

Dados Abertos

Transparência Governamental

Perspectivas, cenários e planejamento

SAS/GDAS

Janeiro/2020

Ao privilégio de participar do novo.

À minha amada Andréa.

*Mais que uma esposa e companheira
inigualável, uma inspiração em todos os
momentos.*

Prefácio

Pensar em tecnologia implica em pensar em meios para melhorar o acesso e a eficiência dos processos desenvolvidos pela sociedade. Seu alcance ultrapassa o horizonte das vantagens pois vez por outra cria demandas que nem pensávamos ter.

É nesse cenário que surge a figura do e-Governo e sua consequência imediata: a e-Democracia. A novidade invade o cenário da participação da sociedade em seu governo ao mesmo tempo em que exige do governo maior eficiência nos serviços ofertados, maior transparência, *accountability*, novas estratégias políticas e processos de governança mais eficientes. Como se não bastasse, há uma elevação no nível da experiência do usuário¹ nas formas de interação entre governos, pessoas e países.

O Brasil tem o mérito de ser o motor deste movimento na América Latina desde 2009 quando iniciou o desenvolvimento de uma política de disseminação de dados governamentais segundo parâmetros e diretrizes dos dados abertos.

Tendo em mente a complexidade deste cenário, tive a iniciativa de escrever este livro tomando como subsídio meu conhecimento e experiências na participação do projeto do Portal de Dados Abertos da Prefeitura de Belo Horizonte.

A cada página apresento e explico elementos que deverão te dar tranquilidade para se aprofundar nos temas aqui tratados. Indico este livro para profissionais de Ciência da Informação, da Computação e demais profissionais que almejam desempenhar funções no seguimento dos dados abertos governamentais.

Boa Leitura!

– *Marcão - Marcus Vinicius Pinto*
Belo Horizonte-Minas Gerais-Brasil

¹ Do inglês *user experience* (UX), é o conjunto de elementos e fatores relativos à interação do usuário com um determinado produto, sistema ou serviço cujo resultado gera uma percepção positiva ou negativa (Wikipedia).

Sumário

1	<u>INTRODUÇÃO</u>	10
2	<u>DADOS GOVERNAMENTAIS ABERTOS X TRANSPARÊNCIA</u>	16
2.1	O CONCEITO DE DGA	18
2.2	PUBLICIDADE X TRANSPARÊNCIA	23
2.3	<i>ACCOUNTABILITY X OPPENNESS</i>	26
2.4	DADOS ABERTOS + TRANSPARÊNCIA = ECOSISTEMA	29
2.5	OPEN KNOWLEDGE BRASIL	32
3	<u>POLÍTICA DE DADOS ABERTOS</u>	33
3.1	INDA	37
4	<u>PLANEJAMENTO DA ABERTURA DE DADOS</u>	43
4.1	OBTER AUTORIZAÇÃO PARA PUBLICAÇÃO	45
4.2	ATRIBUIÇÃO DE LICENÇA DE USO	46
4.3	ELABORAR OS METADADOS	49
4.4	PRIORIZAR DADOS A SEREM ABERTOS	51
4.5	PLANO DE SUSTENTAÇÃO	54
4.6	GOVERNANÇA	55
4.7	USABILIDADE DO PORTAL E DE SEU CONTEÚDO	57
4.8	PROTOCOLO DE ACESSO	61
4.9	COBERTURA DOS CONJUNTOS DE DADOS	62
4.10	API – A DIFÍCIL DECISÃO	62

5	<u>ROTEIRO PARA ELABORAÇÃO DO PLANO DE DADOS ABERTOS</u>	66
5.1	CAPA	68
5.2	INTRODUÇÃO	68
5.2.1	ORIENTAÇÃO	68
5.3	CENÁRIO INSTITUCIONAL	69
5.3.1	ORIENTAÇÃO	69
5.4	OBJETIVOS	70
5.4.1	ORIENTAÇÃO	70
5.5	PROCESSO DE ABERTURA DE DADOS	71
5.5.1	ORIENTAÇÃO	71
5.6	GOVERNANÇA	77
5.6.1	ORIENTAÇÃO	77
5.7	PLANOS DE AÇÃO PARA ABERTURA DE DADOS	80
5.7.1	ORIENTAÇÃO	80
5.8	DIVULGAÇÃO.	83
6	<u>EXEMPLO DE PDA: MINISTÉRIO DA JUSTIÇA- 2017/2018</u>	85
6.1	APRESENTAÇÃO	85
6.2	INTRODUÇÃO	86
6.3	CENÁRIO INSTITUCIONAL	87
6.3.1	INFRAESTRUTURA NACIONAL DE DADOS ESPACIAIS – INDE	89
6.3.2	PLANO DE DADOS ABERTOS 2014-2015	89
6.3.3	LABORATÓRIO DE PARTICIPAÇÃO E INOVAÇÃO – LABPI	91
6.3.4	PLANEJAMENTO ESTRATÉGICO INSTITUCIONAL	91
6.3.5	PLANO ESTRATÉGICO DE TECNOLOGIA DA INFORMAÇÃO E COMUNICAÇÃO – PETIC	91
6.3.6	PLANO DE TRABALHO DA ESTRATÉGIA DE GOVERNANÇA DIGITAL	91
6.3.7	OPEN GOVERNMENT PARTNERSHIP – OGP	92
6.4	OBJETIVOS	92

Dados Abertos: transparência governamental
Perspectivas, cenários e planejamento

6.4.1	OBJETIVO GERAL	92
6.4.2	OBJETIVOS ESPECÍFICOS	92
6.5	DEFINIÇÃO DOS DADOS A SEREM ABERTOS	93
6.6	ESCOPO DO PDA	93
6.7	PRIMEIRA ETAPA PDA/MJ – GRUPO I	94
6.7.1	CATALOGAÇÃO E PRIORIZAÇÃO	95
6.7.2	CRITÉRIOS DE PRIORIZAÇÃO DA ÁREA DE NEGÓCIO	96
6.8	SEGUNDA ETAPA PDA/MJ – GRUPOS II E III	97
6.8.1	CATALOGAÇÃO E PRIORIZAÇÃO	97
6.9	ESTRATÉGIAS PARA ABERTURA	99
6.10	RECOMENDAÇÕES	100
6.11	AVALIAÇÃO DE VIABILIDADE TÉCNICA	101
6.12	METADADOS	102
6.13	GOVERNANÇA	103
6.14	MONITORAMENTO E CONTROLE	104
6.15	PLANO DE AÇÃO	106
6.15.1	PRIMEIRA ETAPA	106
7	<u>PLATAFORMA DE PUBLICAÇÃO – CKAN</u>	<u>107</u>
8	<u>DADOS ABERTOS CONECTADOS</u>	<u>110</u>
9	<u>CONSIDERAÇÕES FINAIS</u>	<u>122</u>
10	<u>LISTA DE SIGLAS E ABREVIATURAS</u>	<u>123</u>
11	<u>GLOSSÁRIO</u>	<u>126</u>
12	<u>PERGUNTAS FREQUENTES</u>	<u>134</u>

13	<u>BIBLIOGRAFIA</u>	<u>157</u>
14	<u>ANEXO 1 – ÁREAS TEMÁTICAS DE ASSUNTO</u>	<u>164</u>
15	<u>ANEXO 2 – EXEMPLOS DE PLANOS DE AÇÃO</u>	<u>168</u>
16	<u>ANEXO 3 – TERMO DE USO DO PORTAL DE DADOS ABERTOS</u>	<u>172</u>

Lista de figuras

Figura 1 – Modelo do relacionamento entre oferta e demanda.	11
Figura 2 - Subdivisões dos dados, Manyika, et al. (2011).....	13
Figura 3 - O que os dados podem produzir.....	13
Figura 4a – Ecossistema de dados abertos – Governo Federal.....	30
Figura 4b – Ecossistema de dados abertos – Governo Federal	31
Figura 5 – Sobre a abertura de dados na Constituição Brasileira.	34
Figura 6 – Sobre a abertura de dados em decretos, portarias, normativas e leis brasileiras.....	35
Figura 7 - Ações realizadas em prol da transparência (ativa e passiva)	36
Figura 8 – Documentos da Política de Dados Abertos	40
Figura 9 – Estrutura do Plano de Dados Abertos (PDA).....	54
Figura 10 – Processo de abertura de dados.....	74
Figura 11 - Distribuição de planos de ação em planos de dados.	82
Figura 12 – Matriz de responsabilidades	83
Figura 13 - Cenário de cinco estrelas de Timothy John Berners-Lee.....	114
Figura 14 – <i>LOD Cloud</i> (imagem de 1970).....	115
Exemplo de janela de abertura de arquivo CSV na ferramenta LibreOffice.....	137

Lista de tabelas

Tabela 1 – Principais demandas de informações governamentais (Ávila, 2015). ...	12
Tabela 2: Classificação de sigilo. Adaptado de BRASIL (2011).	22
Tabela 3 – Marcos do desenvolvimento do Portal Brasileiro de Dados Abertos em 2011.	41
Tabela 4 – Marcos do desenvolvimento do Portal Brasileiro de Dados Abertos em 2012.	42
Tabela 5 - Metadados obrigatórios.	50
Tabela 6 - Metadados desejáveis.	50
Tabela 7- Entidades e elementos do Núcleo de Metadados do Perfil MGB Sumarizado.	73
Tabela 8 - Comparativo entre de dados conectados e outros formatos de dados estruturados.	113
Tabela 9 – Recomendações para publicação de dados abertos governamentais conectados.	120

1 Introdução

*O melhor uso que poderá ser feito com seus dados...
certamente será desenvolvido por outros e não por você.*
Tim Berners-Lee

O Estado Democrático tem entre suas principais características o respeito ao direito do cidadão de conhecer as informações públicas sob sua tutela. É inimaginável a quantidade de informação existente nos ambientes governamentais e, com certeza, este acervo tem um enorme potencial para a sociedade, pois indivíduos e organizações têm a possibilidade de utilizá-los para aprimorar seus serviços, para o projeto de novos produtos e para aprimorar a participação da sociedade nas decisões de governo.

Vários autores se dedicam a discutir e elucidar o valor das informações governamentais como fonte para linhas de atuação estratégicas, planos de trabalho, planejamento de cenários futuros e mudanças no mercado. Dentre eles é possível citar Helbig et al. (2012) que tratam, inclusive, da motivação inicial necessária ao despertar do interesse da sociedade por esta informação e Barreto (2000) que considera que o valor da informação depende diretamente das competências pessoais do consumidor, do nível educacional da sociedade e da força de trabalho como um todo.

A demanda por informação no Brasil pode ser visualizada no modelo proposto por Barreto (2000) em que a mesma é classificada em:

- A. Demanda por informação: configura uma demanda orientada à reflexão, para a reelaboração da informação recebida, da informação de sustentação e do apoio à ciência e tecnologia. É o segmento do mercado de informação com maior escassez, mas é o que concentra o maior esforço de preparação e distribuição da oferta de informação.
- B. Demanda Mantenedora: que pode ser entendida como a informação responsável pela manutenção do indivíduo em seu status quo profissional e social.

C. Demanda Utilitária: a demanda por informação para as transações correntes do indivíduo ao exercer sua cidadania. Percebe-se que nos dias de hoje a oferta de informação é pequena e não cobre a demanda potencial.

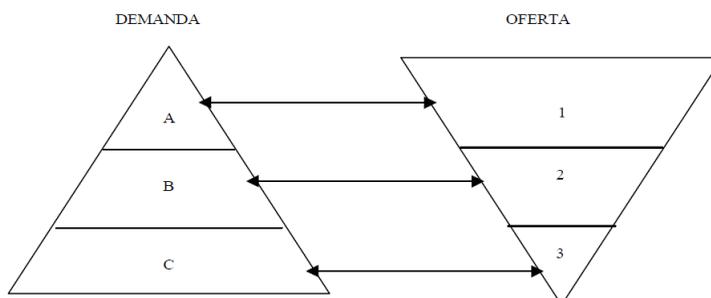


Figura 1 – Modelo do relacionamento entre oferta e demanda.

Na Tabela 1 temos o relacionamento entre alguns tipos de informações públicas e as principais demandas e aplicações em diversos segmentos da sociedade (Ávila, 2015). As instituições governamentais são identificadas como grandes criadores de dados em diversos segmentos, tais como informações de saúde, financeiras, turísticas, geográficas e de segurança. Neste contexto de mercado de demanda e oferta, tendo como *locus* a Internet, podem ser analisadas, concordando com Manyika (2011) e Ubaldi (2012 e apresentado na Figura 2 , em 4 categorias:

- *Big Data* são bases de dados que extrapolam os limites da capacidade de processamento dos sistemas de informação hoje disponíveis. O volume de dados é da ordem de peta bytes.
- Dados abertos são os que estão disponíveis ao cidadão, sem custo ou restrições de qualquer ordem proveniente de qualquer origem, seja privado ou público.
- Dados abertos governamentais são os que instituições governamentais disponibilizam ao cidadão.
- Dados pessoais são os privativos de cada pessoa, tais como imposto de renda, endereço, consumo.

Observe-se que dados abertos podem produzir outros dados abertos, informações, interfaces de interpretação, fatos e serviços (DAVIES, 2010). A Figura 3 ilustra este encadeamento.

Dados Abertos: transparência governamental
 Perspectivas, cenários e planejamento

Segmento	Tipo de Informação Pública	Finalidade
Setor Produtivo	Indicadores Sociais, Econômicos, Demográficos, Planos de Governo, Relatórios Fiscais, Informações Geográficas (imagens aéreas, vetores com distâncias entre localidades, mapas e cartogramas sobre dados socioeconômicos), etc.	Projetos de Consultoria; Expansão e/ou Manutenção de Negócios; Desenvolvimento ou aprimoramento de produtos e serviços
Setor Acadêmico	Indicadores Sociais, Econômicos, Demográficos, Planos de Governo, Relatórios Fiscais, Informações Geográficas (imagens aéreas, vetores com distâncias entre localidades, mapas e cartogramas sobre dados socioeconômicos), etc.	Artigos Científicos; Trabalhos Acadêmicos; Projetos de Pesquisa; Monografias; Dissertações; Teses; Projetos para captação de recursos em instituições de fomento
Setor Público	Indicadores Sociais, Econômicos, Demográficos, Planos de Governo, Relatórios Fiscais, Informações Geográficas (imagens aéreas, vetores com distâncias entre localidades, mapas e cartogramas sobre dados socioeconômicos), Pesquisas acadêmicas, estudos e análises, relatórios de tendência, projeções de cenários.	Diagnósticos governamentais, diagnósticos sobre áreas ou demandas específicas (ex: problemas ambientais); Formulação de planos e programas de governo, execução de ações, monitoramento e avaliação governamental; Publicidade de ações governamentais; Projetos para captação de recursos em instituições de fomento
Imprensa	Dados orçamentários e financeiros; Pesquisas e indicadores socioeconômicos; Dados Populacionais; Relatórios de Monitoramento e Acompanhamento de Ações Governamentais	Matérias e investigações jornalísticas; Publicidade de ações governamentais; Denúncias de não-conformidades em ações governamentais
Sociedade em Geral	Dados orçamentários e financeiros; Pesquisas e indicadores socioeconômicos; Dados Populacionais	Monitoramento e Controle Social do Governo; Elaboração de Projetos para captação de recursos

Tabela 1 – Principais demandas de informações governamentais (Ávila, 2015).

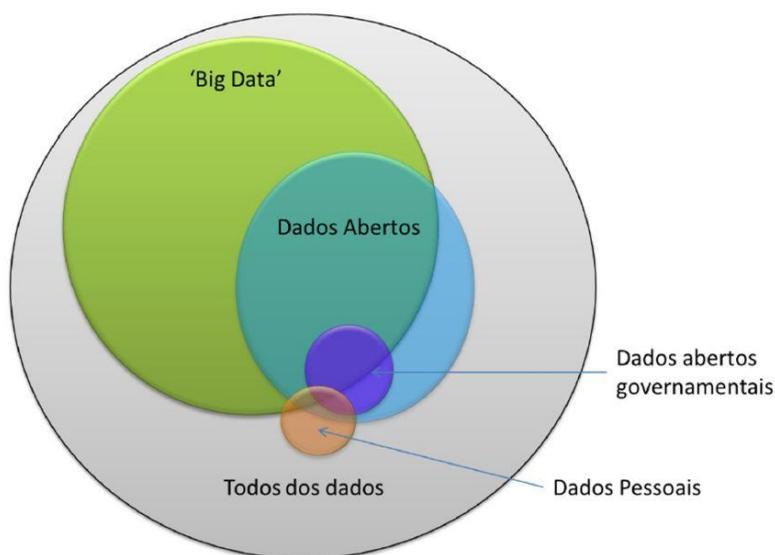


Figura 2 - Subdivisões dos dados, Manyika, et al. (2011)

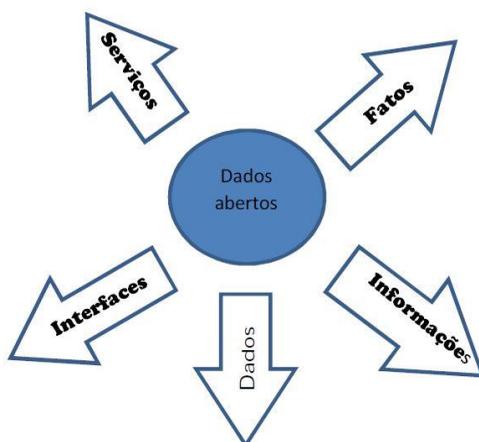


Figura 3 - O que os dados podem produzir.

Dados produzem fatos quando indivíduos buscando em fontes de dados abertas alimentam ações do segmento social que geram resultados no planejamento econômico de empresas, entidades ou indivíduos. Dados dão origem a informações que resultam do cruzamento de bancos de dados que, por seu turno, produzem

informações tabulares, infográficos e relatórios.

Dados geram interfaces de interpretação ao fornecerem meios de interatividade entre um ou mais bancos de dados, tais como: mapas interativos, links com outros dados. Dados produzem outros dados quando são processados e originam outras fontes de dados. Dados geram monitoramento da qualidade de serviços quando os dados abertos são o suporte do fornecimento de serviços online tais como a identificação de anomalias de serviços públicos pela população e comunicação desta para as autoridades.

Não faltam iniciativas bem-sucedidas no mundo e no Brasil com a utilização de dados abertos. Um bom exemplo é o *Citymapper*, aplicativo gratuito para smartphones criado em 2011 em Londres por um ex-funcionário do Google com o objetivo de melhorar a relação da população com o transporte público. O aplicativo trabalha com dados sobre ônibus, trens e metrô, publicados em tempo real pela autoridade de trânsito da cidade, para mostrar o melhor tempo de viagem entre um ponto e outro e quando está prevista a chegada do próximo ônibus. As informações são atualizadas em tempo real e mostram as condições de trânsito, clima e problemas técnicos nos veículos.

A partir do êxito na prestação deste serviço ele foi implantado em Nova York, Paris, Berlim, Washington DC, Madri, Boston, Barcelona, São Francisco, Chicago, Milão, Roma, Cidade do México e em São Paulo. O aplicativo pode trabalhar com qualquer cidade que disponha de dados sobre serviços de transporte público em tempo real.

A cadeia produtiva da informação do setor público tem seu arcabouço legal definido em leis e regulamentos que estabelecem a Política de Informação. Esta política engloba os processos de geração, processamento, armazenamento, transporte, distribuição, busca/recuperação e descarte.

Nas últimas décadas ocorreu uma mudança nesta cadeia produtiva. O enfoque do governo eletrônico deixou de ser a disponibilização de funcionalidades que

aprimoram a prestação de serviços do governo e passou a ser a facilitação da reforma administrativa e a promoção da participação democrática (YILDIZ, 2007)

A abertura dos dados traz benefícios para o cidadão, mas também gera benefícios para o governo, tais como:

- Eficiência nos gastos com transparência passiva: como a publicação de dados públicos atende a várias solicitações de informação que antes demandavam seu tratamento individual isto gera economia nas respostas.
- Redução de gastos com análises sobre os dados do governo: organizações que necessitam tratar dados públicos têm a possibilidade de reduzir gastos com consultorias especializadas.
- Ampliação do mercado de trabalho: as oportunidades de negócio impulsionam a criação ou ampliação de empresas que trabalham com dados da máquina pública.

Face à diversidade das cadeias de produção a Política de Informação é compreendida em segmentos temáticos ou funcionais. Um exemplo atual é o caso da União Europeia em que a aprovação da Diretiva 2003/98/CE do Parlamento Europeu se tornou um forte incentivo ao *Open Government Data*. A diretiva propõe como os cidadãos dos países signatários da União devem proceder para reutilizar os dados governamentais na Europa e deu origem ao *Open Data Reuse*. O Plano de Ação Europeu 2011-2015 foi um dos desdobramentos das iniciativas dos governos.

O Brasil é membro cofundador da Parceria para Governo Aberto e em 2019 obteve a 14^a colocação², em um ranking composto por 122 países e lugares participantes do ranking organizado pela Open Knowledge Internacional, dentre os mais avançados em relação aos dados abertos governamentais.

² <https://www.governodigital.gov.br/noticias/brasil-fica-em-12o-lugar-em-ranking-mundial-de-dados-abertos>. Acessado em 22 nov. 2019.

2 Dados Governamentais Abertos X Transparência

“Transparência significa deixar o Estado, no seu sentido mais amplo, aberto e visível ao cidadão”.

Marcus Vinicius Pinto

Os Dados Governamentais Abertos (DGA) foram utilizados, em um primeiro momento, para cumprir agendas de transparência e controle governamental. Seu potencial como insumo para a criação de produtos e serviços para a sociedade e para o próprio governo tem sido gradativamente reconhecido a ponto de impulsionar a eficiência do próprio setor público. Como exemplo, a série histórica de ocorrências no trânsito³ pode gerar indicadores para compor o cálculo da taxa de risco e da probabilidade de sinistro no trânsito.

No Brasil, os movimentos de governo aberto surgiram a partir de 2005 e se tornaram prioridade após a Lei de Acesso à Informação Pública (LAI), Lei 12.527/2011. Esta Lei regulamenta o acesso a dados e informações detidas pelo governo e incentivou iniciativas de planos e políticas de informação em todo o país por ser um arcabouço jurídico para a obrigação governamental de disponibilizar dados abertos no país. Mesmo que apenas para cumprir obrigações de disponibilização de informação foi possível notar que o cenário havia mudado.

Para um melhor diagnóstico sobre a evolução da abertura de dados dos governos é necessário ampliar o horizonte de análise considerando o contexto em escala mundial. Pesquisando-se os portais de dados abertos disponíveis na internet depara-se com portais de governos nacionais e subnacionais que já disponibilizam seus dados, tais como Estados Unidos, Reino Unido, Austrália e Nova Zelândia (AGUNE, GREGORIO FILHO, BOLLIGER, 2010). Dentre estes países o caso mais representativo é o norte-americano em que há iniciativas tanto no nível federal,

³ Dados abertos do DNIT sobre ocorrências no trânsito estão disponíveis em <http://www.dnit.gov.br/rodovias/operacoes-rodoviaras/estatisticas-de-acidentes>. Último acesso em 21 set. 2019.

quanto nos estaduais.

Segundo Silva (2010), a capital dos Estados Unidos da América, Washington DC, foi bastante vanguardista ao incentivar ao trabalho colaborativo entre governo e sociedade fazendo uso de dados abertos. O presidente Barack Obama, ainda no início de seu primeiro mandato, implementou políticas públicas objetivando a promoção da transparência da máquina pública.

Cabe destacar um memorando enviado pelo presidente, em 21 de janeiro de 2009, aos chefes de governo federal, conhecido como *Transparency and Open Government*, em que se comprometia a criar “níveis sem precedentes de abertura” no governo. Neste memorando o presidente conclamava o primeiro escalão a agir de modo a assegurar a confiança da sociedade no governo, e estabelecendo um sistema de transparência, participação pública e colaboração.

A política *Open Government Initiative* veio assegurar a efetividade deste memorando e teve como uma consequência a implantação do portal *Data.gov* em maio de 2009 com o objetivo de ofertar informações e ampliar o uso criativo dos dados governamentais através de aplicações web. O portal disponibiliza informações de uma miríade de agências e instâncias governamentais organizadas por órgão e estatísticas de visitantes mensais, diários, anuais.

Configurando uma ampliação do rol de iniciativas governamentais há organismos não governamentais atuando na promoção e no apoio à adoção de políticas de dados abertos. O projeto *Open Government Data*, fruto da parceria entre a *Access Info* e a Fundação *Open Knowledge*, e a Fundação *Sunlight* são exemplos desta atuação. Essas iniciativas ampliam a discussão em torno de DGAs e agregam valor ao contexto dos dados abertos com pesquisas e incentivo à criação de soluções e padrões para formatos abertos.

2.1 O conceito de DGA

Por ser um termo ainda muito recente, a definição de DGA ainda não é consenso na literatura. A *Open Definition* (2016) conceitua dados abertos como “Dados que podem ser utilizados livremente, reutilizados, redistribuídos por qualquer pessoa ou organização, com a exigência de atribuição à fonte original e ao compartilhamento pelas licenças em que a informação foi apresentada”.

Para a *Open Knowledge Foundation* (2010) para se considerar determinado conjunto de dados como dados abertos é necessário que o conjunto cumpra três regras fundamentais:

- Disponibilidade de acesso. Se refere aos dados estarem disponíveis na Internet e sem necessidade de software proprietário para sua leitura.
- Reuso e redistribuição. Os dados devem estar livres para reuso e redistribuição sem restrições de divulgação.
- Participação universal. Os dados estão liberados para quaisquer áreas de atuação do usuário. Seja ele uma organização comercial, governamental, grupo de pessoas ou um usuário individual.

O ativista David Eaves (2009) enriquece estes conceitos incorporando os seguintes princípios:

- Se os dados não podem ser encontrados ou indexados eles não existem.
- Se não estiverem disponíveis para acesso livre e em formato legível de máquina eles não podem ser considerados abertos.
- Se há alguma restrição legal ao reaproveitamento dos dados eles também não podem ser considerados abertos.

Para o *OpenGovData*, um grupo de especialistas, há oito princípios para se considerar um conjunto de dados abertos como DGA (OPEN GOV DATA, 2007).

Estes princípios estabelecem que os dados devem ser:

- Completos. Todos os dados públicos estão disponíveis. Dado público é o dado que não está sujeito a limitações válidas de privacidade, segurança ou controle de acesso.

- Primários. Os dados são apresentados tais como os coletados na fonte, com o maior nível possível de granularidade e sem agregação ou modificação.
- Atuais. Os dados são disponibilizados tão rapidamente quanto necessário à preservação do seu valor.
- Acessíveis. Os dados são disponibilizados para o maior alcance possível de usuários e para o maior conjunto possível de finalidades.
- Compreensíveis por máquinas. Os dados são razoavelmente estruturados de modo a possibilitar processamento automatizado.
- Não discriminatórios. Os dados são disponíveis para todos, sem exigência de requerimento ou cadastro.
- Não proprietários. Os dados são disponíveis em formato sobre o qual nenhuma entidade detenha controle exclusivo.
- Livres de licenças. Os dados não estão sujeitos a nenhuma restrição de direito autoral, patente, propriedade intelectual ou segredo industrial. Restrições sensatas relacionadas à privacidade, segurança e privilégios de acesso são permitidas.

Diniz (2010) enriquece o contexto dos DGAs argumentando que são três os agentes da cadeia de valor destes dados, que de maneira colaborativa e estruturada produzem valor para a sociedade. A saber:

1. O órgão público. Responsável pela coleta, organização e armazenamento digital dos dados. Sua importância não advém somente por exercer a figura de guardião das informações mas por ser detentor da autoridade política para decidir quais, como e quando os dados serão liberados ao público.
2. O provedor de tecnologia. Sem tecnologia não há DGAs. Este agente provê a capacidade do dado de ser reutilizado por outros.
3. O agente reutilizador dos dados. Os DGAs só têm valor se houver interesse da sociedade em reutilizá-los. Pode ser qualquer pessoa ou instituição que tenha interesse neles e tenha competência para desenvolver aplicações baseadas na web que façam uso desses dados.

O W3C Escritório Brasil, por seu turno, sugere a qualificação dos três agentes da cadeia de valor com as seguintes responsabilidades:

- Do órgão público: definição dos dados que serão abertos, definição de equipe técnica que receberá capacitação técnica e fará as codificações necessárias, instalação e disponibilização de plataforma de hardware e software necessários ao desenvolvimento do projeto, especificação do processo de abertura dos dados, desenvolvimento e implementação das aplicações necessárias e definição das regras e regulamentação do acesso aos dados.
- Do provedor de tecnologia: orientação e auditoria quanto à aderência aos princípios dos dados abertos, capacitação técnica nas tecnologias necessárias à publicação de dados abertos e articulação de parceiros que oferecerão consultoria técnico-científica ao projeto
- Do agente reutilizador: aderência às regras e regulamentação do acesso aos dados estabelecidas, desenvolvimento de aplicativos que reutilizem estes dados e compartilhamento da solução implementada.

Os benefícios advindos dos DGAs em relação ao controle e participação social e à transparência são óbvios, ao menos em teoria. O cenário da sociedade com acesso a DGAs tende a apresentar um aumento na transparência do governo, criando melhores possibilidades de controle por parte da sociedade sobre as ações governamentais mesmo sendo este um tema novo. O acesso a DGAs também tem o potencial de viabilizar a criação de novas informações e aplicativos a partir destes dados.

Para Diniz (2010): a disponibilização de dados governamentais abertos permite que as informações sejam utilizadas da maneira e conveniência do interessado de tal forma que elas possam ser misturadas e combinadas para agregar mais valor aos dados. O objetivo das informações públicas serem disponibilizadas segundo as regras dos dados abertos é, ainda para o autor, superar as limitações existentes para que usuários de informações do serviço público possam facilmente encontrar, acessar, entender e utilizar os dados públicos segundo os seus interesses e conveniências.

