lesão corporal dolosa – violência doméstica;
- Registros de porte e posse ilegais de arma de fogo, em números absolutos e taxas;

⁹³ Disponível em: <http://forumseguranca.org.br/atlas-da-violencia/>. Acesso em: 27 mar. 2020.

⁹⁴ Disponível em: <https://slack-redirect.net/link?url=http%3A%2F%2Fforumseguranca.org.br%2Fanu%C3%A1rio-brasileiro-seguranca-publica%2F>. Acesso em: 28 mar. 2020.

- Registros de armas de fogo extraviadas, perdidas, furtadas ou roubadas;
- Quantidade de operações da Força Nacional por ano;
- Vitimização de diretores e professores na escola em que trabalham;
- Presos no sistema penitenciário e sob custódia das Polícias e taxas por 100 mil habitantes;
- Total de pessoas privadas de liberdade no sistema penitenciário e sob custódia das Polícias;
- Mortes violentas por causa indeterminada.

Outros indicadores não contemplados acima, mas que também surgem em outras diretrizes específicas de cidades (tal como o Programa Cidades Sustentáveis), são:

- Registro de adolescentes envolvidos em ato infracional ou cumprindo medidas socioeducativas;
- Registro de negligência e abandono de crianças e adolescentes.

4.2.14 Outros

Na dimensão relativa a outros dados, os especialistas que responderam a pesquisa sugeriram dados essenciais em duas grandes áreas: dados científicos e uso de tecnologia. No primeiro grupo, sugeriu-se a divulgação de dados gerados em pesquisas financiadas pelas fundações de apoio estaduais (no caso dos estados) e de pesquisa gerada por servidores da administração municipal (no caso dos municípios), além de dados sobre investimentos e projetos na área de ciência e tecnologia. No segundo grupo, destacou-se a importância de divulgar todos os dados não sensíveis ou pessoais que dizem respeito ao uso de tecnologias de informação e comunicação (TICs) pelo governo, além de divulgar esforços em pesquisas que usam tecnologia (em especial, inteligência artificial) nas mais diversas áreas (saúde, segurança, transportes etc.).

4.3 Considerações da seção

A Seção 4 não pretende ser um levantamento exaustivo e mandatório de dados e indicadores que devem estar abertos por todo estado e/ou município, a depender de sua competência, mas sim um modelo com diretrizes gerais, que pode ser seguido por tais entes federativos. Partindo das experiências de especialistas, bem como de outras iniciativas globais e nacionais, como é o caso dos ODS e do Programa Cidades Sustentáveis, espera-se que esse resultado possa contribuir para a criação de um modelo de referência de dados abertos aplicável a todos os entes da Federação brasileira.

Por fim, vale destacar que alguns especialistas entrevistados ressaltaram que não haveria necessidade de priorizar determinadas bases de dados para abertura ou eleger dados que seriam, em alguma medida, essenciais. Nessa visão, parte-se do pressuposto de que todo dado público não sigiloso deve ser aberto, independentemente se há ou não uma demanda a priori daquela informação pela sociedade. No entanto, tendo em vista que alguns dados já são mais demandados do que outros e que o poder público enfrenta a decisão sobre em que direção enveredar maiores esforços de abertura, o modelo aqui proposto busca oferecer linhas gerais de dados considerados prioritários, sem, no entanto, diminuir a importância dos outros dados que também podem entrar no processo de abertura.

5. Diretrizes de abertura de dados definidas por temas nacionais estratégicos

De forma complementar ao levantamento de indicadores focado na dimensão subnacional, realizado na Seção 4, esta Seção aborda, em nível nacional, a definição de temas e diretrizes de abertura de dados úteis para a construção e a avaliação de políticas públicas, conforme definidos pela eleição de temas nacionais estratégicos. Trata-se, nesse sentido, de um conjunto de diretrizes já previamente definidas, em âmbito nacional, que engendra a necessidade de dados e indicadores essenciais para acompanhamento de metas.

Exemplos desses indicadores seriam a qualidade da água em córregos urbanos, a disponibilidade de iluminação pública, a oferta de coleta de esgoto por rua, os destinos de resíduos sólidos urbanos, a qualidade de pavimentação urbana e rodoviária, a disponibilidade de transporte público, a qualidade de infraestrutura instalada de escolas e creches públicas, os índices de satisfação com merenda escolar por escola, a quantidade de médicos especialistas em hospitais públicos por região, entre outros.

Como o levantamento realizado na seção 4 foi bastante extenso no sentido de mapear dados e indicadores essenciais em diversas áreas, e considerando que vários desses indicadores também poderiam ser fornecidos ou consolidados em nível nacional, apresenta-se aqui uma identificação, para cada área mapeada, de planos nacionais estratégicos que contém informações sobre metas a serem cumpridas e que precisam de dados e indicadores de monitoramento e avaliação.

5.1 Educação

- Plano Nacional de Educação⁹⁵;

⁹⁵ Disponível em: <http://pne.mec.gov.br/18-planos-subnacionais-de-educacao/543-plano-nacional-de-educacao-lei-n-13-005-2014>. Acesso em: 10 set. 2020.

- Política Nacional de Educação Infantil⁹⁶;
- Política Nacional de Educação Especial⁹⁷.

5.2 Saúde

- Plano Nacional de Saúde⁹⁸;
- Política Nacional de Atenção Básica⁹⁹;
- Política Nacional de Promoção da Saúde¹⁰⁰;
- Política Nacional de Ciência, Tecnologia e Inovação em Saúde¹⁰¹.

5.3 Meio ambiente e saneamento

- Plano Nacional de Saneamento Básico¹⁰²;
- Política Nacional do Meio Ambiente¹⁰³;
- Política Nacional de Resíduos Sólidos¹⁰⁴.

⁹⁶ Disponível em: <http://portal.mec.gov.br/seb/arquivos/pdf/polinaci.pdf>. Acesso em: 10 set. 2020.

⁹⁷ Disponível em: <http://portal.mec.gov.br/arquivos/pdf/politicaeducespecial.pdf>. Acesso em: 10 set. 2020.