Apesar da variedade de conceitos é consenso que os dados para serem considerados como Dados Governamentais Abertos precisam estar publicados na

internet, disponíveis para qualquer usuário, livres de licença, livres de restrições de direito autoral, propriedade intelectual ou segredo industrial.

A revisão da literatura sobre DGA permite organizar os conceitos conforme segue:

- Albano e Araújo (2013). É uma solução adotada pelos governos para cumprir exigências de acesso por parte da sociedade. Estas exigências se sustentam em conceitos de transparência, colaboração e participação.
- Barbalho (2014). São os dados gerados nos processos governamentais que passam a ser liberados para acesso pela sociedade.
- Yannoukakou e Araka (2014). Dados produzidos ou encomendados pelo governo ou entidades controladas pelo governo disponíveis livremente para uso e redistribuição.
- Solar et al. (2014) É uma filosofia de trabalho que busca capacitar os cidadãos a acessar os dados produzidos por entidades públicas permitindo acessar, armazenar em bases de dados próprias, redistribuir e integrá-los com outras fontes de dados.
- Kalampokis et al. (2011). São dados do setor público disponíveis livremente em formatos abertos.
- W3C (2016). Refere-se à publicação e disseminação dos dados do setor público na Web, em formato bruto e aberto, compreensíveis logicamente, permitindo sua reutilização em aplicações digitais desenvolvidas pela sociedade.

Um caso a ser analisado em separado neste contexto de abertura de dados é o dos dados sigilosos, que não podem ou não devem ser abertos (BRASIL, 2011, Brasil, 2012b). No Brasil, há legislação específica para este tipo de informação. O inciso III do artigo 4º da Lei de Acesso à Informação - LAI, lei 12.527 de 18 de novembro de 2011, define informação sigilosa como “aquela submetida temporariamente à restrição de acesso público em razão de sua imprescindibilidade para a segurança da sociedade e do Estado.” Adicionalmente, é possível definir o grau de sigilo em que se considera o interesse público na informação para que seja possível aplicar critérios menos restritivos. A LAI trata de outras hipóteses legais de sigilo, tais como segredo de justiça e segredo industrial. A classificação dos graus de sigilo e de seus prazos

são definidos na LAI e estão sumarizados na Tabela 2.

Grau de classificação	Prazos máximos	Quem pode classificar
Ultrassegredo	25 anos, podendo ser prorrogado uma única vez por igual período, apenas pela Comissão Mista de Reavaliação de Informações.	a) Presidente da República b) Vice-presidente da República c) Ministros de Estado e autoridades com as mesmas prerrogativas d) Comandantes da Marinha, do Exército e da Aeronáutica e) Chefes de Missões Diplomáticas e Consulares permanentes no exterior
Secreto	15 anos, sem possibilidade de prorrogação.	Das autoridades que podem classificar informações em grau ultrassegredo, além dos titulares de autarquias, fundações ou empresas públicas e sociedades de economia mista.
Reservado	5 anos, sem possibilidade de prorrogação.	Das autoridades que podem classificar informações em grau secreto e ultrassegredo e daquelas que exerçam funções de direção, comando ou chefia, nível DAS 101.5, ou superior, ou de hierarquia equivalente, de acordo com regulamentação específica de cada órgão ou entidade, observado o disposto na Lei nº 12.527.

Tabela 2: Classificação de sigilo. Adaptado de BRASIL (2011).

A criação do Portal Brasileiro de Dados Abertos, <http://dados.gov.br>, com o objetivo de ser a referência central em DGA para o Brasil, possibilitou a cada órgão ou entidade desenvolver seu catálogo de metadados próprio, em que os dados são classificados/agrupados conforme os assuntos presentes no Vocabulário Controlado de Governo Eletrônico (VCGE) (Brasil, 2016).

Polêmicas em torno da classificação da informação e seu grau de sigilo permanecem em pauta e é inevitável que sejam dependentes dos sistemas de classificação adotados por cada país que tenha seus dados abertos, pois a classificação depende de questões cognitivas dos classificadores, fatores culturais e de segurança.

2.2 Publicidade x Transparência

A Constituição Federal, em seu artigo 37, define os princípios da Administração Pública. No caso dos dados abertos o princípio da publicidade é o que tem mais interseção com o tema. Este princípio estabelece que todos os atos da Administração Pública devem ter divulgação ampla com algumas exceções. Em seu artigo 5º, temos os incisos que tratam da abertura de dados:

- XIV. Do acesso à informação e resguardo do sigilo profissional;
- XXXIII. Do direito à informação particular ou geral existentes nos órgãos públicos ou de caráter público;
- XXXIV. Do direito à certidão;
- LXXII. Do habeas data.

Para Martins Junior (2004) os principais objetivos da transparência são a busca de legitimidade e eficiência dos atos da Administração Pública e a redução do distanciamento com a sociedade. Os conceitos relativos a publicidade e transparência não convergem para um mesmo conceito quando considerados no cenário das tecnologias da informação e comunicação, mas são imprescindíveis para que a informação governamental seja de domínio público e para criar instrumentos de sanção aos agentes públicos.

MELLO (2008), ARAUJO (2006) e DI PIETRO (2006) são alguns dos autores que consideram que o princípio da publicidade pode ser entendido como sinônimo de transparência, pois é comum associar a transparência aos processos de divulgação e acesso às informações da Administração Pública.

A publicidade, segundo ARAUJO (2006), pode ser:

- Geral. É a que está presente na publicação em órgãos oficiais. Em alguns casos ela é obrigatória e deve ser extensa. Um exemplo é o caso dos concursos públicos.
- Restrita. Refere-se às informações de cunho pessoal das pessoas interessadas. Temos como exemplo as comunicações via notificação, citação e registro dos atos em livros administrativos.

E ainda, há três níveis de abertura e transparência dos dados:

- Primeiro nível. Páginas estáticas em sites do governo com relatórios financeiros das finanças públicas que podem ser acessadas em formato estático.
- Segundo nível. Informação disponível em portais mais robustos com acesso a bancos de dados e filtros de pesquisa. Neste nível não é possível se fazer cruzamento entre dados.
- Terceiro nível. Este nível é o único considerado com DGA. Portais com sistemas de consultas e filtros acessando grandes bases de dados em formato aberto. Neste cenário é possível se fazer buscas extensas nos resultados liberados sobre as ações do governo.

Como exemplos de informações que potencializam a transparência das contas públicas e rastreabilidade dos recursos temos os seguintes:

- Orçamento detalhado. É um dado que pode ser utilizado para se validar como os recursos públicos serão aplicados. Neste caso é essencial se ter acesso ao detalhamento, pois o dado fornecido de maneira agrupada dificulta a auditoria do que foi planejado em relação ao que foi aprovado.
- Gastos públicos detalhados. Permitem a comparação do que foi efetivamente gasto com o que foi planejado.
- Resultados eleitorais. Permitem a validação do desempenho de cada candidato em cada região de um município, estado ou do país.

A transparência, amparada por leis, advinda dos dados abertos promove *accountability*⁴, pois fornece ao cidadão informações sobre as ações do governo. O cidadão, de posse destas informações, pode se qualificar para participar das decisões de governo, contribuir com ideias e conhecimentos.

Esse cenário predispõe o governo a ser mais eficaz e eficiente. Mas, apesar de todo este potencial há fatores negativos relacionados ao tema. Zuiderwijk e Janssen (2014b) argumentam que a abertura de dados governamentais pode gerar os seguintes

⁴ Prestação de contas e à definição dos objetos sobre as quais se prestarão contas juntamente com a sua responsabilização, vem juntamente com a transparência das ações que gera confiança entre os governantes e a sociedade.

riscos:

- Violação de privacidade não intencional. Descaracterizar dados para preservar a privacidade de quem o dado se refere e para relativizar dados sensíveis pode demandar muito esforço e investimento.
- Publicação contaminada ou tendenciosa dos dados. Os dados publicados podem ser contaminados por intenções do gestor ou por motivos tendenciosos. Isto faz com que apenas sejam considerados passíveis de abertura dados pouco complexos, pouco sensíveis, que não se perceba possibilidade de gerar danos ou culpa para funcionários públicos.
- Uso e interpretação deturpada. Dados complexos geram risco de má interpretação quando disponibilizados ao público leigo. O uso deturpado pode prejudicar a imagem do provedor dos dados.
- Possibilidade de infringir uma legislação ao abrir os dados. Certos conjuntos de dados não devem ser abertos por razões legais. Informações sensíveis e alguns dados científicos podem causar danos se abertos.
- Respeito à propriedade dos dados. Dados oriundos de parecerias ou cedidos por outras instituições ou Organizações não Governamentais (ONGs) não podem ser considerados como apenas do setor público e estão fora do contexto dos dados abertos.
- Publicação de dados tardios. O tema dados abertos gravita em torno da infraestrutura e seus portais e foca na divulgação de dados e não em incentivar os processos das organizações para abrir seus dados.
- Dados abertos podem gerar consequências negativas. Todo dado aberto tem potencial de fragilizar o governo através de análise de desempenho e de qualidade do serviço tratado no dado.
- Outros lucros com dados abertos. Quem lucra ou é beneficiado com a abertura dos dados não é o cidadão comum, pois ele não possui competência para lidar com esses dados.
- Mensuração do valor dos dados abertos. Não há métricas, parâmetros ou instâncias nas organizações para determinar quais conjuntos de dados têm algum valor.
- Falta de prioridade para a publicação de dados. A publicação dos dados ainda não é uma prioridade para as organizações, pois este procedimento gera apenas custo.

- Informação limitada sobre a política de abertura de dados. As políticas adotadas pelas organizações para nortear a publicação de dados não são publicadas.
- Decisões baseadas em informação de má qualidade. É comum não se ter meios para avaliar a qualidade dos dados gerando riscos na tomada de decisões baseadas em dados de má qualidade.
- Proibição da abertura de dados recentes. Isto faz com que dados sejam publicados apenas após se tornarem irrelevantes para o governo.
- Consequências negativas da transparência. Análise das falhas do governo em dados publicados podem diminuir a confiança no governo e gerar ações diversas das originalmente propostas.
- Responsabilidade. Ainda não está totalmente definido quem é o responsável pela utilização indevida, interpretação contaminada ou baixa qualidade dos dados.

2.3 Accountability x Oppenness

Proximidade e interdependência são conceitos próximos aos conceitos de transparência e *accountability*, sendo que diversos autores definem transparência a partir do conceito de *accountability*. Prado e Loureiro (2006) argumentam que na literatura frequentemente a transparência é associada automaticamente à prestação de contas e *accountability*. Pinho (2008) é um dos autores que aglutina os termos transparência e *accountability*, entendendo como “a transparência, o engajamento dos governantes com a prestação de contas, e também a responsabilização dos governantes pelos seus atos.

De acordo com O’Donnell (1998 apud Angélico, 2012, p. 12), existem duas formas de *accountability*:

- Horizontal: quando uma instituição pública presta contas.
- Vertical: quando a União presta contas à sociedade.

Nas duas formas, o acesso à informação exerce papel fundamental para a fiscalização e a responsabilização de agentes e órgãos públicos. Akutsu (2005), por seu turno, tratando destes termos, afirma que é a indispensável transparência das

decisões e ações tomadas pelos agentes públicos quando se fala em *accountability*. E vai além afirmando que não é suficiente o mero cumprimento de formalidades burocráticas, é necessário demonstrar que a administração agiu com economia, eficiência e honestidade.

Outro que trata do tema é Campos (1990) e argumenta que a “*accountability* não é apenas uma questão de desenvolvimento organizacional ou de reforma administrativa. A simples criação de mecanismos de controle burocrático não tem se mostrado suficiente para tornar efetiva a responsabilidade dos servidores públicos”.

Apesar de ainda não ter sido traduzido para o português e ainda não ter uma conceituação definitiva o termo *accountability* é ainda um tema em debate. Para Prado e Loureiro (2006) ele pode ser considerado como “um princípio que se expressa em mecanismos institucionais que devem constranger os governos continuamente a prestar contas à sociedade e a outros atores no sistema político”. Para estes autores, a transparência das informações públicas é uma das etapas da *accountability* juntamente com a prestação de contas dos governantes e sua responsabilização.

Apesar da grande semelhança entre os conceitos de *accountability* e transparência, o conceito de um não é sinônimo do outro. O conceito de *accountability* apresenta um espectro mais amplo abrangendo prestação de contas, publicidade das ações dos governos e uma atuação de controle sobre a Administração Pública trazendo como resultado incentivos ou sanções para o agente público caso cumpra ou descumpra determinada obrigação.

Um outro viés proposto para o termo transparência, proposto por Laporte, Demchak e Friis (2000), é outro anglicismo: *openness*. O termo pode ser entendido como abertura para o fornecimento de informação livre e universal para seu público-alvo (VAZ, 2002, 2003). Deste termo surgem duas categorias básicas:

- Transparência. A informação sobre a instituição publicada pela própria instituição.

- Interatividade. A facilidade com que as informações publicadas podem ser acessadas pelos cidadãos no site governamental.

O governo *openness* assume que a informação pública pertence ao público e assegura o direito do cidadão à informação. Para tanto, o governo deve implementar meios para que o cidadão possa ter acesso às informações de forma clara e a mais facilitada possível. O cidadão, por seu turno, deve procurar conhecer estruturas e decisões do governo de modo a participar deste. Para ser *openness* se espera que o governo vá além de promover a transparência pela disponibilidade das informações. Ele deve proporcionar meios para que estas informações cheguem aos cidadãos e sejam por eles compreendidas. Um exemplo é o caso da prestação de contas do sistema de saúde, o SUS. O orçamento e a prestação de contas podem estar publicados na íntegra no Portal eletrônico do governo federal gerando possibilidade de consulta por qualquer cidadão. Entretanto, é possível afirmar que o cidadão que não atua no segmento de saúde não conseguirá entender os termos técnicos presentes e os motivos das ações. Apesar disto, segundo critérios de transparência, este agente público cumpre o princípio da transparência com louvor, mas não tem sucesso em gerar compreensão pelos cidadãos.

O Poder Público, mesmo que disponibilize seus dados em um portal, não será efetivamente transparente se as informações forem difíceis de encontrar, não estiverem atualizadas e não forem compreensíveis.

Apesar de não haver métricas ou fórmulas para se avaliar o quanto uma administração é *openness* percebe-se como tal uma administração que publica seus dados, os mantém atualizados e cria facilidades para que o cidadão tenha acesso (LAPORTE, DEMCHAK, FRIIS, 2000).

Ao fim e ao cabo, a abertura de dados deixa de ser apenas uma iniciativa de transparência, e se transforma também em fomento de novos serviços oriundos da interação entre sociedade e governo. Esses serviços podem se dar por novas formas de atuação participativa e colaborativa entre governo e instituições privadas. Os

dados abertos permitem superar a visão do cidadão como simples receptor da informação pública ao permitir que qualquer interessado possa, ao processar livremente os dados governamentais, criar conteúdo a partir da reutilização dos dados.

2.4 Dados Abertos + Transparência = Ecosistema

O Portal da Transparência do Distrito Federal (www.transparencia.df.gov.br) atende determinação constitucional e legal (Lei Complementar nº 131, de 27 de maio de 2009) de dar publicidade a informações sobre a execução orçamentária e financeira dos órgãos e entidades públicas do Poder Executivo Distrital.

Administrado pela Controladoria-Geral do Distrito Federal, o Portal da Transparência é uma ferramenta de participação da sociedade no controle das ações do Governo. Nele estão disponíveis informações sobre os gastos do Governo, receitas públicas arrecadadas, patrimônio público, licitações, contratos, convênios, remunerações dos servidores públicos, entre outras.

Com escopo mais amplo, o Portal de Dados Abertos é o ponto de referência para a busca e o acesso a dados públicos sobre todo e qualquer assunto ou categoria de interesse da população, como saúde, educação, segurança pública, economia, mobilidade urbana, assistência social, esportes, turismo, energia, meio ambiente etc. Ele simplifica o acesso aos dados das mais variadas naturezas, não só orçamentária e financeira, organizando e padronizando a produção e publicação das informações, incentivando a reutilização pela sociedade, por organizações e pelos demais órgãos públicos.

Assim, os portais de transparência e de dados abertos dão origem a um conjunto composto por políticas, portais, padrões, vocabulários, prioridades, tecnologias, instituições públicas e categorias da sociedade em geral. Esse conjunto de elementos interligados pode ser nomeado como Ecosistema de Dados Abertos (GAMA; LÓSCIO, 2014). Essa noção advém do conceito de ecossistema de software proposto por Jansen *et al.* (JANSSEN, CHARALABIDIS, ZUIDERWIJK, 2009): um

conjunto de empresas e produtos funcionando em uma mesma plataforma e compartilhando o mercado de software e serviços entre si.

A metáfora de ecossistema, para Freitas (FREITAS, 2016), permite modelar o universo de DGA interligados, o conjunto de aplicações que orbitam essas bases, o governo e a sociedade como um tipo de ecossistema de software que pode ser compreendido como uma rede de sistemas interligados enviando e recebendo dados via internet. Ding *et al.* (2011) consideram que esse ecossistema dá origem a um sistema em que stakeholders de diferentes tamanhos e papéis encontram, gerenciam, publicam, reutilizam, integram, combinam e consomem dados governamentais abertos em conexão com outras ferramentas e serviços online.

As Figura 4a e 4b, resultado das análises de Freitas (2016) dão materialidade e nos permitem compreender esse ecossistema. Na Figura 4a estão destacados os componentes de primeiro nível do ecossistema de software. Na Figura 4b esses componentes são detalhados.



Figura 4a – Ecossistema de dados abertos – Governo Federal⁵.

⁵ Fonte: Freitas (2016) baseado no modelo de Dawes, Vidiasova e Parkhimovich (2016).

Dados Abertos: transparência governamental

Perspectivas, cenários e planejamento

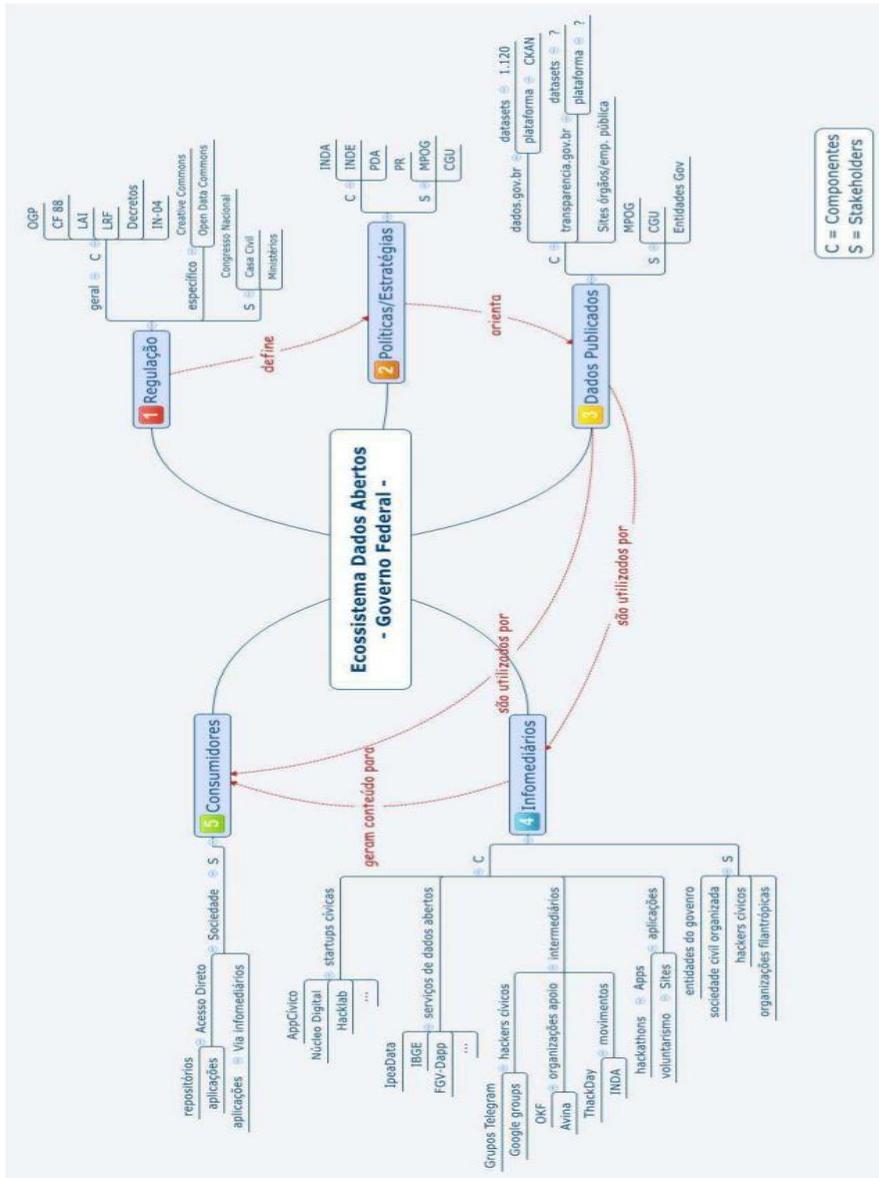


Figura 4b – Ecosistema de dados abertos – Governo Federal⁶

⁶ Fonte: Freitas (2016) baseado no modelo de Dawes, Vidasova e Parkhimovich (2016).

2.5 Open Knowledge Brasil

A Open Knowledge Brasil (OKBr) é uma Organização da Sociedade Civil (OSC) sem fins lucrativos e apartidária, fundada em 2013. A OKBr representa no Brasil a Open Knowledge Internacional a partir de um acordo firmado também em 2013. A OKBr utiliza e desenvolve ferramentas cívicas, faz análises de políticas públicas, oferece treinamento em produção e uso de dados abertos e promove o conhecimento livre para tornar a relação entre governo e sociedade mais transparente e promover uma participação política mais efetiva e aberta.

Durante a última década, a organização tem desempenhado um papel-chave no fortalecimento do ecossistema global de dados abertos e transparência governamental. O portfólio da Open Knowledge Internacional inclui:

- O CKAN. Software livre para catálogo de dados abertos mais bem-sucedido mundialmente.
- O Índice Global de Dados Abertos. Metodologia de avaliação que fornece uma visão abrangente do estado da publicação de dados governamentais abertos no mundo.

Atualmente, segundo a Wikipedia a OKBr desenvolve dois grandes programas. Um deles é o Programa de Ciência de Dados para Inovação Cívica, que hospeda iniciativas como a Operação Serenata de Amor⁷, que utiliza inteligência artificial e mobilização social para fiscalizar os gastos da Câmara dos Deputados. O segundo é a Escola de Dados⁸ que desenvolve capacitações, documentação e eventos, tais como o Coda.Br⁹ e o Cerveja com Dados¹⁰, com foco em organizações da sociedade civil, comunicadores, jornalistas e pesquisadores interessados em métodos digitais de melhor valor agregado. Entre as outras iniciativas destacam-se ainda a

⁷ <https://serenata.ai/> acessado em 02/12/2019.

⁸ <https://escoladedados.org/> acessado em 02/12/2019.

⁹ <https://coda.escoladedados.org/> acessado em 02/12/2019.

¹⁰ <https://www.meetup.com/pt-BR/Cerveja-com-Dados/> acessado em 02/12/2019.

realização do Open Data Index¹¹ no Brasil em parceria com a Fundação Getúlio Vargas, que vislumbra a melhor governança acadêmica, afim de retirar de senso marginal e da alça de mira setores como o ambientalista e sua violência de campo, além de plataformas como o Perfil Político¹² e o portal Orçamento ao seu Alcance¹³.

3 Política de Dados Abertos

*“Mais vale um galo no terreiro,
do que dois na testa.”*

Barão de Itararé

O *Open Government Partnership* (OGP), realizado em setembro de 2011, ficou registrado na história recente do Brasil como o primeiro marco regulatório dos dados que contabilizam a relação ente Estado e Sociedade. O OGP teve lugar na 66^a Assembleia da Organização das Nações Unidas (ONU) para incentivar, em nível global, práticas de *accountability*, acesso público à informação e participação social.

O segundo marco se deu com a aprovação da Lei de Acesso à Informação, LAI, lei 12.527 (BRASIL, 2011), em novembro de 2011, que:

- Assegura o direito fundamental de acesso à informação.
- Segue os princípios da administração pública.
- A publicidade da informação pública é seu princípio maior.
- Trata o sigilo como exceção.
- Conta com a proatividade dos setores públicos na abertura da informação de interesse público.
- Enfatiza a mediação tecnológica como meio para garantir o fluxo de informações.
- Tem como maior objetivo estimular a cultura da transparência e o controle social da administração pública.

¹¹ <http://dapp.fgv.br/tudo-sobre/open-data-index/> acessado em 02/12/2019.

¹² <https://perfilpolitico.serenata.ai/> acessado em 02/12/2019.

¹³ <http://orcamento.inesc.org.br/> acessado em 02/12/2019.

A legislação brasileira, entretanto, não se limita ao contexto destes dois marcos, a legislação é extensa e tem o respaldo nobre da nossa Constituição. As Figuras 5 e 6 apresentam uma síntese da presença da abertura de dados nas leis federais não incluindo leis municipais e estaduais.

Art. 5º, inciso XXXIII	Todos têm direito a receber dos órgãos públicos informações de seu interesse particular, ou de interesse coletivo ou geral, que serão prestadas no prazo da lei, sob pena de responsabilidade, ressalvadas aquelas cujo sigilo seja imprescindível à segurança da sociedade e do Estado;
Art. 37, § 3º	A lei disciplinará as formas de participação do usuário na administração pública direta e indireta, regulando especialmente: [...] II - o acesso dos usuários a registros administrativos e a informações sobre atos de governo, observado o disposto no art. 5º, X e XXXIII;
Art. 216, § 2	Cabem à administração pública, na forma da lei, a gestão da documentação governamental e as providências para franquear sua consulta a quantos dela necessitem.
Leis Complementares	
Lei Complementar nº 131/2009	Lei Capiberibe - acrescenta dispositivos à Lei de Responsabilidade Fiscal
Lei Complementar nº 101/2000	Lei de Responsabilidade Fiscal

Figura 5 – Sobre a abertura de dados na Constituição Brasileira.

Pode-se concluir a partir das informações nos quadros nas figuras 5 e 6 que a abertura de dados após a Constituição de 1988 só retornou ao debate legislativo a partir de 1991. Um caso diverso é o da legislação dos Estados Unidos que, além de antiga, é um dos exemplos mais conhecidos. O *Freedom of Information Act* (FOIA) (1966)¹⁴ e o *Privacy Act* (1974)¹⁵ são garantias de pleno exercício do acesso à informação.

¹⁴ Disponível em <https://www.foia.gov/>. acessado em 9 mai.2019.

¹⁵ Disponível em <https://www.govinfo.gov/content/pkg/USCODE-2012-title5/pdf/USCODE-2012-title5-partI-chap5-subchapII-sec552a.pdf>. Acessado em 9 mai. 2019.

A abertura de dados tem ainda como subproduto dois conceitos relativos à transparência na informação pública:

- **Transparência passiva.** Regula o direito dos cidadãos de obterem informações governamentais e estabelece procedimentos, prazos e meios de recurso em caso de recusa por parte do poder público.
- **Transparência ativa.** Regula as informações mínimas que devem ser publicadas periodicamente pelos governos, sem qualquer demanda por parte dos cidadãos.

Lei nº 12.527/2012	Lei de Acesso à Informação
Lei nº 10.520/2002	Pregão Eletrônico
Lei nº 9.784/1999	Lei do Processo Administrativo
Lei nº 9.507/1997	Rito processual do habeas data
Lei nº 8.159/1991	Política Nacional de arquivos públicos e privados
Decreto nº 7.845/2012	Procedimentos de credenciamento de segurança e tratamento de informações classificadas
Decreto nº 7.724/2012	Regulamenta a Lei de Acesso à Informação no Poder Executivo Federal
Decreto s/nº de 15/9/2011	Institui o Plano de Ação Nacional sobre Governo Aberto e dá outras providências
Decreto nº 6.932/2009	Carta de Serviços ao Cidadão
Decreto 6.170/2007	Sistema de Gestão de Convênios e Contratos de Repasse – SICONV
Decreto nº 5.482/2005	Portal da Transparência e Páginas de Transparência Pública
Decreto nº 5.301/2004	Comissão de Averiguação e Análise de Informações Sigilosas
Decreto nº 4.073/2002	Regulamenta a Política Nacional de Arquivos Públicos e Privados
Normativa nº 4 de 13 de Abril de 2012	Institui a Infraestrutura Nacional de Dados Abertos – INDA é relativa a criação do INDA seus objetivos e sua estrutura
Portaria SLTI/MP nº 92, de 24 de dezembro de 2014	Institui a arquitetura ePING (Padrões de Interoperabilidade de Governo Eletrônico), que define um conjunto mínimo de premissas, políticas e especificações técnicas que regulamentam a utilização da Tecnologia de Informação e Comunicação (TIC) na interoperabilidade de serviços de Governo Eletrônico.
Portaria nº 1.254/2015 (MPOG, CGU)	Institui o Sistema Eletrônico do Serviço de Informação ao Cidadão (e-SIC) no âmbito do Poder Executivo federal
Portaria nº 233/2012 (MPOG, CGU, MF, MD)	Remuneração de servidores e agentes públicos
Portaria nº 516/2010 (CGU)	Cadastro Nacional de Empresas Inidôneas e Sancionadas – CEIS
Portaria nº 262/2005 (CGU)	Relatórios de Auditoria na Internet

Figura 6 – Sobre a abertura de dados em decretos, portarias, normativas e leis brasileiras.

A LAI implementa a transparência passiva através de protocolos de

solicitação de informação nos Serviços de Acesso à Informação (SICs)¹⁶. A solicitação pode ser submetida por qualquer meio legítimo, físico ou eletrônico, com prazos de resposta previstos de no máximo vinte dias, prorrogáveis por mais dez dias, sem qualquer motivação ou justificativa e são gratuitas mediante identificação do requerente. A Figura 7 apresenta as ações realizadas em prol da transparência (ativa e passiva) no Brasil.

	Transparência ativa	Transparência passiva
Voluntariado	ONG Transparência Brasil	ONG Transparência Brasil
Legislação e controle	Lei nº 12.527/2011 ou LAI e Decreto nº 7.724/2012.	Lei nº 12.527/2011 ou LAI e Decreto nº 7.724/2012.
Educação e treinamento	Curso Lei de Acesso à Informação (parceria Senado/UFGM). Criação de guias, cartilhas e manuais para a publicação de dados governamentais. Guia de Abertura de Dados da Inda.	Curso Ouvidoria na Administração Pública (parceria CGU/ILB). Criação de guias para a disponibilização de dados custodiados por instituições públicas.
Instrumentos econômicos	Criação de seções de acesso à informação em portais de instituições públicas federais. Portal de Acesso à Informação. Portal Brasileiro de Dados Abertos. GT da Inda.	Criação do SIC e do e-SIC.
Ator responsável pela análise e disponibilização dos dados	Agente público.	Servidor do SIC ou de Ouvidoria Pública.
Órgão responsável	Órgão público ou instituição relacionada ao setor público.	SIC ou Ouvidoria Pública.

Figura 7 - Ações realizadas em prol da transparência (ativa e passiva)

¹⁶ Disponível em

<https://esic.cgu.gov.br/sistema/site/index.aspx?ReturnUrl=%2fsistema%2fprincipal.aspx>. Acessado em 12 ago. 2019.

Note-se que as políticas de acesso à informação, implementadas pela INDA¹⁷, diferem das políticas de disponibilização da informação, implementadas nos portais de dados abertos. Apesar da mudança de cultura ter início no alto escalão das organizações, é fundamental que os responsáveis diretos pelos sistemas de informação estejam envolvidos no planejamento e desenvolvimento das ações de abertura de dados. Grande parte dos sistemas de informação dos órgãos públicos, assim como seus respectivos bancos de dados, têm sua manutenção terceirizada. Nessas condições é fundamental o entendimento desses atores quanto aos objetivos pretendidos no processo de abertura de dados.

A estratégia de abertura de dados aqui apresentada defende que, uma vez que a sociedade é o maior beneficiário desses dados, deve considerar a participação social processo fundamental na construção dos canais ideais de disseminação desses dados.

3.1 INDA

A Infraestrutura Nacional de Dados Abertos (INDA) é responsável pela regulamentação do Portal Brasileiro de Dados Abertos¹⁸ e é a política que garante e facilita o acesso pelos cidadãos, pela sociedade e, em especial, pelas diversas instâncias do setor público aos dados e informações produzidas ou custodiadas pelo Poder Executivo federal, (BRASIL, 2012b). Esta política decorre do *Open Government Partnership* e do Decreto s/nº de 15 de setembro de 2011 que institui o Plano de Ação Nacional sobre Governo Aberto, o qual estabelece o compromisso do governo de implantar a mesma Infraestrutura Nacional de Dados Abertos.

A proposição da INDA considera como pressupostos:

¹⁷ Instituída pela Instrução Normativa nº 4, de 12 de abril de 2012.

¹⁸ <http://dados.gov.br/>

- Que o direito à informação constitui fundamento básico da democracia e que para o cidadão exercê-lo deve lhe ser facilitado o acesso a informações primárias, íntegras, autênticas e atualizadas.
- Que a adoção de meios eletrônicos para a disponibilização de dados públicos necessita que esses dados sejam publicados de forma que facilite seu reuso e que permitam o acesso simplificado para os seus usuários, premissas presentes nos princípios de dados abertos.
- O disposto no art. 48 da Lei Complementar nº 101, de 4 de maio de 2000, que determina ao Poder Público a adoção de instrumentos de transparência na gestão fiscal em meios eletrônicos de acesso público às informações orçamentárias e prestações de contas.
- A Lei nº 12.527, de 18 de novembro de 2011, que regula o acesso a informações previsto no inciso XXXIII do art. 5º, no inciso II do § 3º do art. 37 e no § 2º do art. 216 da Constituição Federal,

A instrução normativa que institui a INDA (BRASIL,2012b) trata em seu artigo segundo dos objetivos desta política, que são os seguintes:

- Definir, estruturar e coordenar a política de dados abertos, bem como estabelecer o seu modelo de funcionamento;
- Fomentar o ordenamento na geração, armazenamento, acesso, e compartilhamento de dados para uso do Poder Executivo Federal e da sociedade;
- Definir e disciplinar os padrões e os aspectos técnicos referentes à disponibilização e disseminação de dados para uso do Poder Executivo Federal e da sociedade;
- Estimular o compartilhamento de recursos de tecnologia da informação e evitar a duplicidade de ações e o desperdício de recursos na disseminação de dados e informações pelos órgãos e entidades do Poder Executivo Federal;
- Apoiar, capacitar e fornecer suporte para a publicação de dados abertos aos órgãos e entidades do Poder Executivo federal que aderirem à INDA que não possuem prática, cultura e atribuições finalísticas de disseminação de dados;
- Buscar a melhoria contínua da publicação de dados abertos, baseando-se nas melhores práticas concebidas nos cenários nacional e internacional;

- Promover a colaboração entre governos dos diferentes níveis da federação e entre o Poder Executivo Federal e a sociedade, por meio da publicação e do reuso de dados abertos;
- Promover e apoiar o desenvolvimento da cultura da publicidade de dados e informações na gestão pública;
- Disponibilizar tecnologias e apoiar as ações dos órgãos e entidades do Poder Executivo Federal que aderirem à INDA na implementação da transparência ativa por meios digitais; e
- Estimular a participação social na construção de um ecossistema de reuso e de agregação de valor dos dados públicos.

Em seu artigo terceiro a instrução normativa define os atores que integram a INDA:

- Obrigatoriamente o Órgão Central, os Órgãos Setoriais, os Órgãos Seccionais e Correlatos do Sistema de Administração de Recursos de Informação e Informática – SISPI, conforme definido pelo Decreto nº 7.579, de 11 de outubro de 2011; e
- Facultativamente, mediante a assinatura do termo de adesão constante do Anexo pela autoridade competente, os demais órgãos e entidades dos Poderes Executivo, Legislativo e Judiciário, das esferas Federal, Estadual, Distrital e Municipal.

E ainda define que:

- Os cidadãos e entidades da sociedade civil poderão participar de sua implementação independentemente da assinatura de termo de adesão, nos moldes do que dispuser o regimento interno da INDA.
- Entidades privadas nacionais ou internacionais poderão colaborar com a INDA mediante a celebração de termo de cooperação específico para este fim, sem ônus para Administração.

A regulamentação desta política se dá nos documentos listados na Figura 8. Deste conjunto cabe destacar o Plano de Dados Abertos (PDA) que é o documento orientador do Ministério do Planejamento Orçamento e Gestão (MPOG) para as iniciativas de implementação e promoção de abertura de dados de modo a garantir o mínimo de qualidade e a facilitar tanto o entendimento quanto a reutilização das

informações. Neste conjunto estão incluídos os dados geoespacializados.

Documentos Normativos	Descrição
Política de dados abertos	
Lei nº 12,527/2012	Lei de Acesso à Informação
Decreto nº 7,724/2012	Regulamenta a Lei de Acesso à Informação no Poder Executivo Federal
Instrução Normativa SLTI/MP nº 4/2012	Institui a Infraestrutura Nacional de Dados Abertos
Decreto 8,777/2016	Política de Dados Abertos do Poder Executivo
Dados geoespacializados	
Decreto nº 6.666/2008	Institui, no âmbito do Poder Executivo Federal, a Infra-Estrutura Nacional de Dados Espaciais - INDE, e dá outras providências.
Contexto mais amplo	
Decreto de 15 de setembro de 2011	Institui plano de ação nacional sobre governo aberto e da outras providências
Decreto 8,638/2016	Institui a Política de Governança Digital
Documentos de Planejamento	Descrição
Política de dados abertos	
Plano de Ação da INDA	Instrumento de planejamento central da política de dados abertos
Planos de Dados Abertos	Plano de cada organização
Contexto mais amplo	
Plano de Ação do Brasil	Plano do Brasil na Parceria para Governo Aberto
Estratégia de Governança Digital	Aprovada pela Portaria nº 68, do Ministério do Planejamento
Documentos de Orientação	
Cartilha Técnica para Publicação de Dados Abertos no Brasil	
Kit para Dados Abertos	
Guia de Abertura de Dados	
Manual para a Elaboração de Plano de Dados Abertos	
Arquitetura Técnica Referencial de Abertura de Dados	
e-PING	
Acórdãos do TCU	Descrição
Acórdão 2.569/2014	Relatório de levantamento
Acórdão 228/2015	Relatório sistêmico de fiscalização de tecnologia da informação (fiscti)
Acórdão 3,022/2015	Auditoria operacional

Figura 8 – Documentos da Política de Dados Abertos

Desde o início da construção colaborativa da Infraestrutura Nacional de Dados Abertos (INDA), em março de 2011, a sociedade tem contribuído ativamente. Servidores públicos do governo, de universidades e de vários poderes e esferas, em reuniões na Escola de Administração Fazendária (ESAF), definiram metas para a criação da infraestrutura de governo que seria capaz de apoiar quem se

interessasse pela publicação de dados abertos. A proposta visava estabelecer uma plataforma que tornasse viável a utilização dos dados pelos cidadãos. O trabalho foi acompanhado em modo não presencial pelo grupo Transparência Hacker, que há tempos cobrava do governo brasileiro meios para publicação de informações do setor público no formato de dados abertos em modo similar ao de outros países.

Em maio de 2012 o Ministério do Planejamento inaugurou o Portal Brasileiro de Dados Abertos que teve sua construção baseada em soluções livres e de código aberto. A principal ferramenta da plataforma é o *Comprehensive Knowledge Archive Network* - CKAN, desenvolvido pela *Open Knowledge Foundation* para catalogação de dados.

O desenvolvimento do Portal Brasileiro de Dados Abertos se deu nos marcos organizados nas Tabela 3 e 4.

Março	Início à construção colaborativa do que viria a ser a INDA. Primeiro workshop público de construção da INDA.
Abril	Workshop público de construção da INDA – já com transmissão online.
Maior	Verificação e aprovação do Plano de Trabalho da INDA.
Julho	Reunião de acompanhamento do plano de projeto.
Agosto	Segundo Rumble Portal Dados Abertos.
Setembro	Terceiro Rumble Portal Dados Abertos; Quarto Rumble Portal Dados Abertos.
Outubro	Reunião para discussão da Minuta da Instrução Normativa da INDA; Audiência Pública da Instrução Normativa da INDA; I Encontro Nacional de Dados Abertos.
Novembro	Sétimo Rumble Portal Dados Abertos.
Dezembro	Lançamento da versão Beta do Portal.

Tabela 3 – Marcos do desenvolvimento do Portal Brasileiro de Dados Abertos em 2011.

Janeiro	Oitavo Rumble Portal Dados Abertos.
Fevereiro	Cardsorting na Campus Party. Nono Rumble Portal Dados Abertos. 2ª sprint review e retrospectiva do desenvolvimento do dados.gov.br.
Março	3ª sprint review e retrospectiva do desenvolvimento do dados.gov.br. Estudo e desenvolvimento das propostas de arquitetura de informação e redesign da página. 4ª sprint review e retrospectiva do desenvolvimento do portal dados.gov.br. Décimo Rumble Portal Dados Abertos. XI Rumble Portal Dados Abertos.
Maiο	Lançamento da Versão Oficial do Portal.

Tabela 4 – Marcos do desenvolvimento do Portal Brasileiro de Dados Abertos em 2012.

Atualmente, o portal encontra-se em ambiente de produção e conta com atualizações e novos conjuntos de dados são publicados todos os meses. Em novembro de 2019 eram 165 organizações participantes do Portal, 32 aplicativos e 7144 conjuntos de dados. Para mais detalhes de cada etapa executada na construção do Portal Brasileiro de Dados Abertos, consulte o link: <http://dados.gov.br/processo-de-participacao-social-dainda/#linha-do-tempo>.

O principal desafio na publicação de dados abertos é a implementação de um Portal de Dados Abertos na Internet que objetiva ser o ponto central para a busca e o acesso a dados públicos governamentais desta instância de governo.

As orientações para criação do Portal são:

- Estrutura baseada na arquitetura da Web.
- Conjuntos de dados logicamente referenciáveis na Web.

- Conjuntos de dados com opção de busca e localização dentre outras centenas de milhares de conjuntos de dados.

O Portal deve apresentar os seguintes requisitos técnicos:

- Controle de acesso: somente as pessoas autorizadas devem ter acesso para publicação e modificação dos dados e metadados. É recomendado a utilização de mecanismos de autenticação (usuário e senha).
- Integridade: deve ser assegurado que os dados não sejam adulterados durante a transferência.
- Autenticidade: deve ser assegurado que os dados provêm de uma fonte legítima da instituição.

4 Planejamento da abertura de dados

“Os pequenos atos que se executam são melhores do que todos os grandes que apenas se planejam.”

*George Marshall, general norte-americano,
Nobel da Paz em 1953*

A INDA, em seu material de apoio, dispõe de um documento de orientação: o Guia de Abertura de Dados. Este texto norteia a criação do processo de gestão das ações para publicação de dados abertos dentro de uma instituição governamental. Para que um dado de um órgão possa integrar o conjunto de dados da INDA seu responsável deve garantir que este conjunto de dados cumpra as seguintes condições:

- Devem estar em seu formato mais bruto possível.
- Devem estar em formato aberto.
- Devem estar disponíveis para leitura em ferramenta não proprietária, estável e de amplo uso, ou seja, ser publicados em JSON, XML, CSV, ODS, RDF, SVG, GML HTML/RDFA.
- Sua reutilização e redistribuição devem estar livres de impedimento legal.
- Dados estruturados ou em planilhas na sua fonte devem ter sua estrutura original preservada ao máximo.
- Cada conjunto de dados deve possuir uma URL padronizada.

- O Portal não deve utilizar mecanismos antirrobôs, como captchas, para acesso aos conjuntos de dados.
- Os dados abertos devem respeitar a privacidade dos cidadãos conforme citado no § 5º, Artigo 31 da LAI.
- A instituição deve se responsabilizar pelos dados publicados por ela.

Preparar os dados para abertura implica em identifica-los, formatá-los e obter permissão dos patrocinadores no órgão para a publicação. Em teoria a instituição deve publicar o conjunto total de dados sob sua guarda o mais rápido possível. Entretanto, na prática, com as limitações financeiras e técnicas presentes em toda a estrutura governamental os órgãos detentores dos dados não estão organizados para as ações de publicação dos dados nos moldes preconizados pela INDA. Sugere-se como boa prática que a instituição publique as informações conforme dispõe a Lei 12.527 / 2011, em seu Art. 8º, §1º e informações adicionais de interesse público. Um bom conjunto de dados a ser publicado seria o seguinte:

- Dados sobre desempenho de programas, projetos e ações finalísticas da instituição.
- Endereços, informações para contato, horários de funcionamento e estrutura organizacional.
- Dados referentes ao relacionamento com o cidadão.
- Dados referentes a licitações, editais e contratos.
- Registro de repasses, transferências de recursos financeiros e despesas.

A publicação dos dados em um Portal de Dados Abertos necessita dos seguintes artefatos e definições:

- *Dataset* (conjunto de dados). É o arquivo gerado a partir da base de dados original. Precisa seguir normas e padrões de dados abertos para que sua publicação seja aceita no Portal.
- Dicionário de dados. O dicionário de dados corresponde às informações que traduzem cada campo do *dataset* e informa suas propriedades.
- Metadados. Os metadados são informações que detalham o *dataset* e são necessárias à carga do *dataset* no CKAN.
- Área temática. Informação que organiza os *datasets* em temas de acordo com a lógica de organização dos serviços da instituição responsável pelo

dataset. As áreas temáticas da Prefeitura de Belo Horizonte estão listadas no Anexo 1.

- Organização. Informação que identifica a organização responsável pelo *dataset*.

Para que os dados de uma instituição sejam abertos em sua totalidade é essencial o conhecimento estrutural e operacional do órgão. A análise deve considerar dados gerados ou mantidos sob responsabilidade da instituição governamental. O profissional à frente deste trabalho deve analisar os processos de negócio da instituição, identificar todos os grupos de dados produzidos nesses processos e organizá-los no melhor formato de acesso no Portal.

4.1 Obter autorização para publicação

É importante que a instituição tenha segurança em relação à liberação do conjunto de dados. Para tanto é necessário considerar as seguintes recomendações:

- Deve-se fazer uma análise jurídica para se ter uma permissão de publicação formal por parte da instância responsável.
- Deve-se ter certeza de que não há instrumentos jurídicos, portarias, instruções normativas ou afins que proíbam a publicação de certas informações presentes no conjunto de dados em questão.
- Deve-se avaliar se a publicação desses dados pode comprometer algum dos itens a seguir em seu estado natural: segurança nacional, segredos estratégicos de governo, direito de privacidade do cidadão ou legislação.

Caso a publicação do conjunto de dados seja impedida por alguma razão será necessário que se faça uma adequação para que seja possível publicar os dados. Uma adequação relativamente corriqueira é a anonimização, na qual são retirados os dados motivadores da restrição à publicação. Observe-se que a anonimização não é apenas um mascaramento do dado.

Tomando como exemplo o caso dos dados de uma pessoa em uma determinada base de dados a anonimização não se dá apenas alterando o nome da pessoa para algo genérico, pois outros dados da mesma pessoa podem permitir o

cruzamento de dados ou a filtragem do dado resultando em sua identificação.

Este processo precisa seguir protocolos reconhecidamente bem-sucedidos para não se correr o risco do dado ser desanonimizado. Outro aspecto a se observar é que quanto mais anônimo o dado, maior será a redução do valor agregado ao mesmo, podendo se tornar irrelevante e restringir a transparência da instituição.

Apesar destas restrições mantém-se a orientação de que todos os dados da instituição devem ser publicados com exceção daqueles com alguma restrição. Cabe notar ainda que, mesmo não publicado integralmente, um grupo de dados sempre agrega algum valor para a sociedade.

4.2 Atribuição de licença de uso

Tem-se como premissa para dados abertos a definição de que os dados publicados estão liberados para qualquer finalidade de uso. Entretanto, objetivando a melhor utilização dos dados as organizações internacionais responsáveis pelos padrões para dados abertos propõem atribuir licenças de acordo com o assunto e as condições de uso dos dados. Essa atribuição implica disponibilizar as informações de forma completa na Web, em formatos não proprietários, com acesso irrestrito e livre para qualquer cidadão e de modo que qualquer um possa reutilizá-los para qualquer fim, sem restrições. Sem essa licença, os dados não podem ser considerados "abertos", pois não há garantias de que sua utilização esteja resguardada pelos princípios da abertura de dados.

Assim, a atribuição da "licença aberta" é importante para que se proteja o que tem que ser protegido e para que a Lei de Acesso à Informação seja respeitada. A Infraestrutura Nacional de Dados Abertos, do governo federal, ainda não concluiu se o ordenamento jurídico brasileiro é adequado para tratar os dados governamentais de forma aberta, sem a necessidade de licenças, se outras licenças que já existem se enquadram no contexto brasileiro, ou se será necessário criar uma licença específica para o Brasil.

Pode-se citar como exemplo, no caso da legislação brasileira, a Lei de

Direitos Autorais. A lei prevê a proteção a banco de dados que constituam "criação intelectual". Entretanto, por seu turno, a Lei de Acesso à Informação afirma que o governo deve proporcionar o acesso a toda informação que produz ou detém, desde que não seja considerada sigilosa. A interação conflitante dessas leis ainda é objeto de estudo.

A LAI não trata da determinação de qual licença será aplicada ao se publicar dados abertos, mas lista princípios que tendem a direcionar a determinação de uma licença aberta ou termos de utilização compatíveis com os conceitos de abertura de dados. Este direcionamento tem como premissa que os dados governamentais são públicos, configurando o sigilo como exceção.

A compatibilidade de licenças entre as diversas instâncias de governo é outra questão que não pode ser desconsiderada ao se atribuir um licenciamento a um conjunto de dados. Pois, um caso que podemos considerar como exemplo é o de um governo estadual que usa uma licença incompatível com a licença usada por uma administração municipal em seu território gerando conflitos legais no uso de seus *datasets* por parte da sociedade.

A materialização desta hipótese poderia ser o caso em que não seja possível, legalmente, cruzar as bases de dados de escolas, por exemplo, para construir um aplicativo que exhibe as instalações municipais e estaduais. Uma licença ou termos de uso de dados abertos devem ser atribuída de modo a maximizar o grau de compatibilidade com as licenças usadas ou que possam vir a ser usadas por outras esferas e poderes do governo.

No Brasil, a maior parte dos órgãos públicos não atribui qualquer licença ao publicar dados na Internet. Há exceções, contudo, com a utilização das licenças ODbLe DbCL.

As licenças recomendadas são as seguintes (Brasil, 2012d):

- *Open Database License (ODbL)*¹⁹: Qualquer pessoa é livre para copiar, distribuir, usar, criar novas publicações a partir da abertura original, modificar, construir ou transformar os dados (ou o banco de dados). Usuários têm que atribuir a fonte, da forma que for definida na licença e garantir que reusos de seus trabalhos com os dados sigam a mesma regra de atribuição. Nos casos de usos dos dados para fins comerciais, deve ser disponibilizada cópia pública.
- *Database Content License (DbCL)*: mesmas orientações da ODbL, mas aplicável ao conteúdo dos bancos de dados - os dados em si;
- *Creative Commons Zero (CC0)*²⁰: aplicável a contextos em que o conjunto de dados caracteriza-se como de domínio público. Usada quando a análise do gestor aponta para a inexistência de direito autoral e passa a ser “natural” a renúncia permanente aos direitos de copyright e outros relacionados. Nenhuma atribuição à instituição gestora dos dados é esperada por parte dos usuários.
- *Creative Commons Attribution 4.0*²¹: permite ao usuário (qualquer usuário) realizar livremente atividades de compartilhamento (cópia e redistribuição dos dados em qualquer formato ou mídia), de adaptação (incluindo remixagem, transformação e construção de novos produtos) para quaisquer propósitos inclusive comerciais. Ficam os usuários sujeitos às seguintes condições: (i) atribuir crédito apropriado, fornecendo URL da licença e indicando eventuais mudanças; (ii) ao atribuir crédito não é necessário seguir formato padrão, mas é exigido que não seja de forma que leve o leitor a supor que a instituição que licenciou o conteúdo endossa o usuário ou os trabalhos produzidos com os dados licenciados.

É essencial que os cidadãos interessados em usar dados do Portal estejam cientes que as condições que regem cada conjunto de dados podem ser atualizadas, corrigidas e/ou substituídas a qualquer tempo, a critério da instituição publicadora.

¹⁹ Conteúdo integral da licença disponível em <https://opendatacommons.org/licenses/odbl/>

²⁰ Conteúdo integral da licença disponível em <https://creativecommons.org/share-your-work/public-domain/cc0/>.

²¹ Conteúdo integral da licença disponível em <https://creativecommons.org/licenses/by/4.0/>

4.3 Elaborar os metadados

A abertura de dados tem por objetivo permitir o acesso aos dados da administração pública por parte da sociedade. Este objetivo considera que as instituições devem publicar toda informação sobre sua tutela e que esta publicação seja a mais automatizada possível. Entretanto, a publicação não deve resultar em obstáculo ao seu acesso e sim ser a mais amigável e compreensível por aqueles que fazem as consultas a estes dados.

Se um conjunto de dados publicado não está produzindo conclusões ou não está sendo utilizado para algum serviço, então ele não satisfaz o objetivo inicial. É papel do provedor dos dados fornecer artefatos que sejam suficientes para a máxima utilização dos dados.

Este item se refere, então, à publicação de informações que respaldem a utilização do conjunto de dados. Os metadados do conjunto de dados devem ser suficientes para que um usuário externo à instituição compreenda a definição, a validade e o escopo dos dados. É importante informar do que se trata um grupo de dados, o intervalo de tempo entre as publicações, a data de publicação, direitos de uso etc.

Propõe-se para o Portal de Dados Abertos, como padrão os conjuntos de metadados obrigatórios e desejáveis listados nas tabelas 5 e 6.

	Nome do Metadado	Descrição
1	Título	Nome do conjunto de dados.
2	Descrição	Breve explicação sobre os dados.
3	Etiquetas	Lista de palavras-chave relacionadas ao conjunto de dados que são úteis na classificação e busca dele. Ex.: economia, aposentadoria, ENEM.).
4	Tipo de Licença	Para os <i>datasets</i> gerados a partir de bases de dados geradas ou custodiadas pelas instituições publicadoras do setor público sugere-se a licença Open Database License. Para os demais <i>datasets</i> a licença será decidida caso a caso.
5	Catálogo origem	Página (URL) do órgão onde está publicado o

Dados Abertos: transparência governamental
 Perspectivas, cenários e planejamento

	Nome do Metadado	Descrição
		conjunto de dados.
6	Órgão responsável	Nome e sigla do órgão ou entidade responsável pela publicação do conjunto de dados.
7	Identificador	URL persistente que aponta para o recurso na Web.
8	Nome do arquivo	Nome do arquivo do <i>dataset</i> .
9	Formato	Formato do <i>dataset</i> . Ex.: XML, JSON, CSV, etc.
10	Data de Publicação	Data de publicação no Portal.
11	Cobertura temporal	Data ou período à que referem os dados do <i>dataset</i> . Ex.: 09/2017.
12	Nome do mantenedor	Nome do Mantenedor dos dados do <i>dataset</i> .
13	E-mail do mantenedor	E-mail do Mantenedor dos dados do <i>dataset</i> .
14	Versão do arquivo	Número da versão do arquivo do <i>dataset</i> .
15	Frequência de atualização	Frequência temporal com que o recurso é atualizado.
16	Melhor horário	Horário menos sobrecarregado para extração no do sistema de origem (apenas para publicações de <i>Datasets</i> de forma automática): Ex.: 23:00 (vinte e três horas).
17	Cobertura geográfica	Localização ou região geográfica a que se refere o <i>dataset</i> . Caso esta informação se aplique. Ex.: Barreiro.
18	Granularidade geográfica	Precisão geográfica da cobertura geográfica. Caso esta informação se aplique. Ex.: regional.
19	Granularidade temporal	Precisão temporal da cobertura temporal. Caso esta informação se aplique. Ex.: mês.

Tabela 5 - Metadados obrigatórios.

	Nome do Metadado	Descrição
	Documentação	URL de documento que expõe detalhes sobre o conjunto de dados.
	Referências	Relações com outros conjuntos de dados.
	Metodologia	Material com explicação sobre o processo de criação dos <i>datasets</i> .

Tabela 6 - Metadados desejáveis.

O vocabulário utilizado deve ser submetido ao *checklist* proposto pelo W3C (W3C, 2014) e deve seguir vocabulários controlados e ontologias que sejam de conhecimento público. Essa iniciativa deve ser parte das diretrizes da instituição publicadora para composição dos metadados dos dados a serem abertos objetivando facilitar seu entendimento, evitar estrangeirismos, regionalismos e neologismos, além de maximizar a possível conexão com outros conjuntos de dados.

Recomenda-se utilizar o VCGE – Vocabulário Controlado do Governo Eletrônico – (Brasil, 2014b) devido à qualidade de sua estrutura e a ser o vocabulário padrão para dados governamentais no Brasil.

4.4 Priorizar dados a serem abertos

Os defensores do conceito de dados abertos vislumbram o cenário em que todos os dados de uma instituição são publicados o mais rápido possível e mantidos sempre atualizados. Entretanto, a realidade impõe adequações face ao que é de fato possível. Assim, é uma boa prática analisar o potencial de valor dos dados para a sociedade e elaborar um planejamento de publicação priorizando os mais importantes e considerando os recursos disponíveis para esta ação.

Dentre as boas práticas para seleção de dados a serem abertos temos as seguintes orientações:

- Dados catalogados ou criados prioritariamente pela instituição publicadora na estrutura administrativa da sua unidade federativa e que podem ser liberados ao público. Ex.: dados da arrecadação tributária do município.
- Dados que ao serem combinados com outros dados produzam grande valor. Ex.: dados de evasão escolar combinados com dados de adoecimento em determinada região do município.
- Dados que tenham potencial de reuso. Ex.: dados do fluxo de trânsito nas vias preferenciais.
- Dados com potencial para compor séries históricas. Ex.: quantificação de serviços prestados.

- Dados muito requisitados através dos meios de solicitação de acesso às informações públicas pela sociedade. Ex.: número de vagas por tipo de ensino por escola.
- Dados de interesse público frequentemente entregues à imprensa. Ex.: dados de contaminação por dengue.
- Dados não sigilosos regularmente solicitadas por outras instituições. Ex.: dados relativos à qualidade de vida.
- Dados anonimizados que continuam tendo valor para a sociedade. Ex.: quantificações referentes ao atendimento da população em vulnerabilidade social.
- Dados com grau de saneamento que não comprometam o conjunto de dados publicado. Ex.: indicadores de desempenho dos alunos da rede municipal.
- Dados de interesse público em mídia que esteja fora dos padrões de dados abertos, mas que podem ser facilmente convertidas e liberadas. Ex.: mapeamento georreferenciado de infestação de mosquitos transmissores da dengue.
- Dados estáveis com baixo grau de correção após sua publicação para a sociedade. Ex.: série histórica do desempenho da rede hoteleira no município.
- Dados que podem compor rotinas automatizadas. Ex.: dados referentes à posição atual dos ônibus municipais.