⁹⁸ Disponível em: http://bvsmms.saude.gov.br/bvs/publicacoes/plano_nacional_saude_2016_2019_30032015_final.pdf. Acesso em: 10 set. 2020.

⁹⁹ Disponível em: <http://189.28.128.100/dab/docs/publicacoes/geral/pnab.pdf>. Acesso em: 10 set. 2020.

¹⁰⁰ Disponível em: http://bvsmms.saude.gov.br/bvs/publicacoes/politica_nacional_promocao_saude_3ed.pdf. Acesso em: 10 set. 2020.

¹⁰¹ Disponível em: http://bvsmms.saude.gov.br/bvs/publicacoes/Politica_Portugues.pdf. Acesso em: 10 set. 2020.

¹⁰² Disponível em: <https://www.mma.gov.br/cidades-sustentaveis/qualidade-do-ar/item/485-plano-nacional-de-saneamento-b%C3%A1sico>. Acesso em 10 set. 2020.

¹⁰³ Disponível em: http://www.planalto.gov.br/ccivil_03/LEIS/L6938.htm. Acesso em: 10 set. 2020.

¹⁰⁴ Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2007-2010/2010/lei/l12305.htm. Acesso em: 10 set. 2020.

5.4 Infraestrutura e mobilidade urbana

- Política Nacional de Transportes¹⁰⁵;
- Plano Nacional de Logística e Transportes¹⁰⁶;
- Política Nacional de Mobilidade Urbana¹⁰⁷;
- Plano Nacional de Habitação¹⁰⁸.

5.5 Segurança

- Plano Nacional de Política Criminal e Penitenciária¹⁰⁹;
- Plano Nacional de Segurança Pública e Defesa Social¹¹⁰;
- Política Nacional de Segurança Pública¹¹¹.

5.6 Indicadores sociais

- Política Nacional de Assistência Social¹¹²;

¹⁰⁵ Disponível em: <https://www.gov.br/infraestrutura/pt-br/assuntos/politica-e-planejamento/politica-e-planejamento/pnt>. Acesso em: 10 set. 2020.

¹⁰⁶ Disponível em: <https://www.gov.br/infraestrutura/pt-br/assuntos/transporte-terrestre/conheca-o-pnlt>. Acesso em: 10 set.2020.

¹⁰⁷ Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2012/lei/l12587.htm. Acesso em: 10 set. 2020.

¹⁰⁸ Disponível em: https://urbanismo.mppr.mp.br/arquivos/File/Habitacao/Material_de_Apoio/PLANONACIONAL-DEHABITAO.pdf. Acesso em: 11 set. 2020.

¹⁰⁹ Disponível em: http://depen.gov.br/DEPEN/depen/cnppcp/plano_nacional-1/Plano_Nacional_de_Politica_Criminal_e_Penitenciaria_2020_2023_FINAL_.pdf. Acesso em: 26 mar. 2020.

¹¹⁰ Disponível em: <https://www.justica.gov.br/news/collective-nitf-content-1546269936.75>. Acesso em: 10 set. 2020.

¹¹¹ Disponível em: https://www.justica.gov.br/news/politica-nacional-de-seguranca-publica/minuta_pnasp.pdf. Acesso em: 10 set. 2020.

¹¹² Disponível em: http://www.mds.gov.br/webarquivos/publicacao/assistencia_social/Normativas/PNAS2004.pdf. Acesso em: 10 set. 2020.

- Plano Nacional de Segurança Alimentar e Nutricional¹¹³;
- Política Nacional para a População em Situação de Rua¹¹⁴.

5.7 Orçamento público

- Lei de Responsabilidade Fiscal¹¹⁵.

5.8 Propriedade e uso da terra

- Política Nacional de Regularização Fundiária¹¹⁶.

5.9 Judiciário

- Estratégia Nacional do Poder Judiciário¹¹⁷;
- Estratégia da Justiça Federal¹¹⁸.

5.10 Indicadores econômicos

- Estratégia Nacional de Desenvolvimento Econômico e Social¹¹⁹.

¹¹³ Disponível em: http://www.mds.gov.br/webarquivos/arquivo/seguranca_alimentar/caisan/Publicacao/Caisan_Nacional/PLANSAN%202016-2019_revisado_completo.pdf. Acesso em: 10 set. 2020.

¹¹⁴ Disponível em: http://www.planalto.gov.br/ccivil_03/_Ato2007-2010/2009/Decreto/D7053.htm. Acesso em: 10 set. 2020.

¹¹⁵ Disponível em: http://www.planalto.gov.br/ccivil_03/leis/lcp/lcp101.htm. Acesso em: 10 set. 2020.

¹¹⁶ Disponível em: http://www.planalto.gov.br/ccivil_03/_Ato2015-2018/2017/Lei/L13465.htm. Acesso em: 10 set. 2020.

¹¹⁷ Disponível em: <https://www.cnj.jus.br/gestao-estrategica-e-planejamento/estrategia-nacional-do-poder-judicial-2015-2020/>. Acesso em: 10 set. 2020.

¹¹⁸ Disponível em: <http://www.jfsp.jus.br/documentos/administrativo/NUCS/campanhas/permanentes/estrategiajf/CadernoEstrategia20152020editorial.pdf>. Acesso em: 10 set. 2020.

¹¹⁹ Disponível em: <http://www2.planejamento.gov.br/assuntos/planeja/endes>. Acesso em: 10 set. 2020.

5.11 Dados geográficos

- Política Nacional de Geoinformação.

5.12 Ciência, tecnologia e inovação

- Estratégia Nacional de Ciência, Tecnologia e Inovação¹²⁰.

5.13 Contratos e licitações

- Lei n.º 8.666/1993¹²¹;
- Orientações e jurisprudência do Tribunal de Contas da União (TCU).

5.14 Legislativo

- Interlegis¹²².

¹²⁰ Disponível em: http://www.finep.gov.br/images/a-finep/Politica/16_03_2018_Estrategia_Nacional_de_Ciencia_Tecnologia_e_Inovacao_2016_2022.pdf. Acesso em: 10 set. 2020.

¹²¹ Disponível em: http://www.planalto.gov.br/ccivil_03/leis/l8666cons.htm. Acesso em: 10 set. 2020.

¹²² Disponível em: <https://www.interlegis.leg.br/>. Acesso em: 10 set. 2020.

6. Governança dos dados abertos

Modelos de governança de dados abertos de governo continuam sendo objeto de discussão em todo o mundo, devido às diferentes abordagens que incluem administração, política, legislação, padrões tecnológicos e capacidade para disponibilizar dados, entre outros (como, por exemplo, os aspectos culturais e as diferentes formas de como países ou regiões são governados e administrados).

Neste documento, é adotado o entendimento mais comum de que governança de dados abertos é o resultado da interação de regras, padrões, ferramentas, princípios, processos e decisões que influenciam quais dados governamentais são abertos, como e por quem¹²³.

É também um entendimento comum que uma estrutura organizacional – por exemplo, um comitê – deve ser criada para tratar da governança de dados. Esse comitê deve ter “competência para deliberar, dentre outras, sobre as orientações e as diretrizes para a categorização de compartilhamento amplo, restrito e específico, e a forma e o meio de publicação dessa categorização, observada a legislação pertinente e a referente à proteção de dados pessoais”¹²⁴.

6.1 Fundamentação

A iniciativa Open Government Partnership (OGP) demonstra que a relevância do tema de governança de dados abertos está sempre presente quando se discute origem, qualidade, confiança, legalidade da publicação, validade contextual e temporal, entre outros atributos. É importante destacar também que:

- Os recentes desenvolvimentos tecnológicos para análise de grandes quantidades de dados e de perfis individuais aumentaram o nível de consciência (e de críticas) em

¹²³ World Wide Web Foundation. Disponível em: <https://webfoundation.org/>. Acesso em: 10 set. 2020.

¹²⁴ Ver, por exemplo, o Decreto 10.046, de 9 de outubro de 2019. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2019-2022/2019/decreto/D10046.htm. Acesso em: 10 set. 2020.

relação às crescentes assimetrias de poder entre grandes atores digitais, sociedade civil e governos. No entanto, relativamente pouco se sabe sobre como essas tendências tecnológicas podem realmente servir ao bem público ao inovar a formulação de políticas econômicas e de inovação em serviços com o setor público.

- A crescente atenção mundial à inteligência artificial (IA) torna crucial garantir dados de boa qualidade, para treinar novos algoritmos e desenvolver novas aplicações ou serviços. Grandes massas de dados (big data) formam o combustível que alimenta a IA. É preciso desenvolver estratégias nacionais e mundiais, para promover uma formulação de políticas de base de dados ética, justa e centrada no cidadão.

Do lado do governo, independentemente do tipo, se nacional ou subnacional, os principais atores são:

- **Tomadores de decisão:** Definição de quem são os agentes públicos que lideram ou têm autoridade de decisão sobre dados abertos em reuniões, procedimentos, condução, debate, votação e outras questões;
- **Detentores de dados:** Definição de quais são os órgãos ou entidades governamentais que gerenciam e administram dados, bem como garantem a integridade dos dados;
- **Produtores de dados:** Definição de quais são os órgãos ou entidades governamentais que produzem qual tipo de informação do setor público;
- **Garantia de qualidade de dados:** Definição de quem são os atores que garantem que os dados produzidos estão de acordo com os padrões de qualidade e se não há algum tipo de conflito com a publicação como dados abertos;
- **Guardiões ou administradores de dados:** Definição de quem são os atores que controlam a publicação de dados abertos.

Outro aspecto importante sobre a governança de dados a ser considerado é a necessidade de informar para a sociedade civil quem (ou quais atores) está (estão) fazendo o que. A indicação dos atores e de seus papéis, mencionados no parágrafo anterior, ajudam os governos a darem respostas às seguintes questões:

- Quem detém o poder de afirmar autoridade sobre a publicação de dados abertos?
- Que papéis os diferentes atores do governo desempenham para formular políticas e implementá-las?

- Que formas de modelos de governança podem ser derivadas dessas funções e responsabilidades?
- Podemos ver um padrão comum de como o poder de tomada de decisão é distribuído?
- Como esses modelos de governança diferem?
- Quais são os critérios para avaliar o desempenho dos modelos de governança observados?
- Como eles influenciam, por exemplo, a política e a implementação de dados abertos?

Nesse contexto, é sugerido que órgãos e entidades da administração pública definam um modelo de governança multissetorial. O Brasil é internacionalmente reconhecido pelos modelos de governança do Comitê Gestor da Internet (CGI.br)¹²⁵, que inspirou também, parcialmente, a instituição do Comitê Gestor da Infraestrutura dos Dados Abertos (CGINDA).

Um modelo de governança desse tipo pressupõe que os diversos atores interessados em dados abertos devem participar da construção e da gestão da disponibilização dos dados na Web. Assim, a governança sobre a publicação de dados abertos governamentais deve ser exercida de forma transparente, multilateral e democrática, com a participação dos vários setores da sociedade, preservando e estimulando o seu caráter de criação coletiva.

Sugerimos alguns pilares essenciais. Parte desses pilares propostos foi adaptada do documento *Princípios para a Governança e Uso da Internet no Brasil*¹²⁶, do Comitê Gestor da Internet no Brasil. Embora seja dirigido para uso da internet de uma maneira geral, consideramos os pilares válidos para a governança de dados abertos, adicionando um pilar sobre a integridade dos dados. Os pilares propostos são discutidos a seguir.

6.1.1 Participação multilateral

Os diversos atores interessados devem estar representados no modelo de governança e nas instâncias de decisão – pelo menos os quatro setores macro da sociedade: governo, setor privado, universidades e organizações da sociedade civil.

¹²⁵ Disponível em: <https://www.cgi.br/>. Acesso em: 10 set. 2020.

¹²⁶ Disponível em: <http://www.cgi.br/regulamentacao/resolucao2009-003.htm>. Acesso em: 10 set. 2020.

6.1.2 Diversidade

A diversidade dos atores, dentro de cada setor da sociedade civil, deve ser respeitada e preservada, e sua expressão deve ser estimulada, bem como representada nas diversas instâncias da governança.