Cabe lembrar que após sua publicação o conjunto de dados não deve ter seu conteúdo alterado sem ampla divulgação. A alteração pode provocar resultados inesperados e indesejáveis em sua cadeia de utilização.

A clareza do papel de cada pessoa atuante na abertura de dados potencializa sua melhor participação. Os papéis são os seguintes:

1. Cidadão. É o usuário central dos dados disponibilizados e das aplicações. O PDA deve conter ações informativas do processo de abertura como um todo, esclarecendo a priorização. Também deve incentivar a utilização das aplicações que utilizam os dados disponibilizados.
2. Grupos organizados. É uma boa prática incluir esses grupos nas etapas da abertura dos dados, pois podem contribuir para uma melhor elicitação

- dos requisitos técnicos e formas de acesso aos dados. Eles têm a possibilidade de atuar como entidades facilitadoras do acesso à informação presente nos dados abertos.
3. Acadêmicos. Universidades e centros de pesquisa têm capacidade técnica e são parceiros importantes no aprimoramento de práticas e de aplicativos baseados nos dados abertos. O governo pode se beneficiar das análises multidisciplinares realizadas por especialistas deste segmento.
 4. Gestores da administração pública. O público-alvo principal dos dados abertos é o cidadão e suas formas de organização. Entretanto, a abertura de dados das instituições do governo constitui uma iniciativa de interoperabilidade e pode ser utilizada como ferramenta de gestão. Os dados abertos reduzem o trabalho de geração de informação de instituição para instituição e aumentam a utilização de dados publicados, pois o dado publicado pensado em uma instituição pode ser utilizado por toda a máquina pública.
 5. Imprensa. A imprensa tem o potencial de aumentar a visualização dos dados pela sociedade, pois já é uma fonte de informação utilizada por ela.

É uma boa prática organizar o PDA nas seções ilustradas na Figura 9. Cabe destacar:

- Apresentação. Objetiva explicar o propósito do PDA, seu alinhamento com a legislação e a relevância do documento para o cidadão.
- Introdução. Aqui são apresentados os objetivos gerais e específicos da publicação dos dados previstos no PDA.
- Definição dos dados a serem abertos. Local de apresentação dos dados que o órgão se comprometeu a abrir e outros com possibilidade de abertura.
- Estratégia para abertura e publicação em um portal de dados abertos. As etapas que o órgão irá adotar para a execução do processo de publicação estão organizadas nesta seção.
- Plano de ação. Aqui são detalhadas as atividades e produtos que serão derivados do PDA.

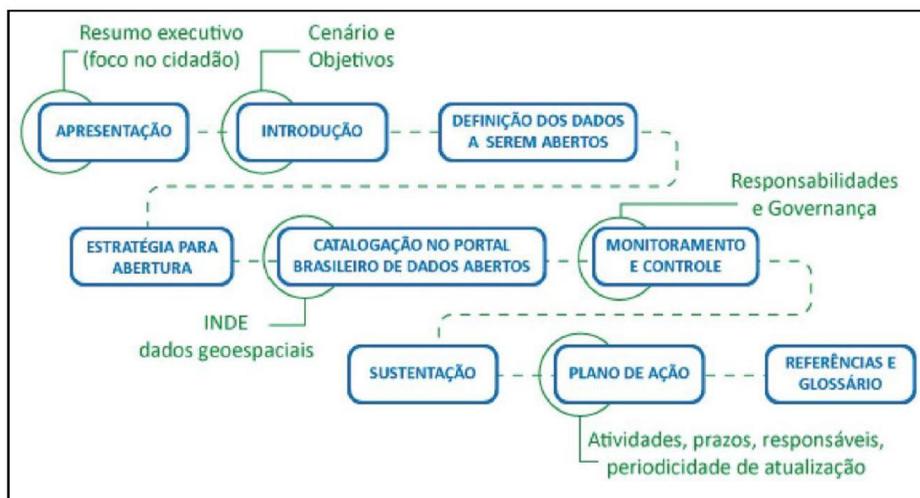


Figura 9 – Estrutura do Plano de Dados Abertos (PDA)

4.5 Plano de sustentação

O maior desafio das instituições públicas que integram o grupo de publicadores de dados abertos é tornar a adesão em um projeto contínuo, sustentado e com manutenção atualizada dos dados. Assim, é recomendável se pensar em um plano de sustentação no início da ação de abrir os dados, pois caso não seja possível mantê-los atualizados a iniciativa perde seu propósito.

O plano deve então considerar dois aspectos. O primeiro aspecto é a continuidade da abertura. Já no início do processo de abertura identifica-se a necessidade de priorizar os dados que serão abertos face à quantidade finita de recursos.

Entretanto, a partir do momento que se considera a abertura como ação imposta por leis é salutar incorporar o valor estratégico desta ação para a instituição inserindo o processo de geração e publicação de dados abertos no conjunto de procedimentos de persistência de dados nos sistemas de informação origem. Nesta ação deve-se prever melhorias na geração e publicação dos dados abertos através da troca de experiências com outras instituições da administração pública e a satisfação dos usuários com o acervo publicado.

O segundo aspecto trata da manutenção dos dados já disponibilizados. Sempre que uma instituição publica um conjunto de dados em formato aberto ela assume o compromisso de mantê-lo disponível e íntegro. Considerando que a maioria dos dados têm perfil de séries históricas gerados ininterruptamente o plano de sustentação deve prever a manutenção da série com publicações em intervalos de tempo regulares, com versões mais recentes sempre que possível e preservando sempre seu valor agregado. A frequência de atualização e seus custos operacionais devem ser considerados quando da definição da estrutura da série temporal.

4.6 Governança

Toda ação de governo precisa responder a princípios da administração pública, tais como legalidade, impessoalidade, moralidade administrativa, publicidade e eficiência. Por ser uma ação de governo a abertura de dados não pode desconsiderá-los e todos têm igual importância. Vejamos como os princípios podem ser situados no contexto dos dados abertos.

Princípio 1 – Legalidade. Publicar os dados da instituição é, por princípio, o cumprimento da legislação em vigor. A publicação deve ser executada conforme os códigos da administração pública e gera responsabilidade civil e criminal como todos os atos da administração.

Princípio 2 – Impessoalidade. O princípio impõe ao gestor público que só pratique o ato para o seu objetivo legal, vedando qualquer prática de ato administrativo sem interesse público ou vantagem para a gestão. No caso dos dados abertos o gestor da instituição não pode interferir na priorização da publicação seguindo suas preferências ou interesses pessoais.

Princípio 3 – Moralidade: Trata de obedecer não somente à lei jurídica, mas também a lei ética da própria instituição. O gestor deve priorizar os dados a serem publicados visando beneficiar a sociedade sem censuras com relação ao desempenho da instituição ou de seus profissionais.

Princípio 4 – Publicidade: todo ato administrativo deverá ser publicado, com exceção dos que possuem sigilo nos casos de segurança nacional, investigações policiais ou de interesse superior da Administração. Este princípio é o que mais têm interseção com as diretrizes dos dados abertos e com a transparência na administração pública.

Princípio 5 – Eficiência: exige que a atividade administrativa seja exercida de maneira perfeita, com rendimento funcional. Sua relação com os dados abertos diz respeito à adequação orçamentária com as plataformas de hardware e software necessárias à publicação dos dados. Deve-se sempre buscar os melhores resultados com o mínimo custo possível.

Esses princípios impregnam os eixos sustentadores do Governo Aberto potencializando a importância dos dados abertos nesse cenário. Os eixos sustentadores são:

- **Transparência.** Trata da promoção da responsabilidade de informar os cidadãos sobre o que o governo está fazendo e que ações se pretende tomar.
- **Prestação de serviços.** Aprimora a eficácia do Estado, encorajando a cooperação entre a sociedade, os diferentes níveis de governo e a iniciativa privada.
- **Participação.** Este é o eixo que apresenta maior relação com os dados abertos. Permite aos cidadãos contribuir com suas ideias e competências, auxiliando o poder público a elaborar políticas eficazes e abrangentes, graças à informação amplamente dispersa em nossa sociedade.

Considerando-se os princípios e os eixos sustentadores faz-se necessário estabelecer instâncias de controle do processo de publicação de dados abertos pela instituição pública de modo a garantir que eles estão sendo cumpridos da melhor forma possível. Uma proposta já institucionalizada é a figura do Comitê Gestor.

Esse comitê materializa a função de governança dos dados abertos na instituição. Cabe lembrar aqui que o termo governança adotado neste texto é o adotado pelo Comitê de Governança da Internet (KURBALIJA 2016): a governança

dos assuntos de qualquer instituição, incluindo instituições não governamentais. Situando a governança no cenário específico dos dados abertos o Comitê Gestor para dados abertos na instituição deve ter como finalidade a gestão da política de dados abertos na mesma. Sua finalidade e composição devem ser definidas em regimento interno a ser definido pela instituição e deve ter representatividade nos diversos setores produtores de dados com potencial de serem abertos na instituição.

A título de referência sugere-se consultar a Instrução Normativa SLTI nº 4/2012²², que institui o comitê Gestor da INDA.

4.7 Usabilidade do Portal e de seu conteúdo

A usabilidade do Portal de dados abertos de uma instituição se traduz no quanto é fácil e intuitivo o uso deste portal como interface de acesso aos conjuntos de dados. Esta usabilidade é fator determinante no sucesso da iniciativa de abrir os dados e deve minimizar impactos negativos de uma infraestrutura tecnológica limitada e da necessidade de se publicar dados em quantidade e com a maior qualidade possível.

O ponto inicial do projeto do Portal é a especificação, com a máxima precisão, de quais são os recursos disponíveis para a plataforma dos dados abertos da instituição. Tais informações farão parte das variáveis que definirão o escopo da hospedagem dos dados e as ferramentas de acesso. Algumas limitações usuais são relativas ao espaço em servidores de dados, à disponibilidade dos servidores de aplicação, à banda de rede e à estabilidade dos sistemas de informação origem.

Outra limitação é a escolha do formato de publicação. Esta escolha, dependendo do volume de dados na origem, pode gerar arquivos muito grandes. A publicação da série temporal e o número de downloads simultâneos, no caso de arquivos muito grandes, podem causar sobrecarga nos servidores. Um exemplo

²² Disponível em <http://wiki.dados.gov.br/IN-SLTI-MP-04-INDA.ashx>. Acesso em 16/12/2019.

comum é a utilização de um arquivo XML ou RDF no lugar de um CSV. Esta escolha pode acarretar arquivos muito maiores entre formatos. A escolha do formato deve equilibrar as limitações de infraestrutura e o valor agregado desejado para os dados.

Os propositores das teorias de dados abertos não elencam os formatos permitidos ou proibidos, mas determinam que, para um dado ser considerado como dado aberto, o conjunto de dados deve estar disponível em um formato de especificação aberta, não proprietário, e estruturado, ou seja, que possibilite seu uso irrestrito e automatizado através da *Web*. Além disso, é imprescindível que seja utilizado um formato amplamente conhecido.

Um erro bastante comum cometido por diversas instituições é a publicação de relatórios em formato PDF. Este formato, não estruturado, inviabiliza ou no mínimo dificulta muito a reutilização daqueles dados.

Recomenda-se, como parte das boas práticas, alguns formatos não proprietários listados na Cartilha Técnica para Publicação de Dados Abertos no Brasil (Brasil, 2012c):

- JSON. É um acrônimo para *Java Script Object Notation*. É um padrão aberto de estruturação de dados baseado em texto e legível por humano. A especificação é a RFC 4627. JSON ganhou maior utilização com o advento do Ajax. A serialização em JSON é muito simples e resulta em uma estrutura pouco verbosa o que se mostra uma ótima alternativa ao XML. JSON possibilita serialização de estrutura de objetos complexos, como listas e subpropriedades. JSON está se tornando o padrão mais utilizado para integração de dados entre repositórios e frameworks, também está se tornando o padrão nativo de armazenamento em alguns bancos de dados modernos.
- XML. É um acrônimo para *Extensible Markup Language*. É um conjunto de regras para codificar documentos com estrutura hierárquica e em um formato legível por máquina. É baseado em texto e tem como principais objetivos simplicidade, extensibilidade e usabilidade. XML é largamente utilizado como formato de troca de dados nos clássicos Web Services SOAP. Possui uma ampla gama de ferramentas associadas, tais como o padrão XSLT que permite transformar para outra estrutura XML ou outro

formato. Apesar de sua ampla utilização, tem sido menos encorajada a utilização desse formato para integração de aplicações na Web, por utilizar mais recursos para transmissão e para o processamento do dados. Em substituição, recomenda-se utilizar JSON.

- CSV. É um acrônimo para *Comma-Separated Values*. São valores separados por vírgula. É um formato para armazenamento de dados tabulares em texto. A codificação é muito simples: cada linha do arquivo representa uma linha na tabela, e as colunas são separadas por vírgula. Campos que podem conter vírgula devem ser delimitados por aspas. CSV é recomendado para representação de estrutura de dados mais simples, de natureza tabular, onde não existem subpropriedades ou listas, gerando um arquivo menor e mais leve para processamento. Arquivos CSV são processáveis diretamente por editores de planilhas, como o OpenOffice e o MS Excel.
- ODS. É um acrônimo para *Open Document Spreadsheet*. Formato não proprietário de arquivo baseado em XML, padronizado pela ABNT sob a norma NBR ISO/IEC 26300:2006. É comumente chamado de planilha, similar ao XLS do MS Office Excel, porém aberto, por isso deve ser utilizado em substituição ao XLS. Planilhas são largamente utilizadas, sendo de fácil utilização e manipuláveis por diversos aplicativos. Apesar de ser um formato estruturado, é muito flexível, possibilitando manipulação e mistura de diversos tipos de dados, como imagens e textos formatados. Para a publicação de dados abertos tabulares, é recomendável a utilização de CSV, pela sua simplicidade e padronização.
- RDF. É um acrônimo para *Resource Description Framework*. É um modelo de dados estruturado em grafos e possui diversos formatos de serialização, tais como RDF/XML, Notation 3 e Turtle. Os formatos baseados em RDF têm seus dados descritos em vocabulários disponíveis na Web. Apesar da grande qualidade dos dados disponibilizados em RDF, a construção de vocabulários para seu uso não é trivial. Numa escala de níveis de qualidade/complexidade de dados abertos, o RDF está no último nível, onde se constituirá a Web semântica.

Outras recomendações sobre formatos²³ constam na Cartilha (Brasil, 2012c)

²³ Para saber mais sobre formatos abertos consulte a e-PING (Padrões de

na qual ainda são considerados como formatos válidos:

- **SVG** - *Scalable Vector Graphics*. Utilizado para dados vetoriais e geográficos.
- **GML** - *Geography Markup Language*. Utilizado para exprimir características e exportação de dados geográficos.
- **HTML/RDFa**. É a incorporação de marcações semânticas com uso de RDF sobre as tags HTML de uma página.
- **GTFS** - *General Transit Feed Specification*. Utilizado para dados de mobilidade e trânsito. Ver detalhes em <http://gtfs.org/>.

A instituição, ao escolher um padrão de publicação, deve analisar sua capacidade tecnológica atual e as limitações advindas das perspectivas futuras. A tendência, neste tipo de plataforma, é de crescimento constante na ocupação do espaço em disco e tráfego de rede intenso.

Tem-se como recomendação e uma boa prática não se publicar conjuntos de dados compactados. Entretanto, em casos de muitos arquivos ou de arquivos com grande capacidade de compressão recomenda-se a utilização de formatos abertos de compactação como o 7Z, TAR/GZIP ou ZIP.

Quanto às responsabilidades advindas do acesso para reuso dos dados, disponíveis no portal de dados abertos de uma instituição da administração pública, os usuários devem observar que (Brasil, 2012d):

- A geração de serviços secundários a partir do acesso ou manipulação dos dados gera para si a responsabilidade integral sobre as informações ou serviços oferecidos. O mesmo se aplica a quaisquer outras atividades baseadas nestes dados e exime a instituição publicadora de toda e qualquer responsabilidade referente às atividades, informações e serviços por ele desenvolvidos.

Interoperabilidade de Governo Eletrônico) em

<http://eping.governoeletronico.gov.br/>. Para a publicação de dados cartográficos e espaciais, consulte padronização na Infraestrutura Nacional de Dados

Espaciais – INDE em <http://inde.gov.br/>.

- Os usuários dos dados do Portal não devem apresentar os dados obtidos de forma incorreta ou enganosa, nem os utilizar para amparar ou sustentar atividades criminosas ou ilícitas ou de favorecimento ostensivo de atividades de igual caráter.
- A instituição publicadora poderá exigir que o usuário cancele a utilização, distribuição e/ou serviço de todo e qualquer conjunto de dados por ele utilizado se for verificada má-fé sem prejuízo da adoção das demais providências legais cabíveis.
- Os usuários dos dados do Portal ao gerarem produtos/serviços com os respectivos dados deverão se responsabilizar pelos conteúdos gerados e prestarem as devidas informações sobre os aplicativos, sistemas, estatísticas ou qualquer informação sobre o que foi produzido.

Por seu turno, a entidade publicadora, enquanto provedor dos dados disponíveis no portal:

- Não garante o fornecimento continuado dos dados nem a sua publicação tempestiva.
- Disponibiliza os dados conforme estão persistidos nos sistemas de informação origem.
- Exime-se de qualquer garantia e obrigação em relação aos dados e à utilização dos mesmos.
- Não será responsável pela utilização feita por pessoas e/ou instituições que transformem e/ou usem direta ou indiretamente os dados e que porventura venham a gerar prejuízos, danos econômicos ou materiais advindos desse uso e/ou transformação dos dados.

4.8 Protocolo de acesso

A definição do protocolo de acesso, do mesmo modo que a escolha do formato de publicação, deve ser precedida de uma análise cautelosa do cenário tecnológico da instituição, pois o esforço por parte do servidor e a banda de transferência de dados são fatores críticos.

Quando o conjunto de dados é naturalmente grande uma boa saída é gerar um pequeno arquivo acessível on-line no portal e o uso de arquivos compactados. Uma outra alternativa é o uso de Torrents, baseado no compartilhamento distribuído dos

arquivos em que o esforço de acesso ao servidor é distribuído com outras pessoas que já fizeram download desses arquivos.

São exemplos de protocolos: HTTP, FTP, Endpoint SPARQL e Bit Torrent.

4.9 Cobertura dos conjuntos de dados

Ao se projetar a cobertura geográfica ou temporal de um conjunto de dados deve-se considerar a eficiência vinculada ao uso de um arquivo do tamanho que será gerado. Caso o uso seja sempre de parte reduzida do todo há que se considerar a opção por arquivos menores compondo o todo.

Um exemplo é a busca por uma escola em uma região de uma cidade. Se o arquivo trata o estado inteiro, a região metropolitana ou a cidade pode ser que para a utilização seja mais fácil para o usuário acessar arquivos segmentados por bairro. Por outro lado, pode ser trabalho para a instituição gerar 400 arquivos de bairro para busca de escolas. Assim, a decisão da cobertura considerada em um pacote é uma decisão estratégica e influi diretamente no tamanho dos conjuntos de dados, sua organização no portal e na usabilidade por parte do usuário.

No portal data.gov dos Estados Unidos da América e no portal data.gov.uk do Reino Unido é usual a distribuição em arquivos por segmento de cobertura geográfica e/ou temporal. Esta organização é sempre informada nos metadados para orientar os usuários e foram adotadas após a percepção de vantagens para o usuário e para a instituição. Cabe destacar que cada instituição deve fazer uma análise individualizada e encontrar outras dimensões relacionadas à natureza do dado gerado ou manipulado.

Exemplos de dimensões de cobertura:

- Geográfica no país: UF, estado, município, regional, bairro.
- Temporal: anual, trimestral, mensal, semanal, diário.

4.10 API – A difícil decisão

Uma alternativa que tem o potencial de maximizar o acesso aos dados publicados no servidor do Portal de Dados Abertos da instituição é o

desenvolvimento de uma API (*Application Programming Interface*) para liberar acesso de informações na Web. Uma API, resumidamente, é uma camada de interação entre uma base de dados e um aplicativo que se alimenta desses dados. A API oferece ao desenvolvedor/ empreendedor usuário um conjunto de chamadas padrão que ele pode utilizar para acessar dados de uma base de dados utilizando requisições na Web para a API.

Apesar de aparentar ser um cenário inofensivo e com grande chance de maximizar a utilização de uma base de dados no servidor de dados abertos, há diversos problemas neste contexto. O desenvolvimento de uma API requer conhecimento técnico apurado e, caso ela seja pública, a definição de comportamentos que tentam antever os casos em que o desenvolvedor/empreendedor usuário precisará dos dados pode ser desafiante.

Uma API apresenta diversas vantagens, como o acesso facilitado e rápido às bases de dados abertas. Em lugar de baixar a base de dados inteira, basta que o programador faça uma chamada simples na Web para extrair a porção que é de seu interesse naquele momento. Ela também agiliza o acesso em tempo real a determinadas partes da base de dados, potencializando o desenvolvimento de aplicativos que dependem de dados atualizados rapidamente.

As APIs podem ser:

- Privadas. É o caso de um desenvolvedor que tenha controle sobre o banco de dados e cria a API para facilitar o acesso aos dados.
- Públicas. Caso em que o responsável por uma base de dados desenvolve uma API para servir à comunidade de desenvolvedores/empreendedores da sociedade. Neste caso é necessário projetar tipos diversos e genéricos de requisições ao banco de dados que possam atender o maior número de aplicações possíveis. Serviços como o Facebook e o Twitter são exemplos de empresas que disponibilizam APIs públicas que promovem a interação, de forma limitada, entre as bases de dados e os programadores de todo mundo.

No caso das instituições governamentais a disponibilização de APIs pode

provocar diversas situações indesejáveis. É preciso analisar detalhadamente se o desenvolvimento de uma API é a melhor opção.

Tomemos como exemplo fictício o caso do *Lincoln Tunnel*²⁴ em Nova York. Este túnel é uma via muito utilizada por veículos de transporte de pessoas e carga entre Manhattan e Nova Jersey. No nosso exemplo A prefeitura de Nova York disponibiliza, em seu Portal de Dados Abertos, o volume de tráfego, minuto a minuto, na região de acesso ao túnel através de uma API pública para acesso a dados sobre as condições das pistas de acesso, volume de tráfego e possíveis interrupções no acesso ao túnel. As informações são utilizadas por empresas que fornecem serviços, via aplicativo, informando sobre tempo de acesso ao túnel e outras análises a partir das informações disponíveis pelas chamadas da API.

O servidor de API teve sua performance muito exigida em várias ocasiões durante os primeiros meses após sua implantação, mas estava sob controle. Entretanto, no primeiro grande feriado, um dia de ação de graças, teve início uma série de sobrecargas no servidor de API causando o travamento dos servidores de banco de dados e a interrupção dos serviços que dependem desse banco de dados. A crise continuou até meados de janeiro quando os analistas concluíram que o problema era causado por serviços das empresas usuárias que haviam incorporado previsões em relação ao trânsito na região do túnel com até 5 horas de antecedência. Nesta situação cada aplicativo instalado em cada celular, realizando várias chamadas de API para acesso aos dados, compunha um enorme conjunto de acessos pela manhã e ao final da tarde. Ou seja, não havia como prever uma infraestrutura tão grande para este número de acessos. E mesmo se fosse previsto o orçamento seria exageradamente alto.

A solução para este caso fictício, e vários outros com o mesmo perfil, não foi

²⁴ Lincoln Tunnel é um túnel de 2,4 km sob o rio Hudson, ligando Weehawken, Nova Jersey, na margem oeste, a Midtown Manhattan, na cidade de Nova York, na margem leste.

ampliar o número de servidores, mas sim desativar as APIs e disponibilizar as informações de tráfego em arquivos .CVS. Com esta solução a sobrecarga foi transferida para o prestador de serviços que teve que alterar sua plataforma baixando o arquivo para seu ambiente e fornecendo o acesso aos usuários do aplicativo direto de seus servidores.

Como dito anteriormente as APIs apresentam muitas vantagens, mas podem provocar situações indesejáveis.

5 Roteiro para elaboração do Plano de Dados Abertos

“O planejamento não diz respeito a decisões futuras, mas às implicações futuras de suas decisões presentes.”

Peter Drucker.

O Plano de Dados Abertos (PDA) tem por função registrar o planejamento de publicação dos dados no Portal de Dados Abertos da instituição. Sua elaboração vem ao encontro do disposto na Lei de Acesso à Informação, na Instrução Normativa SLTI nº 4, de 13 de abril de 2012, que institui a Infraestrutura Nacional de Dados Abertos, bem como os compromissos assumidos pelo Brasil no âmbito do Plano de Ação Nacional sobre Governo Aberto, entre outros normativos que abordam o tema de transparência.

Este capítulo foi elaborado com base no Manual para Elaboração de Plano de Dados Abertos (Brasil, 2014). Recomenda-se que os órgãos construam seus Planos envolvendo todas as suas unidades, considerando os princípios da publicidade e da transparência da administração pública. A priorização dos conjuntos de dados que serão publicados, deve considerar: o grau de relevância das informações para o cidadão, os compromissos anteriormente assumidos pelo órgão, o alinhamento com os instrumentos de planejamento estratégico, as competências e os conjuntos de dados efetivamente mantidos pela instituição. O plano tem periodicidade bienal com revisão semestral e recomenda-se a sua publicação no Portal da instituição publicadora.

Dentre os elementos que devem estar presentes no plano:

- Objetivos (gerais e específicos).
- Legislação e demais normativos aplicáveis.
- Cenário institucional do órgão (demonstrar o alinhamento com compromissos assumidos internamente, perante a sociedade ou outros entes).
- Metodologia de construção e validação do documento.
- Canais de comunicação.

- Formas de interação com a sociedade.
- Cronograma com prazos e responsabilidades.
- Definição dos dados a serem abertos.
- Critérios utilizados para priorização dos dados.
- Estratégia definida para abertura dos dados, com respectivo plano de ação.
- Modelo de Sustentação (de modo a perenizar o fluxo de atualização e manutenção dos dados).
- Estrutura de Governança, forma de monitoramento e controle.
- Canais de Comunicação e Participação Social.
- Metas de melhoria contínua.
- Referências (bibliografia e legislação) e Glossário.

Sugere-se a seguinte estrutura para o PDA.

- Capa
- Introdução
- Contexto institucional
- Objetivos
- Processo de Abertura de dados
- Governança
- Planos de ação para abertura de dados
- Divulgação

Objetivando facilitar a elaboração do Plano de Dados Abertos o texto a seguir seguintes detalha as seções do PDA, seguindo como modelo o PDA do Ministério do Planejamento, Orçamento e Gestão, podendo ser complementado quando necessário pela instituição. O detalhamento conta com uma explicação e um texto exemplificativo da seção. **Para dar sentido aos exemplos e sugestões de textos para as seções do PDA exemplo tomou-se a Prefeitura de Belo Horizonte e sua estrutura hierárquica como estudo de caso.**

5.1 Capa

Na capa do PDA devem constar:

- O nome da instância de governo. Ex.: Prefeitura Municipal de Belo Horizonte.
- Instituição da Administração Pública. Ex.: Secretaria Municipal de Cultura.
- Vigência. Ex.: 2019-2020.
- Data de elaboração do documento.
- Informações sobre versionamento.

5.2 Introdução

5.2.1 Orientação

Nesta seção, deve-se informar que o texto trata do Plano de Dados Abertos da instituição, apresentando a legislação aplicável e os objetivos do documento. É recomendável mencionar como se deu o processo de construção do PDA e apresentar a estrutura do PDA. Sugere-se um texto como este a seguir.

Este texto apresenta o Plano de Dados Abertos - PDA da Secretaria Municipal de Cultura, integrante da estrutura administrativa da Prefeitura de Belo Horizonte - PBH. O PDA contempla o planejamento para publicação de dados, em formato aberto²⁵, de dados públicos²⁶ produzidos ou custodiados nos sistemas de informação da PBH. Além disto, o PDA organiza o planejamento dos processos de publicação e da manutenção de dados abertos da PBH e seus órgãos, orientando os servidores envolvidos nessas atividades.

²⁵ São dados públicos representados em meio digital, estruturados em formato aberto, processáveis por máquina, referenciados na rede mundial de computadores e disponibilizados sob licença aberta que permita sua livre utilização, seu consumo ou cruzamento.

²⁶ Dado público é qualquer dado gerado ou sob a guarda governamental que não tenha o seu acesso restrito por legislação específica.

O PDA é um dos artefatos publicados no Portal de Dados Abertos da PBH.

5.3 Cenário institucional

5.3.1 Orientação

Esta seção trata da avaliação e do detalhamento de elementos internos e externos que impactam a política de abertura de dados da instituição, como os instrumentos e instâncias de Gestão a seguir:

- Planejamento Estratégico Institucional (PEI).
- Planejamento Estratégico de Tecnologia da Informação (PETI).
- Plano Diretor de Tecnologia da Informação (PDTI).
- Estratégia Geral de Tecnologia da Informação (EGTI).
- Plano de Ação da Infraestrutura Nacional de Dados Abertos – INDA.
- Normas da Comissão Nacional de Cartografia – CONCAR - relacionadas à Infraestrutura Nacional de Dados Espaciais – INDE.
- Comitê Interministerial Governo Aberto – CIGA.
- Plano de Ação Nacional sobre Governo Aberto - contexto Parceria para Governo Aberto – *Open Government Partnership*.
- Comitê Estratégico de Tecnologia da Informação – CETI (quando houver).
- Outros instrumentos, políticas e instâncias colegiadas relevantes para o órgão.

Para cada um deles, quando aplicável, recomenda-se explicitar o impacto para o PDA da instituição, metas e indicadores relacionados aos temas abertura de dados e transparência ativa e como se dá o alinhamento com o instrumento. Sugere-se um texto como este a seguir.

O cenário tecnológico de Belo Horizonte, em que ela se prepara para ser uma referência dentre as cidades inteligentes, tem especial potencial de sucesso para as iniciativas que culminam na publicação de dados acessíveis aos cidadãos.