6.1.3 Processo democrático

Os participantes das instâncias de decisão devem ser eleitos por seus pares, exceto os do setor público, que são indicados. Esse processo democrático pode ser mapeado, por exemplo, na forma de um regimento interno de eventual comitê de governança de dados a ser criado (ver, por exemplo, o Decreto n.º 10.403/2020¹²⁷, mencionado mais adiante).

6.1.4 Inovação

A governança da OGP Brasil deve promover a contínua evolução e a ampla difusão de novas tecnologias e modelos de transparência e de controle social.

6.1.5 Colaboração

Os processos de governança da OGP Brasil devem preservar a dinâmica da parceria como espaço de colaboração.

¹²⁷ Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2019-2022/2020/Decreto/D10403.htm. Acesso em: 10 set. 2020.

6.1.6 Transparência

Todas as discussões e decisões levadas nas instâncias decisoras do modelo de governança devem ser públicas, abertas, transmitidas via Web e comunicadas ao público.

6.1.7 Integridade

Deve ser garantida a integridade dos dados disponibilizados, bem como a integridade dos conteúdos relacionados aos conjuntos de dados.

6.2 Legislação associada

O Governo Federal tem extensa legislação sobre governança de dados, sem, contudo, mencionar explicitamente se é extensiva para dados abertos¹²⁸. Em especial, o Decreto n.º 10.403¹²⁹, de 19 de junho de 2020, altera o Decreto n.º 10.046, de 9 de outubro de 2019, que dispõe sobre a governança no compartilhamento de dados no âmbito da administração pública federal e institui o Cadastro Base do Cidadão e o Comitê Central de Governança de Dados. Para este documento, os autores entendem que essa legislação pode também ser aplicada especificamente para dados abertos em conjunto com a legislação relacionada à Infraestrutura Nacional de Dados Abertos (INDA).

A Estratégia de Governo Digital (EGD)¹³⁰ abre oportunidades para não apenas estabelecer plano de dados abertos para órgãos e entidades, mas também para que essas instituições, ao elaborarem seus planos, pensem a governança de dados abertos de forma integrada, levando em consideração a Lei Geral de Proteção de Dados (LGPD).

¹²⁸ Legislação Governança de Dados e Interoperabilidade. Disponível em: <https://www.gov.br/governodigital/pt-br/legislacao/legislacao-governanca-de-dados-e-interoperabilidade>. Acesso em: 10 set. 2020.

¹²⁹ Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2019-2022/2020/Decreto/D10403.htm. Acesso em: 10 set. 2020.

¹³⁰ Instituída pelo Decreto n.º 10.332, de 28 de abril de 2020. Disponível em: <https://www.in.gov.br/en/web/dou/-/decreto-n-10.332-de-28-de-abril-de-2020-254430358>. Acesso em: 10 set. 2020.

Referências

ARY, T.; CARVALHO, P.; WOSGRAU, A. Dados abertos: como fazer um PDA. In: FÓRUM “O CONTROLE NO COMBATE À CORRUPÇÃO”, 1., 2019, Brasília, DF. *Caderno de resumos*. Brasília, DF: Controladoria Geral da União, 2019. Disponível em: <https://repositorio.cgu.gov.br/handle/1/43573>. Acesso em: 11 set. 2020.

BRASIL. Lei n.º 6.938, de 31 de agosto de 1981. Dispõe sobre a Política Nacional do Meio Ambiente, seus fins e mecanismos de formulação e aplicação, e dá outras providências. Brasília, DF, 31 ago. 1981. Disponível em: http://www.planalto.gov.br/ccivil_03/LEIS/L6938.htm. Acesso em: 26 mar. 2020.

BRASIL. Lei n.º 8.666, de 21 de junho 1993. Regulamenta o art. 37, inciso XXI, da Constituição Federal, institui normas para licitações e contratos da Administração Pública e dá outras providências. Brasília, DF, 21 jun. 1993. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/l8666cons.htm. Acesso em: 26 mar. 2020.

BRASIL. Senado Federal. *Interlegis*: Instituto Legislativo Brasileiro. Brasília, DF: Senado Federal, 1997. Disponível em: <https://www.interlegis.leg.br/>. Acesso em: 26 mar. 2020.

BRASIL. Lei Complementar n.º 101, de 4 de maio de 2000. Estabelece normas de finanças públicas voltadas para a responsabilidade na gestão fiscal e dá outras providências. Brasília, DF, 04 maio 2000. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/lcp/lcp101.htm. Acesso em: 26 mar. 2020.

BRASIL. Ministério da Educação. Secretaria de Educação Básica. *Política Nacional de Educação Infantil: pelo direito das crianças de zero a seis anos à Educação*. Brasília, DF: MEC, 2003. Disponível em: <http://portal.mec.gov.br/seb/arquivos/pdf/polinaci.pdf>. Acesso em: 26 mar. 2020.

BRASIL. Ministério das Cidades. *Política nacional de habitação*. Brasília, DF: Ministério das Cidades, 2004. Disponível em: <https://www.cidades.gov.br/images/stories/ArquivosSNH/Arquivos-PDF/4PoliticaNacionalHabitacao.pdf>. Acesso em: 26 mar. 2020.

BRASIL. Ministério do Desenvolvimento Social e Combate à Fome. Secretaria Nacional de As-

sistência Social. *Política Nacional de Assistência Social*. Brasília, DF: MDS, 2004. Disponível em: http://www.mds.gov.br/webarquivos/publicacao/assistencia_social/Normativas/PNAS2004.pdf. Acesso em: 26 mar. 2020.

BRASIL. Ministério do Meio Ambiente. *Plano Nacional de Saneamento Básico*. Brasília, DF: MMA, 2007. Disponível em: <https://www.mma.gov.br/cidades-sustentaveis/qualidade-do-ar/item/485-plano-nacional-de-saneamento-b%C3%A1sico>. Acesso em: 26 mar. 2020.

BRASIL. Ministério da Educação. Secretaria de Educação Especial. *Política Nacional de Educação Especial na Perspectiva da Educação Inclusiva*. Brasília, DF: MEC, 2008. Disponível em: <http://portal.mec.gov.br/arquivos/pdf/politicaeducespecial.pdf>. Acesso em: 26 mar. 2020.

BRASIL. Ministério da Saúde. Secretaria de Ciência, Tecnologia e Insumos Estratégicos. Departamento de Ciência e Tecnologia. *Política Nacional de Ciência, Tecnologia e Inovação em Saúde*. Brasília, DF: Ministério da Saúde, 2008. Disponível em: http://bvsmms.saude.gov.br/bvs/publicacoes/Politica_Portugues.pdf. Acesso em: 26 mar. 2020.

BRASIL. Decreto n.º 7.053, de 23 de dezembro de 2009. Institui a Política Nacional para a População em Situação de Rua e seu Comitê Intersectorial de Acompanhamento e Monitoramento, e dá outras providências. Brasília, DF, 23 jan. 2009. Disponível em: http://www.planalto.gov.br/ccivil_03/_Ato2007-2010/2009/Decreto/D7053.htm. Acesso em: 26 mar. 2020.

BRASIL. Ministério do Desenvolvimento Regional. *Plano Nacional de Habitação*. Brasília: DF: MDR, 2009. Disponível em: https://urbanismo.mppr.mp.br/arquivos/File/Habitacao/Material_de_Apoio/PLANONACIONALDEHABITAO.pdf. Acesso em: 26 mar. 2020.

BRASIL. Lei n.º 12.305, de 2 de agosto de 2010. Institui a Política Nacional de Resíduos Sólidos; altera a Lei no 9.605, de 12 de fevereiro de 1998; e dá outras providências. Brasília, DF, 02 ago. 2010. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2007-2010/2010/lei/l12305.htm. Acesso em: 26 mar. 2020.

BRASIL. Ministério da Saúde. Secretaria de Vigilância em Saúde. Secretaria de Atenção à Saúde. *Política Nacional de Promoção da Saúde*. Brasília, DF: Ministério da Saúde, 2010. Disponível em: http://bvsmms.saude.gov.br/bvs/publicacoes/politica_nacional_promocao_saude_3ed.pdf. Acesso em: 26 mar. 2020.

BRASIL. Lei n.º 12.527, de 18 de novembro de 2011 (Lei de Acesso à Informação). Brasília, DF, 18 nov. 2011. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/l12527.htm. Acesso em: 15 mar. 2020.

BRASIL. Lei n.º 12.587, de 3 de janeiro de 2012. Institui as diretrizes da Política Nacional de Mobilidade Urbana; revoga dispositivos dos Decretos-Leis n.ºs 3.326, de 3 de junho de 1941, e 5.405, de 13 de abril de 1943, da Consolidação das Leis do Trabalho (CLT), aprovada pelo Decreto-Lei n.º 5.452, de 1º de maio de 1943, e das Leis n.ºs 5.917, de 10 de setembro de 1973, e 6.261, de 14 de novembro de 1975; e dá outras providências. Brasília, DF, 03 jan. 2012. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2012/lei/l12587.htm. Acesso em: 26 mar. 2020.

BRASIL. Ministério da Saúde. *Política Nacional de Atenção Básica*. Brasília, DF: Ministério da Saúde, 2012. Disponível em: <http://189.28.128.100/dab/docs/publicacoes/geral/pnab.pdf>. Acesso em: 26 mar. 2020.

BRASIL. Ministério do Planejamento, Orçamento e Gestão. Instrução Normativa n.º 4. Institui a Infraestrutura Nacional de Dados Abertos – INDA. Brasília, DF: MPDG, 2012. Disponível em: <https://www.gov.br/governodigital/pt-br/legislacao/InstrucaoNormativaINDA42012.pdf>. Acesso em: 27 mar. 2020.

BRASIL. Conselho Nacional de Justiça. *Estratégia Nacional do Poder Judiciário*. Brasília, DF: CNJ, 2014. Disponível em: <https://www.cnj.jus.br/gestao-estrategica-e-planejamento/estrategia-nacional-do-poder-judiciario-2015-2020/>. Acesso em: 26 mar. 2020.

BRASIL. Lei n.º 13.005, de 25 de junho de 2014. Aprova o Plano Nacional de Educação - PNE e dá outras providências., Brasília, DF, 14 jun. 2014. Disponível em: <http://pne.mec.gov.br/18-planos-subnacionais-de-educacao/543-plano-nacional-de-educacao-lei-n-13-005-2014>>. Acesso em: 26 mar. 2020.

BRASIL. Ministério da Infraestrutura. *Plano Nacional de Logística e Transportes*. Brasília, DF: Ministério da Infraestrutura, 2014. Disponível em: <https://www.gov.br/infraestrutura/pt-br/assuntos/transporte-terrestre/conheca-o-pnlt>. Acesso em: 26 mar. 2020.

BRASIL. Decreto n.º 8.777, de 11 de maio de 2016. Institui a Política de Dados Abertos do Poder Executivo federal. Brasília, DF, 11 maio 2016. Disponível em: <http://www.planalto.gov.br/>

ccivil_03/_ato2015-2018/2016/decreto/d8777.htm. Acesso em: 26 mar. 2020.

BRASIL. Ministério da Ciência, Tecnologia, Inovações e Comunicações. Centro de Gestão e Estudos Estratégicos. *Estratégia Nacional de Ciência, Tecnologia e Inovação*. Brasília, DF: MCTI, 2016. Disponível em: http://www.finep.gov.br/images/a-finep/Politica/16_03_2018_Estrategia_Nacional_de_Ciencia_Tecnologia_e_Inovacao_2016_2022.pdf. Acesso em: 26 mar. 2020.

BRASIL. Ministério da Saúde. *Plano Nacional de Saúde*. Brasília, DF: Ministério da Saúde, 2016. Disponível em: http://bvsms.saude.gov.br/bvs/publicacoes/plano_nacional_saude_2016_2019_30032015_final.pdf. Acesso em: 26 mar. 2020.

BRASIL. Lei n.º 13.465, 11 de julho de 2017. Dispõe sobre a regularização fundiária rural e urbana, sobre a liquidação de créditos concedidos aos assentados da reforma agrária e sobre a regularização fundiária no âmbito da Amazônia Legal; institui mecanismos para aprimorar a eficiência dos procedimentos de alienação de imóveis da União. Brasília, DF, 11 jul. 2017. Disponível em: http://www.planalto.gov.br/ccivil_03/_Ato2015-2018/2017/Lei/L13465.htm. Acesso em: 26 mar. 2020.