No momento a PBH já conta com um Portal de Dados Abertos disponível na Internet em <http://dados.pbh.gov.br/> que contempla, até o momento, 9 áreas temáticas e 7 secretarias da organização administrativa da PBH. Neste Portal estão disponíveis

70 conjuntos de dados de elevado interesse do cidadão.

O Plano Diretor de Tecnologia da Informação – PDTI é outro norteador importante. Em seu mapa estratégico constam a definição de missão, visão e as perspectivas que contribuem para a consecução da missão. Essas perspectivas estão organizadas em quatro níveis interligados: impactos para o órgão, entregas da TI, processos internos e sustentação. O PDA está diretamente ligado ao “impacto para o órgão” em relação a aperfeiçoar a gestão da informação e do conhecimento e os mecanismos de comunicação e transparência e à “entrega de TI” por aprimorar os meios eletrônicos de fornecimento de informações internamente e com a sociedade.

Institucionalmente o PDA objetiva materializar o compromisso do órgão com a sociedade onde se informa o planejamento em relação a quais dados serão abertos em que tempo.

5.4 Objetivos

5.4.1 Orientação

Esta seção apresenta os objetivos principais e específicos do PDA, tais como:

- Identificar prioridades e disponibilizar dados em formatos abertos, e, sempre que possível, georreferenciados.
- Melhorar a qualidade dos dados disponibilizados.
- Estimular a interoperabilidade de dados e sistemas governamentais pela publicação de dados em formato processável por máquina, conforme padrões estabelecidos.
- Melhorar a gestão da informação e de dados.
- Incrementar os processos de transparência e de acesso a informações públicas.
- Estimular a visualização da informação das ações de governo no território.

Sugere-se um texto como este a seguir.

Tem-se como objetivo Geral do PDA a promoção da abertura de dados na PBH, zelando pelos princípios da publicidade, transparência e eficiência, visando ao aumento da disseminação de dados e informações para a sociedade, bem como a

melhoria da qualidade dos dados disponibilizados, de forma a dar maior suporte à tomada de decisão pelos gestores públicos e ao controle social.

No desenvolvimento do PDA, a PBH buscará sempre:

- *Identificar prioridades e disponibilizar dados em formatos abertos;*
- *Melhorar a qualidade dos dados disponibilizados;*
- *Estimular a interoperabilidade de dados e sistemas governamentais pela publicação de dados em formato processável por máquina, conforme padrões estabelecidos;*
- *Melhorar a gestão da informação e de dados;*
- *Incrementar os processos de transparência e de acesso a informações públicas.*

5.5 Processo de Abertura de dados

5.5.1 Orientação

Nesta seção, recomenda-se detalhar o passo-a-passo relacionado ao processo de publicação de dados, bem como premissas e requisitos de qualidade mínimos relacionados.

A abertura de dados deve se orientar pelos seguintes passos:

1. Levantamento do conjunto de dados candidato à abertura.
2. Priorização e seleção dos dados que serão abertos.
3. Definição de responsáveis pelo preparo e atualização dos dados e detalhamento de plano de ação com metas e prazos.
4. Consolidação da matriz de responsabilidades e definição da governança e do fluxo de aprovação do PDA e revisões.
5. Utilização de metodologia de abertura de dados a ser seguida pelas áreas responsáveis (padrões INDA e INDE).
6. Capacitação dos responsáveis nas áreas de negócio dos dados selecionados para abertura sobre:
 - Processo de publicação de dados abertos;
 - Processo de catalogação dos metadados no dados.gov.br;
 - Processo de catalogação dos metadados na INDE, caso georreferenciados;
 - Definição de arquitetura de abertura para cada sistema.

- Publicização dos dados catalogados, observando-se o uso de URL fixa.

Além de estar alinhado aos princípios e diretrizes mencionados anteriormente neste PDA, o processo de abertura dos conjuntos de dados deve considerar as seguintes premissas:

1. Publicar os dados considerados relevantes para a sociedade o mais rápido possível, no formato disponível e informando as eventuais limitações de qualidade dos dados.
2. Sempre que possível, publicar dados e seus metadados conforme estabelecido no Plano de Ação da INDA, que institui que cada conjunto de dados deve conter, no mínimo:
 - a. Nome ou título do conjunto de dados;
 - b. Descrição sucinta;
 - c. Palavras-chave (etiquetas);
 - d. Assuntos relacionados do VCGE - Vocabulário Controlado do Governo Eletrônico;
 - e. Nome e e-mail do setor responsável pelos dados;
 - f. Periodicidade de atualização;
 - g. Escopo temporal (anual, mensal, diário, bimestral etc. exemplo: dados referentes ao censo de 2011, dados de um indicador mensal);
 - h. Escopo geopolítico (por cidade, por estado, por região).
3. No caso de dados Georreferenciados, deve-se levar em conta as normas e padrões da INDE. A seguir, sumário dos elementos do Núcleo de Metadados do perfil de Metadados Geoespaciais do Brasil (MGB) conforme Tabela 7.

Dados Abertos: transparência governamental
 Perspectivas, cenários e planejamento

Entidade / Elementos	Obrigatoriedade	Entidade / Elementos	Obrigatoriedade
1. Título	obrigatório	12. Tipo de Representação Espacial	opcional
2. Data	obrigatório	13. Sistema de Referência	obrigatório
3. Responsável	obrigatório	14. Linhagem	opcional
4. Extensão Geográfica	condicional	15. Acesso Online	opcional
5. Idioma	obrigatório	16. Identificador Metadados	opcional
6. Código de Caracteres do CDG	condicional	17. Nome Padrão de Metadados	opcional
7. Categoria Temática	obrigatório	18. Versão da Norma de Metadados	opcional
8. Resolução Espacial	opcional	19. Idioma dos Metadados	condicional
9. Resumo	obrigatório	20. Código de Caracteres dos Metadados	condicional
10. Formato de Distribuição	obrigatório	21. Responsável pelos Metadados	obrigatório
11. Extensão Temporal e Altimétrica	opcional	22. Data dos Metadados	obrigatório
		23. Status	obrigatório

Tabela 7- Entidades e elementos do Núcleo de Metadados do Perfil MGB Sumarizado.

4. Publicar os dados seguindo os padrões definidos pela e-PING, pela Infraestrutura Nacional de Dados Abertos - INDA e pela Infraestrutura Nacional de Dados Espaciais – INDE e Governo Eletrônico, quando aplicável;
5. Catalogar os dados abertos da instituição no Portal Brasileiro de Dados Abertos, ponto central de acesso aos dados do Governo Federal, quando aplicável;
6. Catalogar os dados geoespacializados na INDE, quando aplicável;
7. Promover a integração entre os catálogos de metadados INDA e INDE, quando aplicável;
8. Manter os dados publicados atualizados e sincronizados com a origem, com a menor periodicidade e maior granularidade viáveis;
9. A atualização dos dados deve ocorrer, preferencialmente, por meio de sincronização automática, estabelecendo-se um processo contínuo, especialmente no caso de sistemas estruturantes, com ganhos de eficiência em comparação a extrações pontuais;
10. Utilizar, como forma de disseminação, os ambientes do Portal Brasileiro de Dados Abertos do Governo Federal, quando aplicável.

Sugere-se um texto como este a seguir.

O processo de abertura de dados é um processo que tem início nas análises dos dados disponíveis, passa pela sua formatação para publicação e a publicação

no Portal de Dados Abertos da PBH. A Figura 10 ilustra o processo.

Para se estabelecer as metas de qualidade de dados e etapas de priorização, considera-se os seguintes critérios:

- *Publicidade dos dados;*
- *Grau de relevância para o cidadão;*
- *Alinhamento com o Planejamento Estratégico da PBH;*
- *Conjunto de informações e sistemas desenvolvidos e sob a gestão da Prodabel;*
- *Nível de maturidade²⁷ das informações e dados existentes.*

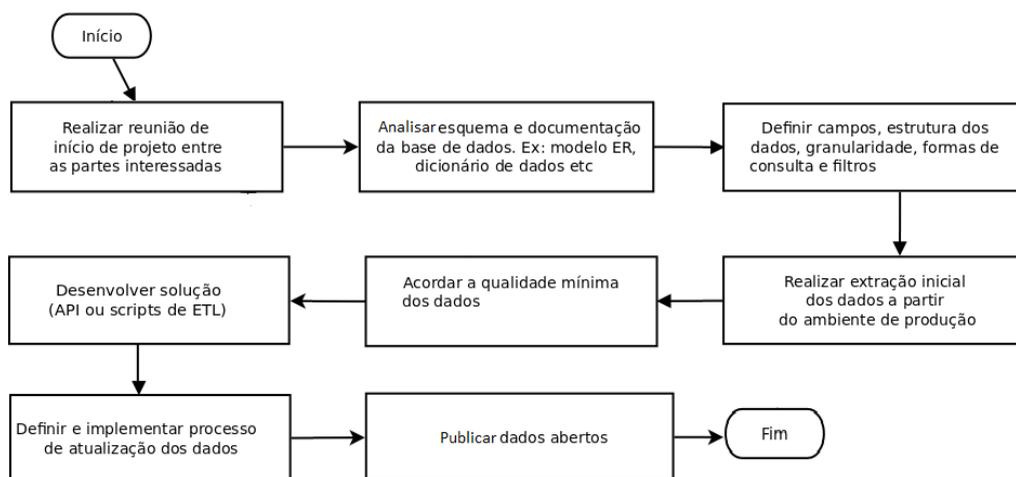


Figura 10 – Processo de abertura de dados

Para o período AAAA²⁸-AAAA²⁹, serão inicialmente abertos os dados que já possuem estrutura e maturidade para abertura. Ademais, outras ações devem ser

27 Níveis que representam os diferentes estados pelos quais uma organização passa à medida que amadurece no conhecimento, organização, qualidade uso e reuso de seus dados. Seguem o Guia de Maturidade de Dados Abertos disponível em <http://www.governoaberto.sp.gov.br/wp-content/uploads/2016/02/Book-Web-Modelo-de-Maturidade-de-Dados-Abertos.pdf>.

28 Inserir o ano inicial do período de vigência do PDA.

29 Inserir o ano final do período de vigência do PDA.

realizadas para garantir a qualidade e o aprimoramento dos conjuntos de dados abertos, como:

- *Levantamento, junto às unidades organizacionais, de todos os conjuntos de dados mantidos pela **PBH**;*
- *Promoção de eventos sobre Dados Abertos, sob responsabilidade da **SUCOM**.*
- *Capacitação, sobre abertura de bases de dados, dos servidores envolvidos nas atribuições do **PDA**;*

As bases de dados divulgadas em formato aberto deverão ser mantidas atualizadas e sincronizadas com a origem primária das informações, com a menor periodicidade e o maior grau de granularidade viáveis. A atualização dos dados se dará, preferencialmente, por meio de sincronização automática, estabelecendo-se um processo contínuo.

*As unidades organizacionais responsáveis pelas bases de dados cuja abertura seja considerada prioritária, de grande interesse para a sociedade ou de utilidade pública, mas que não sejam consideradas com grau de qualidade suficiente, deverão apresentar proposta com as ações de melhoria que se fizerem necessárias, a constar nos Planos de Ação do **PDA**. Devem ser privilegiadas estratégias evolutivas de aprimoramento da qualidade dos dados, para viabilizar o quanto antes a divulgação em formato aberto.*

*Com o intuito de aprimorar constantemente o **PDA**, será realizada sua atualização, ao final do seu período de vigência, com a definição e priorização de novos conjuntos de dados a serem abertos.*

*O controle de qualidade dos dados publicados no Portal da Dados Abertos da **PBH** deve ainda:*

- *Verificar, para efeitos de publicação, se os dados estão de acordo com os padrões da INDA30 e INDE31 em relação a metadados32;*
- *Contatar o responsável pelos dados, caso se verifique que algum dos arquivos catalogados se tornou indisponível;*
- *Identificar e elaborar propostas para possíveis melhorias de qualidade dos dados disponibilizados e novos conjuntos de dados candidatos a abertura de dados.*

Cabe aqui destacar que os responsáveis pela qualidade dos dados abertos devem informar a acurácia e qualidade das informações nos metadados respectivos.

Suas atividades incluem:

1. Atividades de atualização:

- *Garantir a notificação aos usuários sempre que houver atualização ou novas publicações;*
- *Analisar se os dados a serem atualizados ou publicados estão de acordo com os padrões da INDA e INDE;*
- *Testar a usabilidade dos arquivos publicados e notificar os responsáveis por estes arquivos quaisquer problemas de publicação ou de atualização.*

2. Atividades de manutenção:

- *Informar aos gestores do Portal casos de indisponibilidade ou erros nos arquivos catalogados;*

³⁰Infraestrutura Nacional de Dados Abertos. Hospedado em <https://www.governodigital.gov.br/eixos-de-atuacao/dados-abertos/inda-infraestrutura-nacional-de-dados-abertos>

³¹ Infraestrutura Nacional de Dados Espaciais. Hospedado em <http://www.inde.gov.br/geoservicos.html>.

³² Informação que descreve características de determinado dado, explicando-o em certo contexto de uso.

- *Identificar e elaborar propostas para melhorias de qualidade dos dados disponibilizados;*
- *Analisar as solicitações de informações mais recorrentes nos canais de comunicação com os usuários dos dados abertos e sugerir novos conjuntos de dados candidatos à abertura de dados;*
- *Prever formas para garantir a disponibilidade dos dados abertos, bem como a incorporação do processo de abertura dentro da rotina do órgão da PBH;*
- *Garantir não só a sustentabilidade, mas a melhoria contínua do processo de publicação e atualização.*

5.6 Governança

5.6.1 Orientação

Nesta seção, deve-se apresentar de que forma está estruturada a governança do PDA, com papéis e responsabilidades definidas para os principais atores envolvidos.

Sugere-se um texto como este a seguir.

A abertura de dados em instituições governamentais é uma ação interdepartamental que envolve alta gestão, área de gestão das informações, áreas de negócio, entre outras. Essas diversas áreas precisam interagir, de modo que todo o processo seja coerente e contínuo. Além disso, programas de abertura de dados podem implicar em mudanças organizacionais, culturais, e, por fim, tecnológicas.

Nesse contexto, a fim de que a abertura de dados governamentais seja um processo perene e sustentável no âmbito da PBH³³ é importante que a alta administração tenha governança sobre as ações de abertura de dados.

³³ PBH aqui representa o órgão. No caso este exemplo seria referente à Prefeitura de Belo Horizonte.

De acordo com Referencial Básico de Governança Aplicável a Órgãos e Entidades da Administração Pública, elaborado pelo TCU³⁴, “a alta administração da organização possui papel fundamental na divulgação de dados públicos de maneira ampla e aberta. Não é sem motivo que boas práticas apontam o apoio da alta administração como um fator chave para que programas de abertura de dados operem de maneira sustentável e sejam efetivos. A instituição precisa perceber que a alta administração considera a abertura de dados importante e que está disposta a investir os recursos necessários para viabilizá-la”.

A estrutura de governança dos dados abertos proposta para a PBH envolve cinco atores com papéis e responsabilidades definidos:

- *Comitês Gestores, responsáveis pelas diretrizes mestras da política municipal de abertura de dados;*
- *Secretaria Executiva, incumbida da governança dos dados abertos no âmbito da PBH;*
- *Força Tarefa para análise e publicação. A qual se atribui a manutenção e revisão dos conjuntos de dados abertos;*
- *Força Tarefa para Curadoria, encarregada pela curadoria dos metadados;*
- *Áreas Departamentais responsáveis pelos dados e suas regras negociais.*

Por meio da governança, a Secretaria Executiva define diretrizes sobre a publicação dos dados oficiais, estabelece responsabilidades ligadas à abertura de dados, realiza o acompanhamento das ações de abertura de dados avaliando a

34

efetividade dessas ações e procedendo ajustes, quando necessários, para garantir a obtenção dos resultados esperados e administra os recursos para execução de ações ligadas à abertura de dados e gerir as atividades das forças tarefas.

Permanecerá sob a responsabilidade da Secretaria Executiva a função de aprovar, deliberar e acompanhar a execução do PDA no nível estratégico, analisando indicadores, acompanhando as atividades do plano de ação e aprovando seus produtos. Também será seu dever manter o alinhamento do PDA com os instrumentos de planejamento aplicados no âmbito da PBH, em colaboração com as Áreas Departamentais afetadas, persistindo também a incumbência de acompanhar o PDA quanto aos prazos, cronogramas e gestão da comunicação entre os envolvidos.

As Áreas Departamentais são proprietárias e responsáveis pelos conjuntos de dados e portanto, encarregadas de levantar e analisar suas informações, coordenar os processos de abertura, prezando pelo cumprimento do PDA podendo também, propor alterações a este e, prezar pela melhoria contínua da qualidade e confiabilidade dos dados publicados.

O PDA da PBH tem como referência, quanto a melhoria da qualidade dos dados abertos, o documento intitulado Modelo de Maturidade de Dados Abertos³⁵ da INDA que institui os seguintes critérios:

- *Os dados abertos devem conter a possibilidade de serem acessados diretamente, por meio de URL única, ou seja, passível de ser reproduzida e compartilhada, sem necessidade de navegação na página para seu acesso. No caso da PBH a URL única é dados.pbh.gov;*
- *Tabelas mantidas em arquivos PDF (relatórios, por exemplo), devem estar contidas também em arquivos próprios para sua estruturação (como csv), e serem referenciadas por esses relatórios;*

³⁵ Disponível em <http://ceweb.br/publicacao/modelo-de-maturidade-de-dados-abertos/>

- *Dados disponibilizados devem ser gerados em formatos abertos conforme a Cartilha de Publicação de Dados da INDA³⁶;*
- *Os dados publicados devem conter um conjunto mínimo de metadados, ainda conforme a Cartilha de Publicação de Dados da INDA.*

5.7 Planos de ação para abertura de dados

5.7.1 Orientação

Nesta seção, recomenda-se detalhar as atividades e responsabilidades relacionadas ao processo de publicação de dados, bem como prazos e requisitos. Recomenda-se ainda prever as ações com um foco mais macro, que pode ser detalhado em cronograma específico de cada unidade da estrutura da instituição.

Sugere-se um texto como este a seguir.

A gestão eficiente de dados requer muito mais do que investimento em tecnologia. Envolve processos e profissionais capazes de administrar todos os aspectos do ciclo de vida da informação. A gestão é também responsável pela maneira como os dados são criados, armazenados e utilizados pela organização.

Em organizações complexas como a PBH a gestão da informação torna-se um desafio ainda maior quando se analisa a integração de dados entre seus sistemas de informação. No caso de um contexto novo como os Dados Abertos, é necessário utilizar-se de um recurso técnico que agilize e reduza a complexidade das análises necessárias ao resultado desejado.

O Plano de Ação contém uma série de ações para a construção de todos os pilares do conteúdo informacional do Portal de Dados Abertos e é pensado de acordo com o conjunto de dados a ser gerado por cada órgão da PBH.

O PDA é centrado na ideia de áreas temáticas em que para cada área se

³⁶ Disponível em <http://dados.gov.br/pagina/cartilha-publicacao-dados-abertos>

identifica as bases de dados relevantes nos órgãos participantes da área temática na PBH. As demais informações cumprem as funções de responsabilização e de estruturação da base de dados na área temática.

Uma área temática reúne informações que se completam, que formam uma unidade coesa e seguem determinada política de produção, gestão e preservação de dados. São exemplos de áreas temáticas a educação, a saúde, o desenvolvimento humano e a administração. A relação completa de áreas temáticas com seus textos explicativos está no Anexo 1. Assim, o PDA de uma área temática da PBH é o conjunto de Planos de Ação para abertura de dados executados pelos órgãos integrantes desta área temática no período de vigência do PDA da PBH. Por seu turno o PDA da PBH é o conjunto de PDAs de todas as áreas temáticas.

Esta composição está apresentada na Figura 11. O template do plano de ação, que contempla uma matriz de responsabilidades, está apresentado na Figura 12. O Anexo 2 apresenta exemplos de planos de ação preenchidos.

Dados Abertos: transparência governamental
 Perspectivas, cenários e planejamento

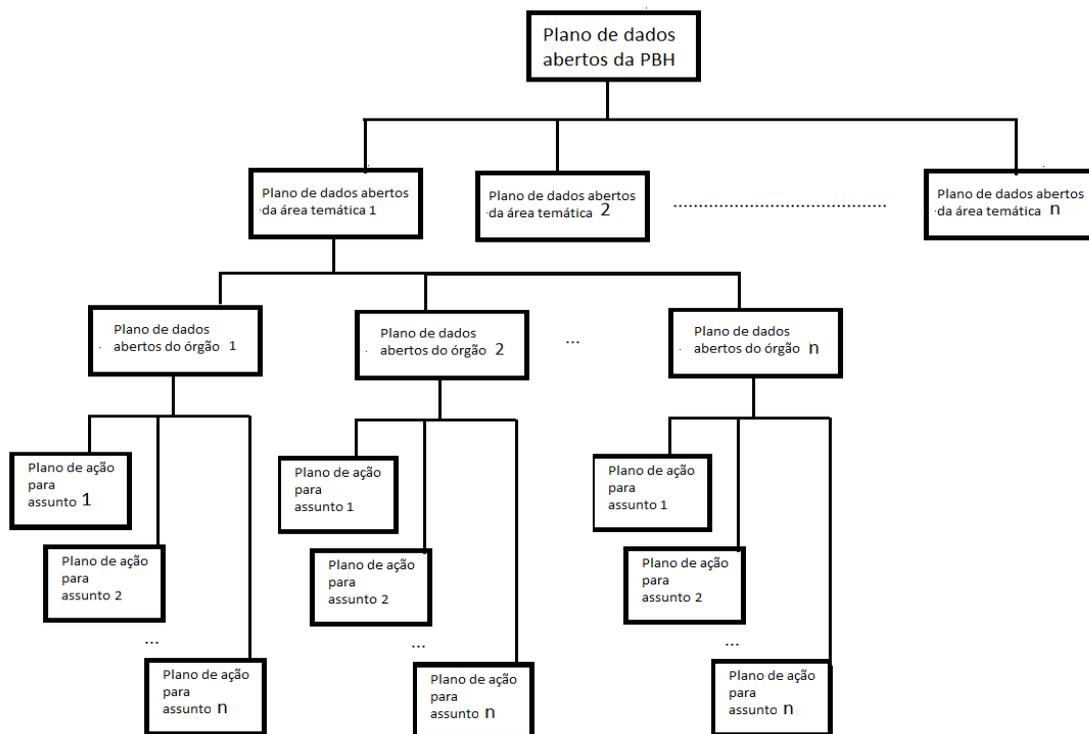


Figura 11 - Distribuição de planos de ação em planos de dados.

<i>Atividade</i>	<i>Metas</i>	<i>Prazo</i>	<i>Responsável</i>
<i>1. Estruturar abertura de dados</i>	<i>1. Identificar os assuntos dos dados a serem abertos.</i>		
	<i>2. Identificar os dados candidatos a serem abertos no assunto.</i>		
	<i>3. Identificar o esforço e os recursos necessários para construir as rotinas de extração de dados e publicação.</i>		
	<i>4. Definir os responsáveis pela qualidade dos dados.</i>		
<i>2. Construção dos dados e</i>	<i>1. Analisar os dados com base na precisão, acurácia, frequência de atualização e formato.</i>		

dicionário de dados	2. <i>Especificar as rotinas de extração de dados.</i>		
	3. <i>Estruturar os Metadados.</i>		
	4. <i>Aprovar conjuntos de dados e Metadados.</i>		
	5. <i>Atualizar o Portal de Dados Abertos da PBH.</i>		
3. <i>Automação de ambiente para leitura e geração dos arquivos</i>	1. <i>Definir como será a extração automática dos dados.</i>		
	2. <i>Construir rotinas necessárias.</i>		
	3. <i>Especificar plano de contingência em caso de erro na geração dos dados.</i>		
4. <i>Publicação dos dados no Portal de Dados Abertos da PBH</i>	1. <i>Publicar os dados e Metadados.</i>		
	2. <i>Realizar o treinamento para os envolvidos.</i>		
	3. <i>Revisar o PDA para validar o que foi feito.</i>		

Figura 12 – Matriz de responsabilidades

5.8 Divulgação.

As exigências legais são um forte incentivo para a publicação dos dados da administração pública, mas pode não ser o bastante para que resultem em algo útil para a sociedade. As instituições que publicam dados abertos podem promover ou participar de ações para encorajar outros a criarem coisas úteis e interessantes com esses dados. É válido que a instituição tome iniciativas próprias do mundo do marketing, tais como:

- Estabelecer parcerias com organizações da sociedade civil para ampliar sua base de participantes e aumentar sua exposição na Internet e na mídia.
- Publicar informativos em listas de e-mail e grupos em redes sociais de comunidades de desenvolvedores que trabalham com tecnologias abertas.
- Organizar ou participar de eventos também pode ser interessante para contatar pessoas interessadas neste tipo de dado. *Barcamp, meetup,*

speedgeek e lightning talk são exemplos deste tipo de evento. Estes eventos devem privilegiar a participação dos cidadãos, de hackers e ativistas e não de outros membros de governo.

- Contatar organizações ou indivíduos que trabalham com, ou estão interessados, na área dos dados abertos.

Cabe destacar que os empresários e as comunidades que utilizam ou que promovem a reutilização de dados da administração pública têm sido indispensáveis na garantia de que os dados sejam política e tecnicamente úteis no impulsionamento de inovações em torno do tema.

Outra forma muito eficiente para nivelar conhecimentos e divulgar os dados abertos são os eventos temáticos que têm como foco o desenvolvimento de aplicativos baseados nos dados. Os eventos geram oportunidades de parcerias por várias razões. Dentre elas:

- É possível ampliar a divulgação das realizações de transparência da instituição pela repercussão nas redes sociais dos participantes.
- Permite conhecer os potenciais usuários dos dados.
- Permite identificar a demanda por diferentes conjuntos de dados, publicados ou não.
- Gera discussões sobre as tendências de desenvolvimento de aplicativos utilizando os dados.
- Configura uma forma de encontro entre prováveis usuários.
- Permite expor os dados para uma ampla audiência (por meio dos blogs ou da cobertura de mídia gerada por conta do evento).

Estes eventos podem ser organizados em diversos formatos desde palestras até demonstrações. Uma das formas atuais com melhor resultado é aquela dirigida pelos próprios participantes. Sua efetividade tem relação direta com o conceito de dados abertos, pois mescla participação e autonomia que, ao fim e ao cabo, são as premissas principais da abertura de dados. Nesses eventos os participantes devem ter oportunidade para:

- Propor os resultados esperados no evento.
- Definir a agenda do evento.

- Apresentar, seu trabalho e no que estão interessados.
- Organizar grupos de trabalho.
- Liderar as equipes.
- Eleger as melhores propostas.

6 Exemplo de PDA: Ministério da Justiça- 2017/201837

6.1 Apresentação

Este documento contém o Plano de Dados Abertos do Ministério da Justiça e Segurança Pública – PDA/MJ, que, conforme conceito apresentado pelo Ministério do Desenvolvimento, Orçamento e Gestão – MP, visa:

“orientar as ações de implementação e promoção de abertura de dados, inclusive os geoespacializados, que deverão obedecer aos padrões mínimos de qualidade, de forma a facilitar o entendimento e a reutilização das informações. É ele quem organiza o planejamento referente à implantação e racionalização dos processos de publicação de dados abertos nas organizações pública”.

Nesse contexto, o presente documento detalha a Política de Dados Abertos e Espaciais no âmbito do Ministério, instituída pela Portaria nº 1.378, de 20 de agosto de 2014, e apresenta orientações estratégicas e operacionais para ações de implementação e promoção da abertura de dados, permitindo a maior transparência das informações e a reutilização dos dados públicos pela sociedade civil.

Sua elaboração vem ao encontro do disposto na Lei nº 12.527, de 18 de novembro de 2011, a chamada Lei de Acesso à Informação – LAI; na Instrução Normativa SLTI nº 4, de 12 de abril de 2012, que institui a Infraestrutura Nacional de Dados Abertos – INDA; bem como no Decreto nº 8.777, de 8 de maio de 2016, que instituiu a Política de Dados Abertos do Poder Executivo federal. Apresenta, ainda, vinculação com os compromissos assumidos pelo Brasil no âmbito da Parceria

³⁷ Disponível em <https://www.justica.gov.br/dados-abertos/plano-de-dados/plano-2.pdf>. Acesso em 22 jan. 2020.

para Governo Aberto (*Open Government Partnership – OGP*), dentre outros normativos que abordam o tema da transparência das informações públicas.

Diante deste cenário, o Plano será o instrumento de planejamento e coordenação das ações de disponibilização de dados pelo Ministério da Justiça e Segurança Pública – MJ, para o biênio 2017-2018.

A partir de diretrizes estratégicas, são definidas no presente documento: o cenário institucional vigente no período de elaboração do plano; os objetivos; o escopo; a metodologia e os critérios de priorização de bases de dados no âmbito do MJ; bases a serem abertas na primeira etapa e metodologia para definição das bases a serem abertas na segunda etapa; estratégia para a abertura; diretrizes e recomendações; governança, monitoramento e controle; bem como o plano de ação para abertura.

O PDA/MJ leva em consideração um amplo conjunto de normativos aplicáveis ao tema e diversas experiências que norteiam e promovem o aperfeiçoamento dos processos de publicação de dados no Ministério e a construção de novos serviços de informação.

Por fim, a institucionalização do Plano, sua governança e revisões deverão ser comunicadas a todo o Ministério e à sociedade civil por meio das páginas oficiais do Ministério da Justiça e Segurança Pública e dos demais meios de comunicação e informação institucionais.

6.2 Introdução

O Plano de Dados Abertos do Ministério da Justiça e Segurança Pública reflete o esforço institucional em atender às necessidades da sociedade no âmbito da transparência da informação em consonância ao que estabelecem as legislações e os normativos vigentes.