BRASIL. Lei n.º 13.709, de 14 de agosto de 2018. Lei Geral de Proteção de Dados Pessoais (LGPD). Brasília, DF, 14 ago. 2018. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/L13709.htm. Acesso em: 15 mar. 2020.

BRASIL. Câmara Interministerial de Segurança Alimentar e Nutricional. *Plano Nacional de Segurança Alimentar e Nutricional*. Brasília, DF: CAISAN, 2018. Disponível em: http://www.mds.gov.br/webarquivos/arquivo/seguranca_alimentar/caisan/Publicacao/Caisan_Nacional/PLAN-SAN%202016-2019_revisado_completo.pdf. Acesso em: 26 mar. 2020.

BRASIL. Ministério da Economia. *Estratégia Nacional de Desenvolvimento Econômico e Social*. Brasília, DF: Ministério da Economia, 2018. Disponível em: <http://www2.planejamento.gov.br/assuntos/planeja/endes>. Acesso em: 10 set. 2020.

BRASIL. Ministério da Justiça e Segurança Pública. *Política Nacional de Segurança Pública*. Brasília, DF: Ministério da Justiça e Segurança Pública, 2018. Disponível em: https://www.justica.gov.br/news/politica-nacional-de-seguranca-publica/minuta_pnasp.pdf. Acesso em: 26 mar. 2020.

BRASIL. Ministério da Segurança Pública. Departamento Penitenciário Nacional. *Plano Nacio-*

nal de Segurança Pública e Defesa Social. Brasília, DF: Ministério da Segurança Pública, 2018. Disponível em: <https://www.justica.gov.br/news/collective-nitf-content-1546269936.75>. Acesso em: 26 mar. 2020.

BRASIL. Ministério dos Transportes, Portos e Aviação Civil. *Política Nacional de Transportes*. Brasília, DF: MTPA, 2018. Disponível em: <https://www.gov.br/infraestrutura/pt-br/assuntos/politica-e-planejamento/politica-e-planejamento/pnt>. Acesso em: 26 mar. 2020.

BRASIL. Decreto n.º 9.903, de 8 de julho de 2019. Altera o Decreto n.º 8.777, de 11 de maio de 2016, que institui a Política de Dados Abertos do Poder Executivo federal, para dispor sobre a gestão e os direitos de uso de dados abertos. Brasília, DF, 08 jul. 2019. Disponível em: http://www.planalto.gov.br/ccivil_03/_Ato2019-2022/2019/Decreto/D9903.htm. Acesso em: 10 set. 2020.

BRASIL. Lei n.º 10.046, de 09 de outubro de 2019. Dispõe sobre a governança no compartilhamento de dados no âmbito da administração pública federal e institui o Cadastro Base do Cidadão e o Comitê Central de Governança de Dados. Brasília, DF, 09 out. 2019. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2019-2022/2019/decreto/D10046.htm. Acesso em: 10 set. 2020.

BRASIL. Ministério da Justiça e Segurança Pública. Departamento Penitenciário Nacional. *Plano Nacional de Política Criminal e Penitenciária*. Brasília, DF: Ministério da Justiça e Segurança Pública, 2019. Disponível em: http://depen.gov.br/DEPEN/depen/cnppcp/plano_nacional-1/Plano_Nacional_de_Politica_Criminal_e_Penitenciaria_2020_2023__FINAL_.pdf. Acesso em: 26 mar. 2020.

BRASIL. Lei n.º 10.403, de 19 de junho de 2020. Altera o Decreto n.º 10.046, de 9 de outubro de 2019, que dispõe sobre a governança no compartilhamento de dados no âmbito da administração pública federal e institui o Cadastro Base do Cidadão e o Comitê Central de Governança de Dados. Brasília, DF, 19 jun. 2020. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2019-2022/2020/Decreto/D10403.htm. Acesso em: 10 set. 2020.

BRASIL. Ministério da Saúde. *Coronavírus: SUS*. Brasília, DF: Ministério da Saúde, 2020. Disponível em: https://play.google.com/store/apps/details?id=br.gov.datasus.guardioes&hl=pt_BR. Acesso em: 23 abr. 2020.

BUENO, S.; LIMA, R. S. (orgs.). *Anuário Brasileiro de Segurança Pública*. São Paulo: Fórum

Brasileiro de Segurança Pública, 2019. Disponível em: <https://slack-redir.net/link?url=http%3A%2F%2Fforumseguranca.org.br%2Fanoario-brasileiro-seguranca-publica%2F>. Acesso em: 28 mar. 2020.

CERQUEIRA, D.; BUENO, S. (coords.). *Atlas da Violência*. São Paulo: Fórum Brasileiro de Segurança Pública, 2019. Disponível em: <http://forumseguranca.org.br/atlas-da-violencia/>. Acesso em: 27 mar. 2020.

CORREA, A. S.; ZANDER, P. O.; SILVA, F. C. Investigating open data portals automatically: a methodology and some illustrations. In: ANNUAL INTERNATIONAL CONFERENCE ON DIGITAL GOVERNMENT RESEARCH, 19., 2018, Delft. *Proceedings... Nova York: Association for Computing Machinery, 2018*. Disponível em: <https://dl.acm.org/doi/10.1145/3209281.3209292>. Acesso em: 10 set. 2020.

GONÇALVES, A. C. G.; VALENTE, M.; PADILLA, N. M. *Orientações para adoção de linguagem clara*. São Paulo: Seade, 2016. Disponível em: http://www.governoaberto.sp.gov.br/wp-content/uploads/2017/12/orientacoes_para_adocao_linguagem_clara_ptBR.pdf. Acesso em: 27 mar. 2020.