A divulgação de dados pela Administração Pública Federal é um dever, decorrente do princípio da publicidade, constante no art. 37 da Constituição Federal de 1988, bem como do direito do cidadão em receber informações públicas, previsto

no inciso XXXIII do art. 5º e no § 2º do art. 216 da Constituição.

Com a entrada em vigor da Lei nº 12.527, de 18 de novembro de 2011, denominada Lei de Acesso à Informação – LAI, e do Decreto nº 7.724, de 16 de maio de 2012, que a regulamentou, o Poder Público passou a ter a obrigação de não apenas publicar as informações na Internet, mas também de disponibilizá-las em formatos abertos e processáveis por máquina, ou seja, seguindo o preceito dos dados abertos.

Como extensão, foi instituída, em agosto de 2014, a Política de Dados Abertos e Espaciais do Ministério da Justiça e Cidadania, por intermédio da Portaria MJ nº 1.378, com o objetivo de disponibilizar dados de forma aberta e estruturada, fomentando o desenvolvimento de novas tecnologias voltadas à construção de ambiente de gestão pública participativa e democrática, garantindo ainda a oferta de dados primários, íntegros, autênticos e atualizados aos cidadãos.

Nesse contexto, foi aprovado também o 1º Plano de Dados Abertos da pasta, visando a abertura de dados, ou seja, a publicação, disseminação e manutenção de dados e informações públicas na Internet, organizados de tal maneira que permita sua reutilização em aplicativos digitais desenvolvidos pela sociedade.

Este segundo PDA/MJ, válido para o biênio 2017-2018, pretende aprimorar o catálogo de bases existente, estabelecer estratégia de abertura de dados, critérios para priorização de abertura e fluxo de governança. Vale destacar, ainda, que a elaboração do presente Plano foi reforçada pela publicação do Decreto nº 8.777, de 11 de maio de 2016, que instituiu a Política de Dados Abertos do Poder Executivo Federal, definiu o instituto do “Plano de Dados Abertos” e tornou sua publicação obrigatória.

Haja vista que um dos pilares da política de abertura de dados é, notavelmente, a participação social. O presente documento, assim como o PDA anterior, foi disponibilizado na plataforma governamental Participa.br, a fim de receber comentários e sugestões da sociedade civil quanto a seu conteúdo.

6.3 Cenário Institucional

Com os princípios da transparência pública e do controle social, a

Constituição Federal de 1988 estabeleceu um novo marco no processo de construção da gestão democrática participativa no Brasil. Ao mesmo tempo, o avanço na área da tecnologia da informação e, sobretudo, a crescente popularização da Internet permitiram o uso de novas tecnologias para a prestação de serviços públicos e implementação de instrumentos de controle e participação social na Administração Pública.

Por conseguinte, o então Ministério da Justiça instituiu, em 2004, o Programa de Transparência e Acesso a Informações, que assumiu importância estratégica na política de transparência do órgão e na promoção da cultura do acesso à informação e abertura de dados públicos, de maneira a assegurar os meios adequados e efetivos para sua divulgação. O Programa, hoje institucionalizado e denominado Serviço de Transparência e Acesso à Informação, integra atualmente a Ouvidoria-Geral do MJ e tem como atribuição coordenar e promover iniciativas relacionadas às ações de transparência do órgão, debruçando-se na implementação da LAI, com destaque para a promoção da política de dados abertos.

Em 2012, após a celebração de um Acordo de Cooperação Técnica, sem transferência de recursos, com o Comitê Gestor da Internet no Brasil – CGI.br, representado pelo Núcleo de Informação e Coordenação do Ponto BR – NIC.br, por meio do Escritório Brasileiro do W3C, foi publicada a primeira base de dados do Ministério no Portal Brasileiro de Dados Abertos do Governo Federal, o Cadastro Nacional de Reclamações Fundamentadas, composta pelas reclamações de consumidores de todo o Brasil finalizadas pelos Procons integrados ao Sistema Nacional de Informações de Defesa do Consumidor – Sindec dos anos de 2009, 2010 e 2011.

Entre os anos de 2012 e 2015, o Ministério e seus parceiros realizaram ainda alguns Concurso de Aplicativos e *Hackathons*, sempre com o intuito de fomentar a utilização de bases públicas de dados, de promover a integração entre governo e sociedade civil, além de facilitar a vida do cidadão e dos servidores públicos a partir

de uma melhor forma de interpretar os dados existentes.

Ao longo desse período, diversas bases de dados públicos foram disponibilizadas no Portal Brasileiro de Dados, do Governo Federal. Tais iniciativas demonstram o propósito do Ministério da Justiça e Segurança Pública em estimular a disseminação de dados e informações governamentais para o livre uso pela sociedade, no desenvolvimento de um estado mais eficiente, com oferta de serviços melhores, de forma colaborativa com a sociedade civil.

Em 11 de maio de 2016, entrou em vigor o Decreto nº 8.777, que instituiu a Política de Dados Abertos do Poder Executivo Federal. Dentre outras medidas, a norma determinou a publicação de Planos de Dados Abertos, bem como estabeleceu que os prazos e os procedimentos previstos para o processamento de pedidos de acesso à informação serão igualmente aplicáveis às solicitações de abertura de bases.

Por fim, deve-se destacar que nos anos de 2016 e 2017, especialmente, o cenário político brasileiro impactou nos trabalhos para o desenvolvimento deste PDA, em razão das sucessivas alterações no comando e na estrutura desta pasta.

6.3.1 Infraestrutura Nacional de Dados Espaciais – INDE

Em 2014, o MJ solicitou sua adesão à Infraestrutura Nacional de Dados Espaciais – INDE, celebrando o Termo de Adesão nos termos do disposto no art. 3º do Decreto 6.666, de 27 de novembro de 2008, declarando, assim, o compromisso em adotar as diretrizes da INDE, estabelecer ações e metas de acordo com o seu Plano de Ação, além de seguir os normativos da Comissão Nacional de Cartografia – CONCAR.

6.3.2 Plano de Dados Abertos 2014-2015

Partindo do aprendizado obtido com as ações já desenvolvidas e com as parcerias firmadas, e decidido a aprimorar a política de dados abertos, conferindo-lhe caráter mais institucional, o Ministério da Justiça e Segurança Pública desenvolveu, em 2014, seu Plano Institucional de Dados Abertos e Espaciais, instrumento que

norteou, nos últimos dois anos, as ações e a sustentabilidade das bases de dados já disponibilizadas em formato aberto.

O PDAE/MJ do biênio 2014-2015, o primeiro lançado pelo órgão, estabeleceu parâmetros iniciais a serem observados no processo de abertura e sustentabilidade das bases de dados da Pasta. Cabe ressaltar que o MJ foi um dos órgãos pioneiros na concepção de seu próprio documento institucional sobre o tema.

O Plano de Dados Abertos anterior estabeleceu diversas metas a serem alcançadas; nem todas as atividades foram cumpridas, muito por conta das mudanças imprevistas no contexto organizacional, bem como no cenário político brasileiro.

Dentre as realizadas nos últimos dois anos, podemos citar:

- Publicação de Portaria instituindo a política de Dados Abertos e Espaciais do MJ;
- Levantamento parcial das bases de dados do Ministério da Justiça;
- Definição de metodologia para publicação das bases;
- Definição, junto às unidades responsáveis pelos dados, da periodicidade adequada para cada base, de acordo com as possibilidades de cada órgão;
- Atualização constante de dados;
- Atualização constante da página de dados abertos vinculada ao Portal do MJ, inserindo as novas bases que foram catalogadas no Portal Brasileiro de Dados Abertos e acrescentando novas informações relevantes para o tema.

Sobre este último ponto, é importante destacar a implementação, neste órgão, do Portal de Dados do Ministério da Justiça e Segurança Pública – Dados.MJ, concebido em cima da plataforma de catálogo de dados CKAN, a mesma utilizada no Portal Brasileiro de Dados Abertos. Seu principal intento é permitir a construção de um inventário, bem como facilitar a abertura do maior número possível de bases. Ainda em fase inicial, o principal desafio do órgão para este ano e o próximo é difundir o uso do Dados.MJ entre todas as unidades e entidades vinculadas, disseminando, assim, a cultura de dados legíveis por máquina, livremente disponíveis e reutilizáveis.

6.3.3 Laboratório de Participação e Inovação – LabPi

Na mesma esteira, em 2015, instituiu-se o Laboratório de Participação e Inovação do Ministério da Justiça– LabPi/MJ, com o objetivo de fomentar ações e práticas no sentido de estimular a participação da sociedade de forma criativa junto ao órgão.

6.3.4 Planejamento Estratégico Institucional

O presente Plano de Dados Abertos encontra-se alinhado à missão institucional, que é trabalhar para a consolidação do Estado Democrático de Direito, assim como está em conformidade com o Mapa Estratégico 2015-2019 do Ministério da Justiça.

Além disso, encontra-se alinhado ainda, ao objetivo estratégico habilitador “ampliar ações de participação social e de comunicação com a sociedade” e ao objetivo fundamento “aprimorar e integrar a gestão e a governança institucional”.

6.3.5 Plano Estratégico de Tecnologia da Informação e Comunicação – PETIC

O Plano Estratégico de Tecnologia da Informação e Comunicação a ser elaborado para o ciclo 2017-2020, será alinhado com a Estratégia de Governança Digital – EGD em vigor, cujos indicadores possuem enfoque, além da abertura dos dados, na prestação de serviços públicos digitais e na participação da sociedade civil.

6.3.6 Plano de Trabalho da Estratégia de Governança Digital

O Plano de Dados Abertos é um dos escopos do Plano de Trabalho para Governança Digital, firmado entre o Ministério da Justiça e o Ministério do Planejamento, Orçamento e Gestão.

6.3.7 Open Government Partnership – OGP

A abertura de dados no âmbito desta pasta também se encontra relacionada à Parceria para Governo Aberto, que é uma “iniciativa internacional que pretende difundir e incentivar globalmente práticas governamentais relacionadas à transparência dos governos, ao acesso à informação pública e à participação social”.

No 2º Plano de Ação do Brasil, a abertura de dados foi um dos compromissos firmados pelo Ministério da Justiça e Segurança Pública. Já o 3º Plano, traz diversas iniciativas em Governo Aberto, dentre as quais podemos destacar ações voltadas à participação social, a disponibilização de estruturas de transparência ativa e o efetivo atendimento das demandas criadas por força da Lei de Acesso à Informação.

Ao todo, o 3º Plano de Ação Nacional elenca 16 compromissos, dos quais o Ministério da Justiça e Segurança Pública será responsável pela coordenação de 1 (um) compromisso e colaborará com a implementação de outros 6 (seis), sendo 1 (um) relacionado à Dados Abertos no Governo Federal e 1 (um) relacionado à Transparência sobre Recursos Públicos.

6.4 Objetivos

6.4.1 Objetivo Geral

Promover o planejamento e a coordenação das ações de disponibilização e sustentabilidade de dados abertos, no Ministério da Justiça e Segurança Pública, zelando pelos princípios da publicidade, transparência e eficiência, e visando o aumento da disseminação de dados e informações para a sociedade em formatos reutilizáveis, de maneira a incentivar a participação social e promover a melhoria da qualidade dos dados disponibilizados.

6.4.2 Objetivos Específicos

No âmbito do Ministério da Justiça e Segurança Pública:

- Aprimorar o catálogo com as bases de dados sob a guarda do MJ;

- Identificar prioridades, a partir de critérios objetivos;
- Selecionar, na medida do possível, com a participação da sociedade civil, os dados a serem disponibilizados em formatos abertos;
- Propiciar o conhecimento público sobre as bases de dados mantidas pelo Ministério da Justiça e Segurança Pública
- Aperfeiçoar mecanismos e procedimentos de disponibilização dos dados públicos;
- Estimular o cruzamento de dados e a produção de conhecimento em proveito da sociedade e do poder público;
- Fomentar o desenvolvimento de aplicações pela sociedade civil para reutilização dos dados públicos disponibilizados;
- Melhorar a gestão da informação e de dados;
- Melhorar a qualidade dos dados disponibilizados;
- Estimular a interoperabilidade de dados e sistemas governamentais pela publicação de informações em formato processável por máquina, conforme padrões estabelecidos;
- Promover a sustentabilidade dos dados publicados;
- Incrementar os processos de transparência e de acesso a informações públicas;
- Buscar as melhores práticas para acesso e integração de dados.

6.5 Definição dos Dados a Serem Abertos

Diante da atual conjuntura de reformulação das diretrizes políticas e do processo de gestão pelo qual passa o Ministério da Justiça e Segurança Pública, adotou-se a estratégia de manter o Plano de Dados Abertos composto por duas etapas claramente subdivididas. A primeira etapa englobará as unidades do Grupo I, enquanto a segunda etapa, as unidades dos Grupos II e III

6.6 Escopo do PDA

O escopo do Plano de Dados Abertos engloba os procedimentos e metodologias descritos para priorização, seleção e abertura de bancos de dados em formatos não proprietários e com uma licença aberta.

Além disto, o Plano envolverá todas as unidades organizacionais do Ministério da Justiça e Segurança Pública, com exceção das unidades vinculadas,

Cade e Funai.

Tendo em vista a quantidade de Secretarias e Departamentos que compõem o MJ, e com o intuito de facilitar o entendimento do presente trabalho, as unidades serão alocadas em 3 (três) grupos, conforme tabela abaixo:

GRUPO I

Secretaria Nacional de Segurança Pública – Senasp

Secretaria Nacional de Justiça e Cidadania – SNJC

Secretaria Nacional de Políticas sobre Drogas – Senad

Secretaria Nacional do Consumidor – Senacon

Departamento Penitenciário Nacional – Depen

GRUPO II

Gabinete do Ministro – GM

Secretaria Executiva – SE

Consultoria Jurídica – CJ

Comissão de Anistia – CA

Secretaria de Assuntos Legislativos – SAL

GRUPO III

Secretaria Extraordinária de Segurança para Grandes Eventos – Sesge

Departamento de Polícia Federal – DPF

Polícia Rodoviária Federal – PRF

Arquivo Nacional – AN

6.7 Primeira Etapa PDA/MJ – Grupo I

A primeira etapa de elaboração do Plano de Dados Abertos abrangerá as unidades do Grupo I, que abarcam uma vasta gama de temas, como: segurança pública, enfrentamento às drogas, defesa do consumidor, política penal, cidadania, migração, dentre outros.

6.7.1 Catálogo e Priorização

A catalogação, priorização e seleção das bases de dados a serem abertas na primeira etapa de elaboração do PDA observará os seguintes procedimentos:

1. Catalogação: contato com os gestores das unidades para apresentação de catálogo prévio de bases de dados, obtido em 2014, bem como de resultado de um levantamento realizado pela CGTI contendo possíveis bases de dados, com o objetivo de que a relação seja confrontada, complementada e atualizada, resultando, ao final, em um catálogo confiável de bases de dados em poder da unidade;
2. Avaliação qualitativa das bases de dados para priorização: as unidades gestoras das bases de dados identificadas no passo “1” deverão aplicar os critérios de priorização da área de negócio (item 3.2.2) e devolver o catálogo atualizado e avaliado para o Serviço de Transparência e Acesso à Informação da Ouvidoria-Geral do Ministério da Justiça e Segurança Pública;
3. Saneamento do catálogo: para o adequado saneamento do catálogo e identificação das bases de dados potencialmente aptas à abertura, serão aplicados os seguintes filtros:
 - a) Planilha – as bases de dados que estejam em formato de planilha serão retiradas do catálogo em razão do potencial de vulnerabilidade de suas informações. Este tipo de base de dados deverá ter tratamento diferenciado, razão pela qual não fará parte do escopo deste PDA, mas poderá ser contemplado nos próximos ou trabalhados paralelamente ao longo da execução do Plano;
 - b) Base sigilosa – as bases avaliadas pelas unidades gestoras como “sigilosas” serão excluídas do catálogo e poderão ser avaliadas com maior cautela posteriormente;
 - c) Bases já abertas – considerando que o levantamento incluía também bases que já estão disponibilizadas em formato aberto, estas também serão retiradas do catálogo.
4. Ranking: após o saneamento das planilhas, os pontos dos critérios de priorização serão somados e classificados em ordem decrescente, criando, desta forma, um ranking para cada Secretaria que compõe o Grupo I;
5. Seleção da base de dados para abertura: com o resultado do ranking, a primeira colocada em cada planilha será submetida à unidade gestora como base a ser aberta durante a vigência do PDA 2017-2018.

Na fase de implementação do Plano de Dados Abertos, serão aplicados pela Coordenação-Geral de Infraestrutura e Governança de Tecnologia da Informação, com o apoio do Serviço de Transparência e Acesso à Informação e da unidade gestora da base, os critérios de avaliação de viabilidade técnica (item 4.3). Nesse sentido, caso a base selecionada exija um esforço desproporcional da área de tecnologia da informação, de forma que possa prejudicar a abertura das outras bases previstas no PDA, e sempre que tecnicamente justificado, será selecionada a base subsequente.

O mesmo ocorrerá se for verificado no momento da avaliação de viabilidade técnica que a base de dados não possui maturidade suficiente para sua abertura, de modo a comprometer a qualidade e a integridade do dado.

6.7.2 Crítérios de Priorização da Área de Negócio

Obrigação Legal para Abertura

A. O dado ou informação tem previsão para abertura:

- não possui previsão (1);
- deve ser disponibilizado em razão da determinação de algum órgão de controle (2);
- deve ser divulgado em razão de algum acordo/compromisso internacional ou legislação (3)

Nível de Transparência

A. Situação atual do dado:

- disponível em formato aberto (1);
- não disponível (2);
- disponível em formato não aberto (3).

B. Nível de transparência da base:

- dados sigilosos/pessoais (1);
 - dados mistos (2);
 - dados ostensivos - sem nenhum tipo de restrição de acesso (3).
-

Nível Estratégico

- A. Potencial de utilização dos dados para decisão e monitoramento de políticas públicas:
- nulo ou mínimo (1);
 - médio (2);
 - alto (3).
- B. Potencial de utilização do dado como indicador ou componente de indicador de planejamento estratégico:
- baixo (1);
 - médio (2);
 - alto (3).
- C. Potencial de utilização por outros órgãos públicos:
- baixo (1);
 - médio (2);
 - alto (3).
-

- Governo (1);
 - Sociedade (2);
 - Governo e Sociedade (3).
- D. Frequência de demanda dos dados pela população (ouvidoria/SIC/comunicação social ou diretamente na área):
- baixa (1);
 - média (2);
 - alta (3)
-

6.8 Segunda Etapa PDA/MJ – Grupos II e III

A segunda etapa de elaboração do Plano de Dados Abertos abrangerá as unidades dos Grupos II e III, que tratam de temas como: grandes eventos, rodovias federais, produtos químicos, anistia política, dentre outros.

6.8.1 Catálogo e Priorização

A catalogação, priorização e seleção das bases de dados a serem abertas na segunda etapa de elaboração do PDA observará os seguintes procedimentos;

1. Catalogação: contato com os gestores das unidades para apresentação de catálogo prévio de bases de dados, obtido em 2014, bem como catálogo

- resultado de um levantamento realizado pela CGTI contendo possíveis bases de dados, com o objetivo de que a relação seja confrontada, complementada e atualizada, resultando, ao final, em um catálogo confiável de bases de dados em poder da unidade;
2. Avaliação qualitativa das bases de dados para priorização: as unidades gestoras das bases de dados identificadas no passo 1 deverão aplicar os critérios de priorização da área de negócio (item 3.2.2) e devolver o catálogo atualizado e avaliado ao Serviço de Transparência e Acesso à Informação da Ouvidoria-Geral do Ministério da Justiça e Segurança Pública;
 3. Saneamento do catálogo: para o adequado saneamento do catálogo e identificação das bases de dados potencialmente aptas a abertura, serão aplicados os seguintes filtros:
 - a) Planilha – as bases de dados que estejam em formato de planilha serão retiradas do catálogo em razão do potencial de vulnerabilidade de suas informações. Este tipo de base de dado deverá ter outro tipo de tratamento, razão pela qual não fará parte do escopo deste PDA, mas poderá ser contemplada nos próximos ou trabalhadas paralelamente ao Plano;
 - b) Base sigilosa – as bases avaliadas pelas unidades gestoras como “sigilosas” serão excluídas do catálogo e poderão ser avaliadas com maior cautela posteriormente;
 - c) Bases já abertas – considerando que o levantamento incluía também bases que já estão disponibilizadas em formato aberto, estas também serão retiradas do catálogo.
 4. Ranking: após o saneamento das planilhas, os pontos dos critérios de priorização serão somados e classificados em ordem decrescente, criando, desta forma, um ranking unificado para o Grupo II e um ranking para cada Secretaria que compõe o Grupo III;
 5. Consulta pública: as bases consideradas aptas em cada um dos rankings serão submetidas à priorização de abertura pela sociedade por intermédio de uma consulta pública realizada virtualmente;
 6. Seleção da base de dados para abertura: com o resultado da consulta pública, a primeira colocada em cada planilha das unidades do Grupo III

e as duas primeiras colocadas no ranking do Grupo II serão submetidas à unidade gestora como bases a serem abertas durante a vigência do PDA 2017-2018.

Na fase de implementação do Plano de Dados Abertos, serão aplicados pela Coordenação-Geral de Infraestrutura e Governança de Tecnologia da Informação, com o apoio do Serviço de Transparência e Acesso à Informação e da unidade gestora da base, os critérios de avaliação de viabilidade técnica (item 4.3). Nesse sentido, caso a base selecionada exija um esforço desproporcional da área de tecnologia da informação, de forma que possa prejudicar a abertura das outras bases previstas no PDA, e sempre que tecnicamente justificado, será selecionada a base subsequente.

O mesmo ocorrerá se for verificado no momento da avaliação de viabilidade técnica que a base de dados não possui maturidade suficiente para sua abertura, de forma a comprometer a qualidade do dado.

Após a conclusão do levantamento previsto para a segunda etapa, as bases de dados consideradas potencialmente aptas para abertura serão incluídas no primeiro catálogo de base de dados, constituindo-se, assim, o catálogo completo de base de dados do MJSP.

A segunda etapa do Plano de Dados Abertos deverá ser finalizada até fevereiro de 2018 e as bases selecionadas nesta etapa serão acrescentadas ao Plano de Ação inicial, acarretando, portanto, na revisão do presente PDA. Ressalta-se que as bases selecionadas serão igualmente objeto de publicação em formato não proprietário e com licença aberta até o final de 2018.

6.9 Estratégias para Abertura

A estratégia para abertura dos dados selecionados observará os seguintes procedimentos:

- a) Definição de 1 (um) interlocutor, lotado no Gabinete do Secretário ou Diretor-Geral, e 1 (um) suplente, de cada Secretaria ou Direção-Geral,

- responsáveis pelo monitoramento da execução do Plano de Dados Abertos em sua unidade e pela validação da base de dados antes de sua publicação;
- b) Consolidação de matriz de responsabilidade;
 - c) Observância ao modelo de governança estabelecido;
 - d) Utilização de metodologia de abertura de dados a ser seguida pelas áreas responsáveis (padrões INDA e INDE), sempre que possível;
 - e) Capacitação dos responsáveis nas áreas de negócio das bases selecionadas para abertura;
 - f) Definição de arquitetura de abertura para cada base de dados a ser aberta;
 - g) Publicação dos dados selecionados, observando-se o uso de URL fixa, no Portal dados.mj e no Portal Brasileiro de Dados Abertos.

Pela metodologia apresentada, somando-se a primeira e a segunda etapa, serão abertas, durante o período de vigência do presente plano, até 9 (nove) bases.

O processo de abertura dos conjuntos de dados observará as mesmas diretrizes estabelecidas no Decreto nº 8.777, de 2016; na Portaria MJ nº 1.378, de 2014, bem como no Plano de Ação da INDA.

6.10 Recomendações

a. Utilizar o CKAN como repositório, ferramenta de levantamento de bases e sincronização automática dos novos recursos, sempre que possível, mantendo os dados publicados atualizados e sincronizados com a origem, com a menor periodicidade e maior granularidade viáveis;

b. Publicar dados e seus metadados, conforme estabelecido no Plano de Ação da INDA;

c. Buscar a facilidade do intercâmbio de dados entre órgãos e entidades da administração pública federal e as diferentes esferas da federação, bem como o compartilhamento de recursos de tecnologia da informação, de maneira a evitar a duplicidade de ações e o desperdício de recursos na disseminação de dados e informações;

d. Publicar os dados do Ministério seguindo os padrões definidos pela e-PING, pela Infraestrutura Nacional de Dados Abertos – INDA e pela Infraestrutura

Nacional de Dados Espaciais – INDE e Governo Eletrônico, na medida da capacidade técnica do MJ;

e. Buscar a completa integração entre o Portal de Dados do MJ e o Portal Brasileiro de Dados Abertos, ponto central de acesso aos dados do governo federal, para permitir o intercâmbio de dados, mantendo URL fixa;

f. Catalogar os dados geoespacializados no Diretório Brasileiro de Dados Geoespaciais, na medida da capacidade técnica do MJ;

g. Unificar fluxos e procedimentos para a publicação de dados abertos e espaciais, atendendo de forma conjunta padrões da INDA e INDE, na medida do possível

6.11 Avaliação de viabilidade técnica

Na fase de implementação, serão avaliadas ainda os seguintes critérios de viabilidade técnica:

1. Capacidade de Operacionalização para Abertura
 - a. Tamanho estimado em pontos de função:
 - i. mais de 100 [1];
 - ii. entre 20 e 99 [2];
 - iii. menos de 20 [3].
 - b. Necessidade de Integração:
 - i. acima de 2 sistemas [1];
 - ii. até 2 sistemas [2];
 - iii. não integra com outros sistemas [3].
 - c. Tecnologia envolvida:
 - i. não há domínio da tecnologia [1];
 - ii. não é a tecnologia padrão, mas há domínio técnico [2]
 - iii. tecnologia padrão MJC [3].
 - d. Complexidade para extrair os dados:
 - i. combinar várias tabelas de dados e envolver cálculos [1];
 - ii. combinar várias tabelas com extração parcial dos dados [2];
 - iii. combinar várias tabelas sem tratamento parcial ou acesso direto [3].

2. Qualidade do Dado

- a. Existência de dicionário de dados do Modelo Entidade Relacionamento:
 - i. não possui dicionário de dados [1];
 - ii. possui dicionário de dados parcial [2];
 - iii. possui dicionário de dados completo [3].
- b. Existência de documentação do sistema:
 - i. não possui documentação do sistema [1];
 - ii. possui documentação do sistema, porém desatualizada [2];
 - iii. possui documentação do sistema atualizada [3].
- c. Fonte institucional dos dados a serem abertos:
 - i. coleta, alimentação e extração por uma ou mais de uma fonte(s) manual(is) [1];
 - ii. coleta, alimentação e extração por mais de uma fonte sistematizada [2];
 - iii. coleta, alimentação e extração de dados única e corporativa [3].

Caso a base selecionada exija um esforço desproporcional devido à obtenção do valor mínimo de pontuação na maioria dos critérios de viabilidade técnica, prejudicando a abertura das outras bases previstas no PDA, será selecionada a base com colocação subsequente.

6.12 Metadados

A publicação dos dados com seus respectivos metadados será considerada uma prioridade neste 2º Plano de Dados Abertos, tendo em vista que são “informações úteis para identificar, localizar, compreender e gerenciar dados”, sendo, portanto, imprescindível seu correto registro.

“Quando documentamos os metadados e os disponibilizamos, estamos enriquecendo a semântica do dado produzido, agregando seu significado real, e dando suporte à atividade de Administração de Dados executada pelo produtor desse dado”.

Neste contexto, cada conjunto de dados deverá conter, no mínimo:

- a) Nome ou título do conjunto de dados;

- b) Descrição sucinta;
- c) Palavras-chaves (etiquetas);
- d) Assuntos relacionados no VCGE – Vocabulário Controlado do Governo Eletrônico;
- e) Nome e e-mail do setor responsável pelos dados;
- f) Periodicidade de atualização no portal de dados;
- g) Escopo temporal;
- h) Escopo geopolítico (por cidade, por estado, por região); e
- i) Data da última atualização.

Ademais, toda base de dados cadastrada no Portal de Dados Abertos do MJ e no Portal Brasileiro de Dados Abertos do Governo Federal deverá conter ainda um arquivo composto pelo dicionário de dados da base, descrevendo cada campo que a integra.

6.13 Governança

A estrutura de governança do Plano de Dados Abertos do MJ será composta pela autoridade de monitoramento da Lei de Acesso à Informação; pela Ouvidoria-Geral, por intermédio do Serviço de Transparência e Acesso à Informação; pela Coordenação-Geral de Infraestrutura e Governança de Tecnologia da Informação; pelo Comitê de Governança Administrativo; e, precipuamente, por todas as unidades do Ministério da Justiça e Segurança Pública.

O quadro a seguir representa a estrutura de governança do PDA/MJ:

Autoridade de Monitoramento da LAI

Supervisionar a execução do PDA, nos termos do § 4º do art. 40 do Decreto nº 8.777, de 2016

Comitê de Governança Administrativa

Instância composta pelos chefes de gabinete das Secretarias, responsável pelas deliberações relacionadas ao PDA, como revisão e atualização.

Ouvidoria-Geral

Monitorar a execução do Plano de Dados Abertos; dar suporte às áreas para garantir a implementação do plano; e fomentar a cultura de abertura de dados na instituição

CGTI

Fomentar a política de abertura de dados; Integrar as ações de abertura de dados às ações de tecnologia de Informação, especialmente, as ações referentes à Governança Digital; Dar suporte técnico e tecnológico para a implementação do PDA;

Unidades MJC

Tutela dos dados e responsabilidade pelas informações e qualidade das bases; publicação periódica no portal Dados.MJ das bases já abertas e das bases previstas nestes documentos; manter catálogo de base de dados atualizado; informar sempre sobre a disponibilidade de novas bases de dados com potencial de abertura.

Observação: sempre que a base de dados estiver submetida à uma instância de gestão ou governança não prevista neste PDA, sua abertura deverá ser submetida, para conhecimento ou autorização, conforme o caso, a esta instância.

6.14 Monitoramento e Controle

Tendo em vista a necessidade de articulação com os vários atores responsáveis, no âmbito do órgão, a fim de impulsionar e garantir o cumprimento da Política de Dados Abertos, estão previstos encontros com os envolvidos, ao longo do período durante o qual estiver em vigor o presente documento, nos seguintes moldes:

- Reunião inicial da Ouvidoria-Geral e da Coordenação de Tecnologia da Informação com cada uma das áreas de negócio, conforme a metodologia estabelecida no item 6.3.2, com o objetivo de apresentar a Política de Dados Abertos, a necessidade de elaboração e execução de um Plano de Dados Abertos, a necessidade de construir um catálogo de bases de dados da instituição, bem como propor a aplicação de critérios para avaliação e priorização das bases de dados;

- Reuniões da Ouvidoria-Geral junto à área de negócio, com a presença da CGTI e periodicidade sujeita à conveniência e necessidade específica, visando estabelecer tratativas necessárias à abertura de cada base;
- Reuniões mensais da Ouvidoria-Geral com a CGTI, com a finalidade de acompanhamento dos processos pertinentes à implementação do Plano de Dados Abertos;
- Participação, no mínimo, semestral, da Ouvidoria-Geral e da CGTI em encontros do Comitê de Governança Administrativa, com o objetivo de alinhar as informações sobre a implementação deste Plano, além de deliberar sobre atualizações no PDA, sempre que necessário.

6.15 Plano de Ação

6.15.1 Primeira Etapa

Secretaria Gestora	Nome da Base	Descrição da Base	Prazo para Abertura	Unidade Responsável pela Base
Secretaria Nacional do Consumidor	Recall 3.0	Nova base do Sistema de Recall, como resultado da atualização no aplicativo. O objetivo maior deste sistema é garantir a celeridade e a eficiência do intercâmbio de informações entre Senacon, cidadãos, entidades parceiras e mercado. O principal escopo do Recall é eliminar riscos à saúde e segurança dos consumidores, e a garantia do direito à informação aos consumidores.	Jun/2018	Coordenação de Consumo, Seguro e Saúde
Secretaria Nacional de Justiça e Cidadania	Sistema Gerencial da Classificação Indicativa - Classind	Sistema setorial denominado CLASSIND, no qual são cadastrados requerimentos de classificação de jogos eletrônicos, RPG e obras audiovisuais.	Dez/2017	Coordenação de Classificação Indicativa / Departamento de Justiça, Qualificação, Títulos e Classificação
Secretaria Nacional de Segurança Pública	Sinesp - JC Estatística	O Sistema Nacional de Estatísticas de Segurança Pública e Justiça Criminal (SINESPJC) foi desenvolvido com o objetivo de reunir as informações de segurança pública e justiça criminal produzidos pelos Estados da Federação.	Dez/2017	Coordenação-Geral de Pesquisa e Análise da Informação

7 Plataforma de publicação – CKAN

A adoção do CKAN (*Comprehensive Knowledge Archive Network*) como ferramenta de publicação de dados abertos é hoje mais que uma orientação, é uma indicação, face à sua utilização por quase a totalidade dos portais do mundo. Pode-se citar, por exemplo, os portais:



- data.gov - Portal de dados abertos do governo norte-americano.
- data.gov.uk - Portal de dados abertos do governo britânico.
- dados.gov.br - Portal de dados abertos do governo brasileiro.
- dados.pbh.gov.br – Portal de dados abertos da prefeitura de Belo Horizonte.
- dados.al.gov.br - Portal de dados abertos do Governo de Alagoas
- web.transparencia.pe.gov.br/dados-abertos - Portal de dados abertos do governo de Pernambuco.
- dados.recife.pe.gov.br - Portal de dados abertos da prefeitura de Recife.
- dados.ufrn.br - Portal de dados abertos da UFRN (Universidade Federal do Rio Grande do Norte).
- opendata.smart.ufs.br - Portal de dados abertos da UFS desenvolvido academicamente.

Outras instâncias do CKAN podem ser consultadas no seu site³⁸.

O CKAN é em si mesmo um portal para a construção de portais de dados abertos e apresenta as seguintes características:

- É uma plataforma *open source*.
- Desenvolvido em Python.
- É suportado por uma comunidade de desenvolvedores que busca corrigir falhas e inserir novas funcionalidades periodicamente.

³⁸ CKAN Instances - <<https://ckan.org/about/instances/>>

- É utilizada para publicação e gerenciamento de coleções de dados.
- Tem como funções primárias os processos de publicação, de compartilhamento, de descoberta e de uso dos dados.
- A publicação e o gerenciamento dos dados são realizados por meio de uma interface *Web*, pela utilização de uma API ou por planilhas customizadas.
- A busca e a descoberta de informações utilizam um sistema de buscas baseado nos buscadores tradicionais. A utilização de metadados na descrição dos dados maximiza a recuperação dos registros
- Seus principais utilizadores são instituições privadas ou governamentais que disponibilizam suas informações em formato aberto.
- Disponibiliza ferramentas de trabalho cooperativo que permitem que os usuários interajam e colaborem em relação aos dados.
- A funcionalidade de visualização *on-line* dos *datasets* permite verificar as características dos dados publicados e validar seu conteúdo.
- Permite download e upload de conjuntos de dados em vários formatos (JSON, CSV, XLS, XLSX, ODS, PDF, dentre outros).
- Disponibiliza customização através de extensões para adequar o CKAN às particularidades de cada instituição.
- Disponibiliza um histórico de ações em que todas as ações realizadas na plataforma são armazenadas para posterior consulta.
- Disponibiliza APIs em que um servidor de requisições realiza busca e acesso às informações nos *datasets* publicados.

O desenvolvimento do CKAN foi uma iniciativa da *Open Knowledge Foundation*³⁹, iniciado em março de 2006, sob a licença *open source Affero GNU GPL v3.0*. O CKAN tem como foco a publicação de dados de uma instituição na Web utilizando uma estrutura de banco de dados com as seguintes informações:

- Descrição dos dados/conjunto de dados armazenados em seu banco de dados.
- *Tags* de classificação da informação.

39 Organização sem fins lucrativos que promove conhecimento livre fundada em maio de 2004, em

- Órgãos e áreas temáticas da instituição.
- Formato em que os dados estão disponíveis.
- Informações pessoais e de contato com o responsável pelo dado.
- Informação sobre a disponibilidade e a licença de uso.
- Assuntos com que se relaciona.

Cada *dataset* contém informações extraídas de uma fonte de dados do sistema de informação origem. A título de exemplo podemos citar *datasets* contendo informações sobre compra de equipamentos, informações sobre folha de pagamento da instituição ou informações sobre o orçamento bianual da instituição.

A estrutura do *dataset* é baseada em dois tipos de arquivos:

- Os metadados que descrevem e dão significado a cada campo do arquivo de dados. O padrão adotado para a representação dos objetos digitais é composto por: título, identificador único, descrição, histórico de revisão, visualização de dados, campos extras, licença, *tags* e grupos.
- O conjunto de dados, nos formatos recomendados, tais como o CSV.

Para desenvolver o Portal Dados Abertos da Prefeitura de Belo Horizonte, o CKAN foi escolhido, além dos motivos já listados, por ser recomendado pela plataforma de cidades inteligentes FIWARE, que é uma infraestrutura em nuvem de plataforma aberta para IoT, financiada pela Comissão Europeia e pela União Europeia, baseada no OpenStack⁴⁰, plataforma *open source* para computação em nuvem.

Para ter acesso a informações sobre a instalação do CKAN você pode consultar o Manual de Instalação do CKAN publicado *on-line* pelo IBICT⁴¹ e para informações sobre o uso da plataforma consulte o Guia do Usuário CKAN também publicado pelo IBICT⁴².

⁴⁰ <https://www.openstack.org/>

⁴¹ Disponível em http://wiki.ibict.br/index.php/Manual_de_instala%C3%A7%C3%A3o. Acessado em 11/12/2019.

⁴² Disponível em

8 Dados abertos conectados

“A mágica dos dados abertos é que a mesma abertura que permite a transparência também permite a inovação.”

Ademir Piccoli.

Afirmar que houve um enorme crescimento no volume de dados produzidos nos últimos anos no mundo da internet, nas empresas, nas instituições governamentais e até mesmo em uma residência é algo muito fácil de ser constatado.

Dentre as estatísticas disponíveis sobre este crescimento encontra-se estimativas de que mais de 70% dos dados presentes na Web foram criados em anos recentes. E mais, o crescimento registrado não ocorreu apenas no volume, mas também no nível de detalhamento.

Segundo o DATA REVOLUTION GROUP (2014), este cenário foi potencializado pelas novas tecnologias advindas de diversas iniciativas tecnológicas. Dentre as mais relevantes podemos citar:

- IoT (BARNAGHI, 2012),
- Big Data (HITZLER; JANOWICZ, 2013),
- Web Semântica (BERNERS-LEE et al, 2001),
- Governo Aberto (MEIJER; CURTIN; HILLEBRANDT, 2012).

Na esteira do aumento no nível de detalhamento dos dados surgiram a Web semântica e conceitos tais como: dados conectados (*Linked Data*) e dados abertos conectados (*Linked Open Data*).

Considerando-se que a produção dos dados apresenta uma tendência a crescer astronomicamente e que a descentralização desta produção também se ampliará em igual escala, um desafio já possível de ser antecipado é a conexão de todo este volume

de dados. Sem esta conexão a utilidade de todo este potencial fica seriamente comprometida. Seja para a tomada de decisão da própria administração pública, seja para a participação da sociedade, seja para pesquisadores ou para a mídia especializada.

Assim, mesmo que os dados estejam disponíveis em formato aberto pelas instituições governamentais permanecem os desafios para que estas instituições disponibilizem meios de acesso a bases de dados de origem distribuída e que estes meios consigam superar as limitações humanas em conectar, processar, classificar e utilizar uma oferta tão grandes de dados e informações.

Alcantara *et al.* (2015), alguns dos autores que tratam este tema, afirmam que a publicação de dados em formato aberto e estruturado não se bastam para permitir sua utilização pela sociedade. Isto implica em que estes dados sejam publicados de acordo com padrões e vocabulários controlados que permitam a interoperabilidade dos dados na *Web* através de serviços automatizados.

Um dos conceitos mais promissores para solucionar este problema é o dos dados conectados. Segundo a Wikipedia, o conceito de *Linked Data*, dados ligados entre si, trata de um conjunto de práticas introduzidas por Tim Berners-Lee em suas notas sobre a arquitetura web "*Linked Data*", com função de publicar e estruturar dados na Web. Essas práticas, ao ter um crescimento expressivo em sua utilização, deram origem à criação do que ficou conhecido como web de dados, *Web of Data*. No contexto de Web Semântica, a função não é somente conectar os dados, mas também fazer com que a pessoa e a máquina possam explorar a web de dados.

Dados conectados para serem assim considerados, segundo Berners-Lee (2006), precisam atender a quatro princípios:

- Devem ser usados Identificadores Universais de Recursos (*Universal Resource Identifier* - URIs) em substituição aos nomes para as coisas a serem publicadas;
- Devem ser usadas URIs HTTP para que os usuários possam localizar estes nomes;

- A URI acessada deve prover informação útil usando padrões como o RDF (*Resource Description Framework*) ou o SPARQL;
- As URIs devem incluir hiperlinks para outras URIs, para que os usuários possam descobrir novos dados relativos ao assunto.

Dados Abertos Conectados seguem os mesmos princípios aplicáveis aos dados conectados (BANDEIRA et al., 2014), incorporando obrigatoriamente requisitos dos Dados Abertos como as permissões e formatos para uso e reuso livre, podendo ser redistribuídos por qualquer pessoa - sujeitos, no máximo, à exigência de atribuição da fonte e compartilhamento.

Assim, dados abertos conectados podem ser definidos como “um conjunto de práticas para publicação de dados abertos que incluem hiperlinks para outros dados abertos, mediante o uso de URIs que garantem que a partir de um dado, possam ser acessados outros dados relacionados”.

A iniciativa ISA, *Interoperability Solutions for European Public Administrations*, propõe uma comparação entre as características dos Dados Abertos Conectados com outros formatos de dados estruturados (ISA, 2014) que merece destaque. A comparação está listada na Tabela 8.

A evolução dos dados abertos de convencionais para conectados foi estabelecida pela escala de maturidade 5-Stars Linked Open Data (BERNERS-LEE, 2006) a partir da observação da pluralidade de conceitos no universo dos DGAs. Tim Berners-Lee propôs um cenário de cinco estrelas para permitir a classificação das publicações de dados abertos.

A base para esta classificação é o conceito de dados abertos conectados ou *Linked Open Data*. A proposta implica em que, a cada nova estrela conquistada, os dados sejam progressivamente mais fáceis de serem utilizados por todos.

Dimensão	Situação Atual (Dados não-conectados)	Situação desejada (Dados Conectados)
Compartilhamento de dados:	Dados são compartilhados utilizando o XML (eXtensible Markup Language)	Dados são compartilhados utilizando o RDF (Resource Description Framework)
Validação de dados:	Utilização do <i>XML Schema (XSD)</i> como meio para validação dos dados	Regras (ex: SWRL) e SPARQL utilizados como meio para validar dados e fazer inferências
Significado de dados:	Estrutura centralizada para prover significado aos dados, mediante a reutilização de arquivos XSD	Estruturas descentralizadas para associação de significado aos dados como vocabulários e outros dados de referência
Provimento de dados:	Serviços <i>Web</i> especializados baseados em SOAP permitem o acesso aos dados	Serviços RESTful leves de dados conectados permitem o acesso aos dados
Integração de dados:	Integração de Sistemas	Conexões de dados baseadas em elementos semânticos
Relacionamento com outros conceitos:	Dados e esquemas são considerados completos (mundo fechado)	Dados e esquemas são considerados incompletos (mundo aberto)
Endereçamento de dados:	Cada recurso (ou entidade) de dados possui um único identificador no nível dos sistemas de informação	Cada recurso (ou entidade) de dados possui identificadores comuns e conectados, por diferentes sistemas de informação, no nível da <i>Web</i>

Tabela 8 - Comparativo entre de dados conectados e outros formatos de dados estruturados.

De acordo com o *Open Knowledge Brasil* (2013), a definição das estrelas, ilustrada na Figura 13, é a seguinte:

- Uma estrela: os dados estão disponíveis na web, em qualquer formato, sob uma licença aberta. Um exemplo seria um relatório em formato pdf disponível em um site.
- Duas estrelas: os dados estão disponíveis na web **como dados estruturados legíveis por máquina**, sob uma licença aberta. Um exemplo seria uma planilha Excel disponível para download em um site.

- Três estrelas: os dados estão disponíveis na web como dados estruturados, **em formato não proprietário**, legíveis por máquina, sob uma licença aberta. Um exemplo seria uma planilha “.csv” disponível para download em um site.
- Quatro estrelas: os dados estão disponíveis na web como dados estruturados, em formato não proprietário, legíveis por máquina, sob uma licença aberta, **utilizando uma URL com arquitetura que possibilita identificar os dados**. Um exemplo seria um Portal com os dados organizados e disponíveis para acesso e para download.
- Cinco estrelas: dados disponíveis na web como dados estruturados, em formato não proprietário, legíveis por máquina, sob uma licença aberta, utilizando uma URL com arquitetura que possibilita identificar os dados, **conectando seus dados com dados de outros portais para prover contexto**. Um exemplo seria um Portal com os dados organizados de previsão do tempo com links de outros sites de meteorologia.

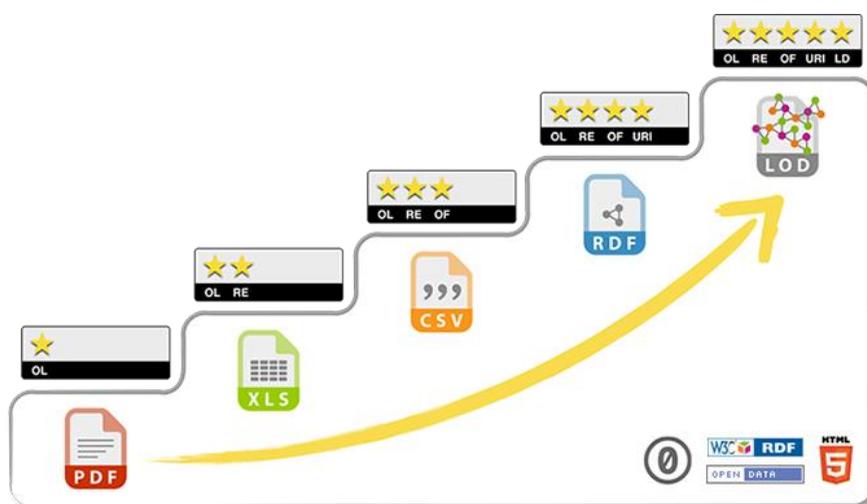


Figura 13 - Cenário de cinco estrelas de Timothy John Berners-Lee.

Os pesquisadores Richard Cyganiak, Insight Centre for Data Analytics at NUI Galway⁴³, e Anja Jentsch, Hasso Plattner Institut⁴⁴, criaram, então, o projeto *LOD*

⁴³ Disponível em <http://www.insight-centre.org>.

⁴⁴ Disponível em <http://www.hpi.de>.

Cloud, The *Linking Open Data Cloud Diagram*, para acompanhar a evolução dos dados abertos conectados.

O *LOD Cloud* cataloga e disponibiliza uma imagem que mostra os *datasets* que foram publicados como dados conectados, por contribuintes da comunidade de dados abertos conectados, *Linking Open Data*, e outros indivíduos e organizações. A “nuvem” de dados é baseada em metadados coletados mediante a curadoria de contribuintes para o Hub de Dados organizados na “*LOD Cloud*”.

A imagem disponibilizada pelo projeto, baseada em um banco de grafos, é interativa e cada conjunto de dados é representado por um círculo contendo um hiperlink para sua página inicial. Em março de 2009, o *LOD Cloud*, conforme a Figura 14, era composta por 1239 *datasets* e 16.147 *links*.

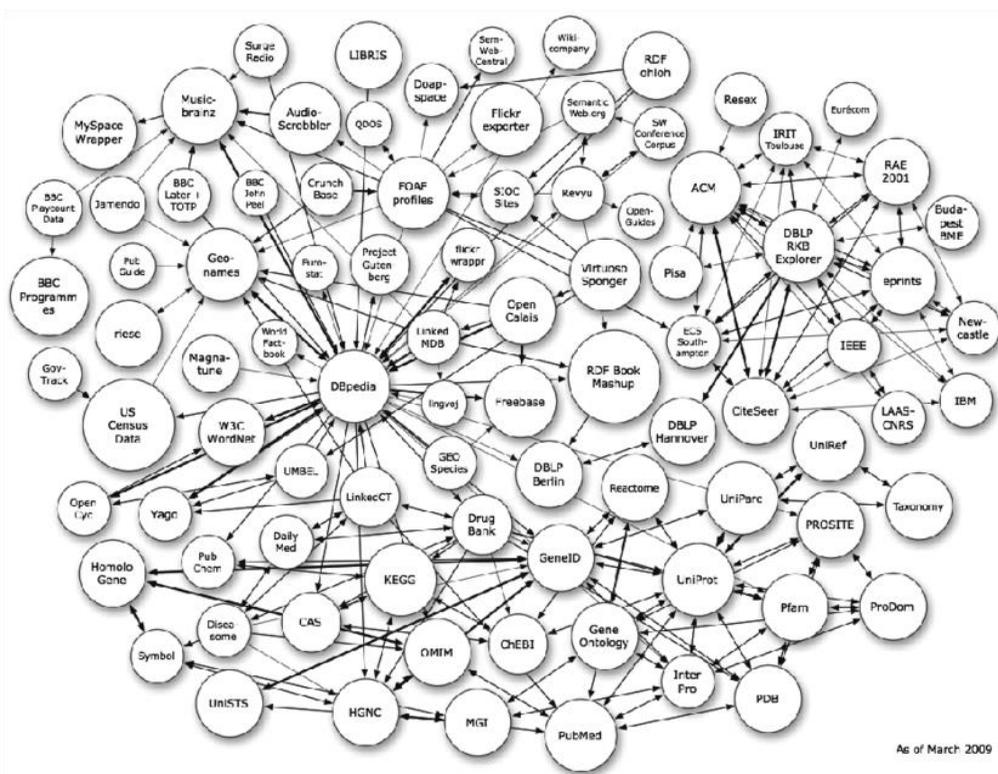


Figura 14 – *LOD Cloud* (imagem de 1970)⁴⁵.

⁴⁵ Disponível em <https://lod-cloud.net/>.

Entretanto, apesar da maior oferta de dados públicos, a falta de mecanismos para cruzamento dos dados dificulta seu uso por parte da sociedade. O atual modelo de publicação de dados estruturados na Web se mostra insuficiente, pois as bases de dados de um mesmo assunto em origens diferentes não apresentam mecanismos de compatibilização entre chaves de acesso (Hyland e Wood, 2011).

Este desafio se torna monumental no caso de grandes arquivos de dados, os dados abertos, governamentais ou não, desconsideram os requisitos semânticos que potencializam seu processamento automatizado.

Hyland e Wood (2011) exemplificam esta limitação citando um trecho do discurso do pesquisador Robert Schaefer, ao tratar de mudanças climáticas e situações críticas face ao clima no espaço na Conferência Internacional para Governo Aberto em 2010:

Dispor de dados abertos governamentais é algo ótimo, mas é muito difícil dar sentido aos dados sem o conhecimento do contexto de tais dados. É necessário o envolvimento de analistas e cientistas para gerar conhecimento a partir de tais dados. Só assim é possível repassar estes dados aos gestores políticos.

Um caso típico de dados publicados em o devido cuidado são os Portais de Dados Abertos. A publicação dos dados acontece sem atenção aos padrões, às melhores práticas e aos metadados. O resultado, via de regra, é um conjunto de *datasets* desconexos que dificulta a sua localização, sua reutilização e integração. Assim, é flagrante que dar sentido a estes dados é uma tarefa hercúlea.

Os Dados Abertos Conectados Governamentais (DACG) não são uma panaceia, mas são uma resposta possível a este problema no contexto de Governo Eletrônico. DACG consiste numa forma de indexar e conectar dados abertos governamentais que seguem os princípios estabelecidos por Berners-Lee (2006) e podem compor uma nuvem de DACG nos moldes da *LOD Cloud*.

A diversidade de conjuntos de dados, tecnologias de geração, processamento, persistência, recursos semânticos e padrões torna uma nuvem de dados abertos

conectados governamentais muito complexa, mas ela tem o potencial de suprir as necessidades de consumidores como os listados na Tabela 1,

Nesta nuvem as instituições governamentais fornecem seus *datasets* como serviços DACG eliminando a necessidade de os usuários executarem downloads de dados. O objetivo desta nuvem é processar conjuntos de dados completos, independentemente de seu tamanho. Como consequência direta é natural que geradores de dados e consumidores interajam, sem muitos protocolos regulatórios, gerando cooperação entre os participantes da nuvem.

Ávila (2015), em sua análise em relação às melhores práticas para publicação de dados abertos conectados propõe um conjunto de 70 recomendações, organizadas em relação às 10 BLPDs⁴⁶ (W3C, 2014) conforme listagem na Tabela 9.

BPLD	Recomendações
1. Preparar partes interessadas (stakeholders)	<ol style="list-style-type: none">1. Identificar os benefícios para a abertura de dados.2. Identificar as Partes Interessadas.3. Definir perfis profissionais a serem envolvidos.4. Definir grupos de usuários dos dados.5. Elaborar um plano de ações para publicação dos dados.6. Capacitar os envolvidos.
2. Selecionar conjuntos de dados	<ol style="list-style-type: none">7. Analisar a estrutura organizacional.8. Estabelecer diretrizes que orientem a priorização de dados a serem abertos.9. Realizar consultas aos usuários sobre a demanda de dados.10. Identificar os dados que serão abertos.11. Definir nível de maturidade da abertura.12. Analisar o nível de sigilo dos dados e informações.13. Analisar relatórios anuais e documentação existente.14. Analisar o esforço para abertura de dados.15. Fazer e validar mapa de responsabilidades entre conjuntos de dados e unidades de negócio responsáveis.

⁴⁶ *Best Practices for Publishing Linked Data.*

Dados Abertos: transparência governamental
 Perspectivas, cenários e planejamento

	<p>16. Identificar e analisar sistemas de informação que poderão ser objeto da abertura de dados.</p> <p>17. Identificar dados que podem ser conectados.</p>
3. Modelar os dados	<p>18. Gerar cópias de segurança das bases de dados que serão abertas.</p> <p>19. Higienizar os dados.</p> <p>20. Estabelecer rotinas de conversão de dados para formatos legíveis por máquina</p> <p>21. Anonimizar dados sensíveis.</p> <p>22. Modelar rotinas automatizadas (ETL).</p> <p>23. Analisar se os dados serão conectados ou não.</p> <p>24. Estabelecer ou aprimorar documentação de dados (esquemas, vocabulários e ontologias).</p>
4. Especificar uma licença apropriada	<p>25. Adotar Licenças Não restritivas.</p> <p>26. Estabelecer questões-chave para definição de licenças</p> <p>27. Apresentar opções de licenças a serem adotadas.</p>
5. Estabelecer bons identificadores universais (URIs) para dados conectados	<p>28. Utilizar URIs para conectar os dados.</p> <p>29. Estabelecer URIs persistentes, que não se alterem em nenhum momento.</p> <p>30. Proporcionar pelo menos um recurso de dados em formato que seja legível por máquina para cada URI.</p> <p>31. Usar URIs como nomes para as coisas.</p> <p>32. Estabelecer Design simplificado de URIs.</p> <p>33. Utilizar identificadores relacionados a informações do mundo real.</p> <p>34. Usar URIs HTTP para que recursos de dados possam ser encontrados via Web por pessoas e máquinas.</p> <p>35. Estabelecer URIs neutras.</p> <p>36. Utilizar datas em URIs com moderação.</p> <p>37. Utilizar hashes (#) em URIs cautelosamente.</p> <p>38. Garantir que as URIs das entidades (conjuntos de dados ou recursos) sejam diferentes das URIs das páginas que apresentam estes recursos para a leitura feita por humanos.</p>
6. Utilizar	<p>39. Estabelecer metadados obrigatórios.</p>

Dados Abertos: transparência governamental
Perspectivas, cenários e planejamento

vocabulários padrão	<ul style="list-style-type: none">40. Criar um esquema de dados para cada conjunto de dados.41. Incentivar o reuso de vocabulários.42. Publicar esquemas de dados em arquivos diferentes.43. Determinar linguagens para expressar esquemas de dados.44. Estabelecer critérios de escolha de vocabulários.45. Certificar que os dados estão conectados a outros conjuntos de dados.46. Desenvolver ou utilizar ontologias para estruturar a semântica dos dados.
7. Converter e enriquecer dados	<ul style="list-style-type: none">47. Converter dados para múltiplas finalidades e usos.48. Adotar rotinas ETL para enriquecimento de dados.49. Estabelecer bons links com outros conjuntos de dados.50. Permitir o envolvimento de várias pessoas na identificação de como os dados a serem convertidos se relacionam com outros dados.51. Utilizar rotinas automatizadas de conversão de dados, como a triplicação, quando possível.52. Converter dados em várias serializações RDF.
8. Prover acesso automatizado aos dados	<ul style="list-style-type: none">53. Disponibilizar bases completas para download (dumps).54. Estabelecer um Mapa de Decisões Tecnológicas.55. Desenvolver uma API.56. Desenvolver um <i>endpoint</i> SPARQL.
9. Anunciar os conjuntos de dados para o público	<ul style="list-style-type: none">57. Publicar metadados junto aos dados58. Estabelecer dados tecnicamente e legalmente abertos.59. Disponibilizar os dados com o menor custo possível ao usuário, preferencialmente de modo gratuito na internet.60. Divulgar dados em meios complementares (Catálogos, FTP, Torrent).61. Divulgar dados em seções destacadas de sítios de governo.62. Estabelecer recursos de consulta parcial da base de dados como uma API ou Webservice.63. Estabelecer visualizações e demais recursos de exploração dos dados.64. Melhorar os dados para que sejam mais facilmente encontrados por máquinas.

	65. Disponibilizar dados conectados em servidores de triplas.
10. Estabelecer um contrato social para os dados publicados	66. Estabelecer mecanismos de monitoramento e avaliação da oferta de dados disponibilizados ao público. 67. Estabelecer espaços para recebimento do feedback do usuário, preferencialmente publicando dados de uma pessoa e/ou telefone de contato para esclarecimento de dúvidas sobre o uso e disponibilidade dos dados. 68. Disponibilizar leis e atos normativos que explicitem aos usuários às obrigações dos governos em publicarem dados com qualidade e disponibilidade. 69. Estabelecer com clareza que o processo de publicação contempla etapas de manutenção e atualização dos dados. 70. Utilizar tecnologias que mantenham os dados conectados disponíveis, atualizados e abertos

Tabela 9 – Recomendações para publicação de dados abertos governamentais conectados.

Deste conjunto cabe destacar a recomendação número 10 que trata da seleção dos conjuntos de dados que serão publicados. Segundo o W3C (2014), a prioridade deve ser atribuída de acordo com os seguintes critérios:

- Dados catalogados ou criados pela instituição publicadora.
- Dados que ao serem combinados com outros dados produzam grande valor.
- Dados com potencial de reuso.
- Que informação é entregue com maior frequência aos cidadãos através dos meios de solicitação de acesso às informações públicas?
- Que informações consideradas pela instituição como de interesse público são entregues à imprensa com maior frequência?
- Que informações só o entregues para outras instituições regularmente e que podem ser ofertadas amplamente ao público?
- Que informações da sua instituição atendem aos requisitos de dados abertos e podem ser publicadas?
- Que informações da sua instituição ainda não atendem aos requisitos de dados abertos, mas que podem ser facilmente convertidas para dados abertos?