GRAY, J.; DIETRICH, D.; MCNAMARA, T.; POIKOLA, A.; POLLOCK, R.; TAIT, J.; ZIJLSTRA, T. *Manual dos Dados Abertos: Governo*. Tradução: Comunidade Transparência Hacker, Esfera.mobi e Everton Zanella Alvarenga. São Paulo: W3C Brasil, 2011. Disponível em: https://www.w3c.br/pub/Materiais/PublicacoesW3C/Manual_Dados_Abertos_WEB.pdf. Acesso em: 24 ago. 2020.

HERMAN, I.; ADIDA, B.; SPORNY, M.; BIRBECK, M. (eds.). *RDFa 1.1 Primer - Third Edition: rich structured data markup for web documents*. 2015. Disponível em: <https://www.w3.org/TR/rdfa-primer/>. Acesso em: 26 mar. 2020.

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA. *Ciência Aberta é Vida: Diretório de fontes de informação científica de livre acesso sobre o Coronavírus*. Rio de Janeiro: IBICT, 2020. Disponível em: <http://diretoriodefontes.ibict.br/coronavirus/>. Acesso em: 24 abr. 2020.

KIM, J.J.; HAUSENBLAS, M. *Five stars Open Data*. [S. l.: s. n.], 2015. Disponível em: <https://5star-data.info/en/>. Acesso em 20 mar. 2020.

KLEIN, R. H.; KLEIN, D. C. B.; LUCIANO, E. M. Identificação de mecanismos para a ampliação da transparência em portais de dados abertos: uma análise no contexto brasileiro. *Cadernos*

Ebape.br, v. 16, n. 4, p. 692-715, 2018. Disponível em: <http://dx.doi.org/10.1590/1679-395173241>. Acesso em: 21 mar. 2020.

LÓSCIO, B. F.; BURLE, C.; OLIVEIRA, M. I. S.; CALEGARI, N. *Fundamentos para publicação de dados na Web*. São Paulo: Comitê Gestor da Internet no Brasil, 2018. Disponível em: <https://ceweb.br/publicacao/livro-fundamentos-dados-web/>. Acesso em: 13 ago. 2020.

LÓSCIO, B. F.; BURLE, C.; CALEGARI, N. *Boas práticas para dados na Web: Recomendação do W3C de 31 de janeiro de 2017*. São Paulo: W3C Brasil, 2019. Disponível em: <https://w3c.br/traducoes/DWBP-pt-br/>. Acesso em: 10 set. 2020.

NIEVINSKI, F. G. *Dados “meio” abertos: sobre o uso e reuso dos dados governamentais brasileiros*. [S. l.]: Open Knowledge Brasil, 2013. Disponível em: <https://br.okfn.org/2013/08/28/dados-meio-abertos-sobre-o-uso-e-reuso-dos-dados-governamentais-brasileiros/>. Acesso em: 23 mar. 2020.

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*. Fisciano: ROUT-TO-PA, 2015. 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: 11 set. 2020.

PASQUETTO, I.; RANGLES, B.; BORGMAN, C. On the Reuse of Scientific Data. *Data Science Journal*, n. 16, p. 8, 2017. Disponível em: <https://doi.org/10.5334/dsj-2017-008>. Acesso em 23: mar. 2020.

PIRES, M. T. *Guia de Dados Abertos*. São Paulo: Comitê Gestor da Internet no Brasil, 2015. Disponível em: <https://ceweb.br/publicacao/guia-de-dados-abertos/>. Acesso em: 25 mar. 2020.

REINHARD, N.; GERMANO, E. C. *Guia de Incentivo ao (Re)uso de Dados Abertos*. São Paulo: Comitê Gestor de Dados Abertos, 2016. Disponível em: <https://ceweb.br/publicacao/guia-de-incentivo-ao-reuso-de-dados-abertos/>. Acesso em: 29 mar. 2020.

RUEDIGER, M. (Coord.). *Modelos de abertura de dados: potencialidades e limites*. Brasília: FGV DAPP, 2019. Disponível em: <https://www.gov.br/cgu/pt-br/governo-aberto/a-ogp/planos-de-acao/4o-plano-de-acao-brasileiro/compromisso-2-docs/marcos-2-e-3-modelos-de-abertura-de-dados.pdf>. Acesso em: 10 set. 2020.

SILVA, N. *Maturidade em Dados Abertos: entenda as 5 estrelas*. [S. l.]: Open Knowledge Brasil, 2013. Disponível em: <https://br.okfn.org/2013/01/17/maturidade-em-dados-abertos-entenda-as-5-estrelas/>. Acesso em: 22 mar. 2020.

TENOPIR, C.; DALTON, E. D.; ALLARD, S.; FRAME, M.; PJESIVAC, I.; BIRCH, B.; POLLOCK, D.; DORSETT, K. Changes in data sharing and data reuse practices and perceptions among scientists worldwide. *PLoS ONE [online]*, v. 10, n. 8, 2015. Disponível em: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0134826>. Acesso em: 11 set. 2020.

SANDT, S.; DALLMEIER-TIESSEN, S.; LAVASA, A.; PETRAS, V. The Definition of Reuse. *Data Science Journal*, v. 18, n. 1, p. 22, 2019. Disponível em: <http://doi.org/10.5334/dsj-2019-022>. Acesso em: 23 mar. 2020.

Anexo

Questionário utilizado para a seção 4: Diretrizes de Abertura para Dados Públicos Estaduais e Municipais

Conteúdo mínimo de dados abertos nas esferas municipal e estadual

Este breve questionário tem o objetivo de mapear, juntamente a especialistas em um conjunto de áreas de interesse público, quais dados são considerados prioritários para disponibilização em formato aberto nas esferas estadual e municipal. Ele chegou até você pois consideramos que seu conhecimento e sua experiência podem contribuir para chegarmos a um conteúdo mínimo necessário em cada dimensão de análise para estados e municípios no Brasil.

O questionário é muito rápido e simples, não vai tomar mais do que cinco minutos! Agradecemos desde já sua colaboração. Se possível, compartilhe o questionário com outras pessoas que você considere aptas a respondê-lo.

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. Em outras palavras, dados abertos devem ser completos, primários, atuais, acessíveis, processáveis por máquina, sem discriminação de acesso, disponível em formatos não proprietários e com licenças que permitam o seu reuso.

Trata-se de uma atividade que compõe a execução do Compromisso 2 - Ecossistema de Dados Abertos - do 4º Plano de Governo Aberto do Brasil. O objetivo deste compromisso é consolidar um marco de referência de abertura de dados que contemple não apenas os padrões de abertura e suas características, como também um conteúdo mínimo de dados que cada estado e município deve disponibilizar em formato aberto.

O questionário foi elaborado conjuntamente pelas seguintes instituições: Controladoria-Geral da União (CGU), Diretoria de Análise de Políticas Públicas da Fundação Getulio Vargas (FGV DAPP), Open Knowledge Brasil (OKBr), Centro de Tecnologia da Informação Renato Archer (CTI) do Ministério da Ciência, Tecnologias, Inovações e Comunicações (MCTIC),

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

- Endereço de e-mail: _____ [obrigatório]
- Nome completo: _____
- Filiação institucional: _____
- Escolha uma área de expertise: _____ [obrigatório]
 - Possibilidades: Educação; Saúde; Segurança; Meio ambiente e saneamento; Indicadores sociais; Indicadores econômicos; Orçamento público; Contratos e licitações; Infraestrutura e mobilidade urbana; Legislativo; Judiciário; Dados geográficos; Propriedade e uso da terra; Outro.

Para cada escolha de área de expertise, o entrevistado encontra um preâmbulo: “Dados considerados prioritários para disponibilização em formato aberto na área de _____ por parte de estados e municípios”. E responde as perguntas:

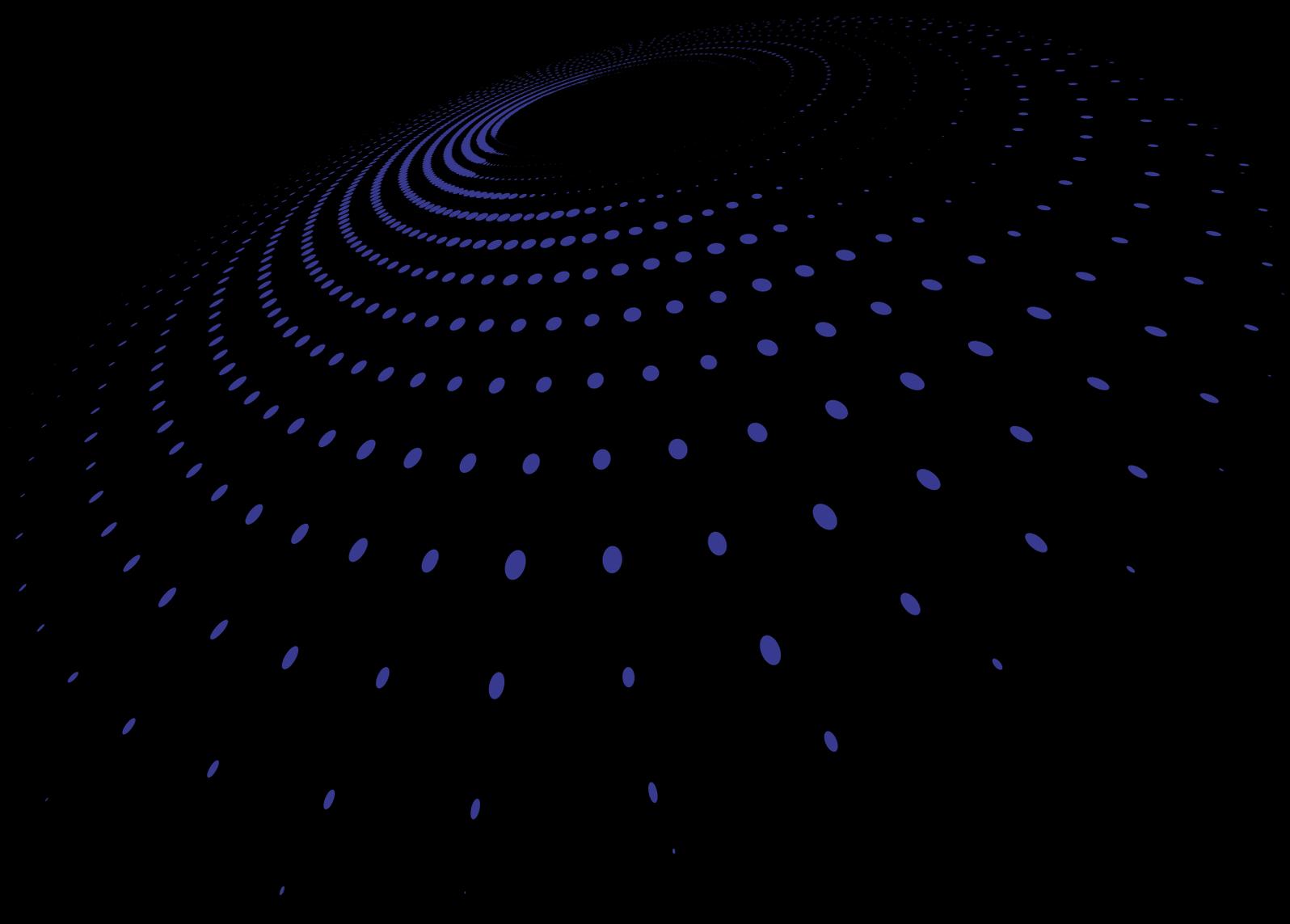
- Do ponto de vista da administração estadual, quais dados você considera essenciais para disponibilização em formato aberto? _____ [obrigatório]
- Do ponto de vista da administração municipal, quais dados você considera essenciais para disponibilização em formato aberto? _____ [obrigatório]

Caso tenha escolhido “outro” na escolha de área:

- Qual a área sugerida? _____ [obrigatório]

Por fim, o entrevistado era convidado a ajudar na divulgação (aplicação da metodologia bola de neve):

- Nos ajude com a divulgação!
 - Indique abaixo um ou mais e-mails de outras pessoas que poderiam estar aptas a preencher este questionário: _____



Em parceria com

CONTROLADORIA-GERAL DA UNIÃO GOVERNO FEDERAL

ceweb.br nic.br



MINISTÉRIO DA
**CIÊNCIA, TECNOLOGIA
E INOVAÇÕES**