- Que informações são solicitadas habitualmente e que exigem um processamento de dados para serem entregues?

Além desses critérios deve-se priorizar dados de natureza geoespacial, saúde, legislação, população e demografia por serem dados muito demandados.

9 Considerações finais

"A mudança não acontecerá se nós esperarmos por outra pessoa ou se esperarmos por algum outro momento. Nós somos as pessoas pelas quais esperávamos. Nós somos a mudança que buscamos."

Barack Obama

A disponibilização de dados governamentais abertos tem o potencial de promover a inclusão digital ao fornecer dados em formatos padronizados e abertos que podem ser transformados e colocados à disposição em formato que respeite os padrões de acessibilidade possibilitando a sua utilização por pessoas com deficiência. Tem o potencial de prover mais transparência ao prover dados que podem ser reutilizados por terceiros livremente permitindo uma visão mais ampla das ações de governo. E se os dados governamentais abertos forem “confrontados” com outros dados de outras fontes, será possível obter novos significados sobre o desempenho do governo, o que demanda maior responsabilidade dos agentes públicos. Somente os casos de uso em andamento em diversos países e em fase embrionária no Brasil poderão concluir se o potencial dos dados governamentais aberto é realizável.

A diversidade brasileira habilita a política externa do País a dialogar com as distintas realidades encontradas ao redor do mundo. De um lado, por exemplo, unimo-nos aos demais países em desenvolvimento na busca da ampliação do acesso à Internet, ainda concentrado e caro; de outro, somos reconhecidos internacionalmente pela capacidade inovadora e criativa do brasileiro na Internet, exemplificada pela participação ativa do País na rede, sua grande comunidade de desenvolvedores de softwares livres, bem como pela informatização de nosso sistema bancário e de nosso processo eleitoral.

“Abertura”, a palavra-chave do desenvolvimento da Internet e de sua governança, inspira a prática cotidiana da política externa brasileira para a Sociedade da Informação. E inspira também essa edição do “Congresso Governo Eletrônico e Software Livre – CONSEGI”, que tem cumprido importante papel na aproximação do governo, a diplomacia incluída, com os temas mais atuais do mundo digital.

10 Lista de siglas e abreviaturas

1. API - *Application Programming Interface*.
2. APT - GET - *Advanced Packaging Tool*.
3. BLPD - *Best Practices for Publishing Linked Data*.
4. CC0 - *Creative Commons Zero*.
5. CETI - Comitê Estratégico de Tecnologia da Informação.
6. CIGA - Comitê Interministerial Governo Aberto.
7. CKAN - *Comprehensive Knowledge Archive Network*.
8. COAT - Coordenação de Articulação, Geração e Aplicação de Tecnologia.
9. CONCAR - Comissão Nacional de Cartografia.
10. CSS - *Cascading Style Sheets*.
11. CSV - *Comma-Separated Values*
12. DAC - Dados Abertos Conectados.
13. DACG - Dados Abertos Conectados Governamentais.
14. DAG - Dados Abertos Governamentais.
15. DbCL - *Database Content License*.
16. EGTI - Estratégia Geral de Tecnologia da Informação.
17. ESAF - Escola Superior de Administração Fazendária.
18. GML - *Geography Markup Language*.
19. GNU - General Public License.
20. GTFS - *General Transit Feed Specification*.
21. HTML – *Hyper Text Markup Language*.

22. IBICT - Instituto Brasileiro de Informação em Ciência e Tecnologia.
23. INDA - Infraestrutura Nacional de Dados Abertos.
24. INDE - Infraestrutura Nacional de Dados Espaciais.
25. IoT – *Internet of Things*.
26. JDK - *Java Development Kit*.
27. JSON - Notação de Objetos JavaScript.
28. LAI – Lei de Acesso à Informação.
29. *LOD Cloud – Linking Open Data Cloud*.
30. MGB - Metadados Geoespaciais do Brasil.
31. MPOG - Ministério do Planejamento, Desenvolvimento e Gestão.
32. ODbL - *Open Database License*.
33. ODS – *Open Document Spreadsheet*.
34. OKBr - *Open Knowledge Brasil*.
35. OKI – *Open Knowledge Internacional*.
36. ORCID - *Open Researcher and Contributor ID*.
37. PBH – Prefeitura de Belo Horizonte.
38. PDA - Plano de Dados Abertos.
39. PDF - *Portable Document Format*.
40. PDTI - Plano Diretor de Tecnologia da Informação.
41. PEI - Planejamento Estratégico Institucional.
42. PETI - Planejamento Estratégico de Tecnologia da Informação.
43. RDF - *Resource Description Framework*.

44. RDFa - É a incorporação de marcações semânticas com uso de RDF sobre as tags HTML de uma página.
45. SBGD - Sistema Gerenciador de Banco de Dados.
46. SMTP - *Simple Mail Transfer Protocol*.
47. SNJ - Secretaria Nacional de Juventude.
48. SOLR - *Open Source Enterprise Search*.
49. SVG - *Scalable Vector Graphics*.
50. TXT - Arquivo de texto.
51. URI - Universal Resource Identifier/Identificador Universal de Recursos.
52. URL - Universal Resource Locator/Localizador Padrão de Recursos.
53. VCGE – Vocabulário Controlado do Governo Eletrônico.
54. W3C - *World Wide Web Consortium/Consórcio da World Wide Web*.
55. XML - *Extensible Markup Language*.

11 Glossário

1. *Accountability*. Prestação de contas. É a definição dos objetos sobre as quais se prestarão contas juntamente com a sua responsabilização, vem juntamente com a transparência das ações que gera confiança entre os governantes e a sociedade.
2. Anonimação. Tornar um dado anônimo não sendo possível associá-lo a uma pessoa. Esta prática é usual quando se trata de dado sigiloso.
3. Área temática. Informação que organiza os datasets em temas de acordo com a lógica de organização dos serviços da instituição responsável pelo dataset. As áreas temáticas da Prefeitura de Belo Horizonte, utilizadas como parte do exemplo desenvolvido neste texto, estão listadas no Anexo 1.
4. *Barcamp*. Rede internacional formada em torno de um modelo de conferência, geralmente referenciado com desconferência.
5. *Big Data*. É o termo em Tecnologia da Informação (TI) que trata sobre grandes conjuntos de dados que precisam ser processados e armazenados,
6. *Checklist*. É um instrumento de controle, composto por um conjunto de condutas, nomes, itens ou tarefas que devem ser lembradas e/ou seguidas.
7. Cidadão. É o usuário dos dados disponibilizados no Portal de Dados Abertos e das aplicações.
8. *Citymapper*. Aplicativo gratuito para smartphones criado em 2011 em Londres por um ex-funcionário do Google com o objetivo de melhorar a relação da população com o transporte público.
9. Cobertura. Abrangência geográfica ou temporal de um dado.
10. Dado público. Qualquer dado gerado ou sob a guarda governamental que não tenha o seu acesso restrito por legislação específica.
11. Dado. Sequência de símbolos ou valores, representados em algum meio, produzidos como resultado de um processo natural ou artificial. Entende-se que dados são observações ou o resultado de uma medida (por investigação, cálculo ou pesquisa) de aspectos característicos da natureza, estado ou condição de algo de interesse, que são descritos através de representações formais e, ao serem

apresentados de forma direta ou indireta à consciência, servem de base ou pressuposto no processo cognitivo

12. Dados Abertos. São os dados que estão disponíveis ao cidadão, sem custo ou restrições de qualquer ordem proveniente de qualquer origem, seja privado ou público.
13. Dados Governamentais Abertos (DGA). São os dados abertos que instituições governamentais disponibilizam ao cidadão.
14. Dados pessoais. Dados privativos de cada pessoa.
15. Dados sigilosos. São os dados submetidos temporariamente à restrição de acesso público em razão de sua imprescindibilidade para a segurança da sociedade e do Estado”.
16. *Dataset*. Conjunto de dados. É o arquivo gerado a partir da base de dados original. Precisa seguir normas e padrões de dados abertos para que sua publicação seja aceita no Portal.
17. Dicionário de dados. Acervo que descreve e define o significado de todos os dados que compõem o universo de informação de um sistema e permite fazer a verificação de consistência entre os vários modelos.
18. Direito autoral. Direito de propriedade sobre uma criação.
19. Ecossistema de *software*. Uma rede de sistemas interligados enviando e recebendo dados via internet.
20. ePing. Concebida como uma estrutura básica para a estratégia de governo eletrônico, aplicada ao governo federal – Poder Executivo, não restringindo a participação, por adesão voluntária, de outros Poderes e esferas de governo.
21. Estado Democrático. Baseado no cumprimento por parte dos governos das normas de Direito e no chamado Estado social de Direito ou Estado de bem-estar social, que compreende uma série de medidas que devem ser atendidas pelo Estado soberano para tornar digna a vida da população..
22. Ferramenta de acesso. Programa ou aplicativo utilizado para acessar os dados em um servidor de dados.

23. FIWARE. Infraestrutura em nuvem de plataforma aberta para IoT. Estrutura financiada pela Comissão Europeia e pela União Europeia, baseada no *OpenStack*⁴⁷, plataforma open source para computação em nuvem.
24. Fontes de dados. São os dados normalmente provenientes de sistemas transacionais e que devem ser extraídos e carregados na Área de Transporte de Dados (Área Intermediária ou Stage). Fontes de dados estruturadas de outras formas e dados de fontes externas à empresa também podem vir a ser considerados.
25. Formato aberto: formato de arquivo não proprietário, cuja especificação esteja documentada publicamente e seja de livre conhecimento e implementação, livre de patentes ou qualquer outra restrição legal quanto à sua utilização.
26. Formato. Estrutura do dado que permite que seja compreendido por um programa leitor ou interpretador.
27. Gestão do Tempo do projeto. Inclui os processos necessários para alcançar a conclusão do projeto a tempo.
28. Gestão dos Riscos do projeto. Inclui os processos relacionados com o planejamento da gestão de riscos, a identificação e a análise de riscos, as respostas aos riscos e o seguimento e controle de riscos de um projeto.
29. Gestor de dados. Profissional da instituição publicadora responsável por estabelecer as diretrizes do processo de abertura na instituição.
30. Governança. A governança dos assuntos de qualquer instituição, incluindo instituições não governamentais.
31. Hacker. Indivíduo que se dedica, com intensidade incomum, a conhecer e modificar os aspectos mais internos de dispositivos, programas e redes de computadores.
32. Hospedagem de dados. Estrutura de tecnologia responsável por armazenar dados. Um banco de dados exerce esta função.

⁴⁷ <https://www.openstack.org/>.

33. INDA. A Infraestrutura Nacional de Dados Abertos (INDA) é um conjunto de padrões, tecnologias, procedimentos e mecanismos de controle necessários para atender às condições de disseminação e compartilhamento de dados e informações públicas no modelo de Dados Abertos, em conformidade com o disposto na ePING.
34. LAI. Lei de Acesso à Informação. Lei nº 12.527, de 18 de novembro de 2011, é uma lei ordinária federal que regulamenta o art. 5º, XXXIII, art. 37, §3º, II e art. 216, §2º da Constituição Federal de 1988, que asseguram o direito fundamental de acesso às informações produzidas ou armazenadas por órgãos e entidades da União, Estados, Distrito Federal e Municípios.
35. Lei de Acesso à Informação. Ver LAI.
36. Licença aberta. É uma forma de o titular de direitos de autor (criador ou outro titular de direitos) conceder ao público em geral autorização para utilizar a sua obra.
37. Licença de uso. É o documento que define os limites de uso que um usuário pode ter em relação a um produto de terceiros.
38. Linked Data: conjunto de boas práticas para publicação de dados estruturados de maneira a facilitar seu compartilhamento e integração.
39. *Linked Open Data*. Dados abertos conectados.
40. *Meetup*. É um encontro informal em que as pessoas conversam de pé, facilitando a circulação e o networking.
41. Metadado. Os dados, especialmente os dados digitais, assumem muitas formas. As conversas por voz, as mensagens de texto ou as redes sociais comunicam dados. As transações digitais bancárias ou comerciais envolvem a transferência de dados. Conteúdo na web, entretenimento digitalizado e transferido, bancos de dados e repositórios de informações de todos os tipos são exemplos de publicações de dados. Os metadados descrevem o que são esses dados: eles fornecem informações sobre esses dados. Isso é bastante simples. No entanto, se analisarmos isso em detalhe, descobrimos que "descrever" os dados

é um exercício técnico rigoroso e um problema carregado de implicações sócio-políticas (PISCITELLO, 2016).

42. **Negociação.** Implica debater com outros para entender-se com eles ou chegar a um acordo.
43. **Ontologia.** É uma especificação formal e explícita de uma conceitualização compartilhada. Nessa definição, "formal" significa legível para computadores; "especificação explícita" diz respeito a conceitos, propriedades, relações, funções, restrições, axiomas, explicitamente definidos; "compartilhado" quer dizer conhecimento consensual; e "conceitualização" diz respeito a um modelo abstrato de algum fenômeno do mundo real.
44. **Órgão.** Integram a estrutura do Estado e das demais pessoas jurídicas como partes desses corpos vivos, dotados de vontade e capazes de exercer direitos e contrair obrigações para a consecução de seus fins institucionais.
45. **Origem dos dados.** Todo elemento de dado precisa identificar sua origem ou o processo que o gera. Esta identificação é muito importante para que se possa ter informações sobre a fonte geradora do dado. Esta informação deve ser única, ou seja, cada dado deve ter uma e somente uma fonte de origem.
46. **Papel.** É uma definição precisa e bem delimitada de uma função e responsabilidades a serem desempenhadas por uma ou mais pessoas. Um papel pode ser desempenhado por mais de uma pessoa simultaneamente, por exemplo. um arquiteto (papel) pode ser interpretado por um grupo de desenvolvedores com habilidades complementares. Para desempenhar um determinado papel de maneira adequada o desenvolvedor deve possuir as habilidades necessárias à sua interpretação.
47. **Participação.** Um dos três eixos sustentadores do Governo Aberto. Permite aos cidadãos contribuir com suas ideias e competências, auxiliando o poder público a elaborar políticas eficazes e abrangentes, graças à informação amplamente dispersa em nossa sociedade.
48. **PDA.** Plano de Dados Abertos. É o documento orientador para as ações de implementação e promoção de abertura de dados, inclusive geoespacializados obedecendo a padrões mínimos de qualidade, de forma a facilitar o entendimento e a reutilização das informações da instituição da administração

pública. Sua elaboração vem ao encontro do disposto na Lei de Acesso à Informação (LAI), na Instrução Normativa SLTI nº 4, de 13 de abril de 2012 (que institui a Infraestrutura Nacional de Dados Abertos), o Decreto Presidencial nº 6.666, de 27 de novembro de 2008 (que institui a Infraestrutura Nacional de Dados Espaciais), bem como dos compromissos assumidos pelo Brasil no âmbito do 2º Plano de Ação Nacional sobre Governo Aberto, entre outros normativos que abordam o tema de transparência.

49. Perfil de acesso. Forma de identificação de usuários que podem ler, atualizar, excluir ou inserir dados na base.
50. Planejamento. Fase do ciclo de vida de um projeto composta por processos para planejar e administrar com sucesso um projeto; para isto, desenvolvem o Plano de gestão, que inclui principalmente o alcance do projeto, seus custos, as atividades a serem desenvolvidas, os recursos humanos e materiais necessários, as comunicações necessárias, os riscos a serem enfrentados, a qualidade a ser garantida e as necessidades de aquisições e compras.
51. Plano de Ação. Documento que descreve o que será feito, em que prazo, por quem.
52. Plano de Dados Abertos. Ver PDA.
53. Plano. Documento que descreve em linhas gerais como um objetivo será alcançado e o que será necessário para alcançá-lo.
54. Portal de Dados Abertos. Sítio eletrônico de referência para a busca e o acesso aos dados públicos, seus metadados, informações, aplicativos e serviços relacionados.
55. Prestação de serviço. Um dos três eixos sustentadores do Governo Aberto. Aprimora a eficácia do Estado, encorajando a cooperação entre a sociedade, os diferentes níveis de governo e a iniciativa privada.
56. Procedimento. Descrição passo-a-passo de uma sequência de tarefas para a realização de uma atividade. Descreve tarefas a serem executadas e identifica regras para desenvolvê-las.

57. Processo iterativo. Em um contexto de ciclo de vida de software, é o tipo de processo que envolve o gerenciamento de uma cadeia de versões (releases) executáveis.
58. Processo. É um conjunto de ações e atividades inter-relacionadas levadas a cabo para alcançar um conjunto previamente definido de produtos, resultados ou serviços.
59. Produto. Um artigo produzido que é quantificável e que pode ser um elemento terminado ou um componente.
60. Projeto. Esforço temporal levado a cabo para criar um produto, serviço ou resultado único.
61. *Python*. É uma linguagem de programação de alto nível, interpretada, de script, imperativa, orientada a objetos, funcional, de tipagem dinâmica e forte.
62. Recursos: Recursos humanos especializados, equipes, serviços, fornecimentos, matérias primas, materiais, orçamentos ou recursos necessários para realizar as atividades do cronograma.
63. Resolução de problemas. Implica tanto a definição do problema (causas e sintomas) quanto a tomada de decisões (analisar o problema para identificar soluções viáveis e tomar uma decisão).
64. Responsabilidade de informações. Deve ser identificado o responsável por cada elemento de dados do data warehouse e também o responsável pela entrada de metadados.
65. Resultado. Saídas ou documentos gerados a partir de uma fonte de informação.
66. Servidor de aplicação. É um computador que disponibiliza um ambiente para a instalação e execução de certas aplicações, centralizando e dispensando a instalação nos computadores clientes.
67. Sistemas de Informação. Um sistema, automatizado ou manual, que compreende pessoas, máquinas, e/ou métodos organizados para coletar, processar, transmitir e disseminar dados que representam informação para o usuário. Oracle é um sistema de banco de dados que surgiu no final dos anos 70,

quando Larry Ellison vislumbrou uma oportunidade que outras companhias não haviam percebido ao encontrar a descrição de um protótipo funcional de um banco de dados relacional e descobriu que nenhuma empresa tinha se empenhado em comercializar essa tecnologia.

68. *Speedgeek*. É o processo de participação utilizado para ver rapidamente um número de apresentação com período fixo para cada uma.
69. *Template*. Modelo de documento com apenas a apresentação visual (apenas cabeçalhos por exemplo) e instruções sobre onde e qual tipo de conteúdo deve entrar a cada parcela da apresentação — por exemplo conteúdos que podem aparecer no início e conteúdos que só podem aparecer no final.
70. *Torrent*. Maneira de compartilhar arquivos entre usuários, sem que o arquivo precise estar em um servidor (computador central de alto desempenho).
71. *Transparência*. Um dos três eixos sustentadores do Governo Aberto. Trata da promoção da responsabilidade de informar os cidadãos sobre o que o governo está fazendo e que ações pretende tomar.
72. *Usabilidade*. Termo usado para definir a facilidade com que as pessoas podem empregar uma ferramenta ou objeto a fim de realizar uma tarefa específica e importante.
73. *Wikipedia*. Projeto de enciclopédia colaborativa, universal e multilíngue estabelecido na internet sob o princípio wiki.

12 Perguntas frequentes

1. A equipe responsável pela elaboração do PDA deve ser composta apenas por profissionais de tecnologia?

Não. A equipe deve ser composta por diversos profissionais da instituição publicadora, tais como:

- Integrantes da alta direção,
- Gestores das áreas envolvidas no PDA.
- Funcionários que detêm o conhecimento sobre os dados produzidos e custodiados pela instituição.
- Profissionais de TI responsáveis pelos sistemas de informação que serão origem dos dados a publicar.
- Profissionais de TI responsáveis pela estruturação do Portal de Dados Abertos da instituição.

2. A governança do PDA é exercida diretamente pela alta direção da instituição publicadora?

Não. A estrutura de governança dos dados abertos envolve atores com papéis e responsabilidades definidas para que não se tenha o risco do PDA ser objeto de manobra política ou de manipulação para demonstração de resultados filtrados ou censurados pela direção da instituição publicadora. Um grupo de participantes da governança deve ter participantes tais como:

- Comitês Gestores, responsáveis pelas diretrizes mestras da política de abertura de dados da instituição;
- Secretaria Executiva, incumbida da organização da governança dos dados abertos;
- Força Tarefa para análise e publicação. A qual se atribui a manutenção e revisão dos conjuntos de dados abertos;
- Força Tarefa para Curadoria, encarregada pela curadoria dos metadados;
- Áreas Departamentais responsáveis pelos dados e suas regras negociais.

3. A minha instituição já disponibiliza os dados gratuitamente na Internet. Posso dizer que são dados abertos?

A gratuidade não significa abertura. É fundamental que sejam explicitamente garantidos os princípios mencionados a seguir para um recurso ser considerado aberto.

- Deve ser permitida a reutilização e redistribuição dos dados. Qualquer restrição à sua edição, modificação e redistribuição impede que sejam considerados abertos.
- Não pode haver restrições a áreas específicas. Por exemplo, um dataset não é aberto quando é apenas permitido o uso pessoal, não-comercial e/ou acadêmico.
- É preciso que estejam disponíveis em formatos abertos de modo a não obrigar o público a utilizar ferramentas de software específicas para a eles poder aceder.

4. A que se refere o Contexto Institucional no PDA?

O Contexto Institucional trata da avaliação e do detalhamento de elementos internos e externos que impactam a política de abertura de dados da instituição. Tem-se como exemplos desses elementos os instrumentos e instâncias de Gestão a seguir:

- Planejamento Estratégico Institucional (PEI).
- Planejamento Estratégico de Tecnologia da Informação (PETI).
- Plano Diretor de Tecnologia da Informação (PDTI).
- Estratégia Geral de Tecnologia da Informação (EGTI).
- Plano de Ação da Infraestrutura Nacional de Dados Abertos – INDA.
- Normas da Comissão Nacional de Cartografia – CONCAR - relacionadas à Infraestrutura Nacional de Dados Espaciais – INDE.

5. A que tipo de informação os cidadãos podem ter acesso pela Lei de Acesso?

Com a Lei de Acesso, a publicidade passou a ser a regra e o sigilo a exceção. Dessa forma, as pessoas podem ter acesso a qualquer informação pública produzida ou custodiada pelos órgãos e entidades da Administração Pública. A Lei de Acesso, entretanto, prevê algumas exceções ao acesso às informações, notadamente àquelas cuja divulgação indiscriminada possa trazer riscos à sociedade ou ao Estado.

6. Como abrir arquivos CSV?

Os arquivos do tipo CSV constituem-se em arquivos texto, onde os valores são delimitados por um caractere, normalmente a vírgula, ponto e vírgula ou tabulação. Menos frequentemente, é possível também que os campos sejam delimitados por uma largura fixa de caracteres.

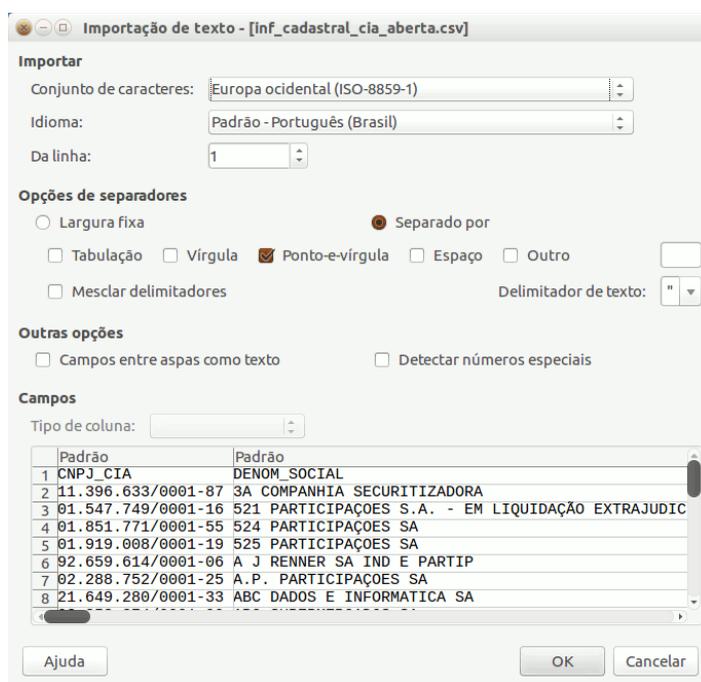
Podem ser abertos em qualquer editor de planilhas, como o LibreOffice/OpenOffice ou Excel. Nesse caso, dependendo da versão do Excel, pode ser necessário executar um passo adicional para converter texto em colunas.

Como o formato CSV não segue um padrão único, geralmente é necessário ajustar parâmetros, tais como o caractere delimitador (ex.: vírgula, ponto e vírgula ou tabulação) e a codificação de caracteres (geralmente UTF-8 ou ISO-8859-1), de acordo com a forma em que o arquivo tiver sido gravado. Por exemplo, o aparecimento de caracteres estranhos no lugar da acentuação das palavras é um sinal de que a codificação de caracteres selecionada ao abrir o arquivo deve ser substituída.

Caso o arquivo CSV seja muito grande, pode não ser viável abri-lo em um editor de planilhas, como o Excel ou LibreOffice, pois pode não haver memória suficiente para fazê-lo. Nesses casos, recomenda-se utilizar ferramentas especializadas, frequentemente utilizadas em projetos de ciência de dados, mineração de dados ou big data. Outra possibilidade é utilizar uma

linguagem de programação, tal como Python ou R, para fazer operações com os dados.

Arquivos CSV também podem ser importados em sistemas gerenciadores de bancos de dados (SGBD). Consulte o manual do seu SGBD⁴⁸ caso deseje fazer a importação de uma tabela a partir de um arquivo CSV.



Exemplo de janela de abertura de arquivo CSV na ferramenta LibreOffice.

7. Como abrir arquivos JSON?

Arquivos JSON, a depender do seu tamanho, podem ser visualizados diretamente no navegador. Entretanto, caso o arquivo seja grande, o uso intensivo de memória e processamento podem causar travamento da tela.

Para fazer algo de útil com os arquivos, é necessário utilizar ferramentas especializadas, frequentemente utilizadas em projetos de ciência de dados, mineração de dados ou big data. Outra possibilidade é utilizar uma

⁴⁸ Sistema Gerenciador de Banco de Dados

linguagem de programação, tal como Python ou R, para fazer operações com os dados.

A estrutura de dados do formato JSON foi especialmente concebida para ser de leve processamento e fácil manipulação em linguagens de programação.

8. Como abrir arquivos XML?

Arquivos XML, a depender do seu tamanho, podem ser visualizados diretamente no navegador. Entretanto, caso o arquivo seja grande, o uso intensivo de memória e processamento podem causar travamento da tela.

Para fazer algo de útil com os arquivos, é necessário utilizar ferramentas especializadas, frequentemente utilizadas em projetos de ciência de dados, mineração de dados ou big data. Outra possibilidade é utilizar uma linguagem de programação, tal como Python ou R, para fazer operações com os dados.

A estrutura de dados do formato XML é especialmente favorável nas situações em que os dados precisam ser mesclados com conteúdo de texto. Por isso, é considerada uma linguagem de markup.

9. Como eu posso compartilhar meus aplicativos e visualizações?

A Internet é o melhor ambiente para o compartilhamento de conteúdo digital. A forma mais rápida para você compartilhar é publicando em um blog ou qualquer outro ambiente acessível publicamente na Internet, e assim espalhar o endereço, a URL, do seu aplicativo. Além disso, tem-se a intenção de ordenar e democratizar o acesso a serviços e aplicativos construídos sobre os dados abertos disponibilizando-os em um portal.

10. Como se nivela o conhecimento dos participantes em um projeto de elaboração de um PDA?

Para nivelar o conhecimento sugere-se que se faça oficinas que tratam do tema, se promova cursos sobre a Política de Dados Abertos e demais

legislação. Além disto a participação em eventos do tema tende a gerar redes de conhecimento com parceiros de outras instituições.

11. É necessário que uma instituição tenha um PDTI para elaborar seu PDA?

Não. Mas é interessante que a publicação de dados abertos passe a integrar a rotina da instituição e, como consequência, seja item do seu planejamento.

12. Em que consistem os indexadores?

Utiliza-se tags a associação ao órgão publicador e uma área de assunto.

13. Existe alguma estrutura recomendada para um PDA?

Sim. Sugere-se a seguinte estrutura para o PDA.

- Capa.
- Introdução.
- Contexto institucional.
- Objetivos.
- Processo de Abertura de dados.
- Governança.
- Planos de ação para abertura de dados.
- Divulgação.

14. Há alguma orientação sobre como deve ser o processo de abertura dos dados?

Sim. A abertura de dados deve se orientar pelos seguintes passos:

1. Levantamento do conjunto de dados candidato à abertura.
2. Priorização e seleção dos dados que serão abertos.
3. Definição de responsáveis pelo preparo e atualização dos dados e detalhamento de plano de ação com metas e prazos.
4. Consolidação da matriz de responsabilidades e definição da governança e do fluxo de aprovação do PDA e revisões.
5. Utilização de metodologia de abertura de dados a ser seguida pelas áreas responsáveis (padrões INDA e INDE).
6. Capacitação dos responsáveis nas áreas de negócio dos dados selecionados para abertura sobre:

- Processo de publicação de dados abertos;
- Processo de catalogação dos metadados no dados.gov.br;
- Processo de catalogação dos metadados na INDE, caso georreferenciados;
- Definição de arquitetura de abertura para cada sistema.

7. Publicização dos dados catalogados, observando-se o uso de URL fixa.

15. No caso de dados abertos quais são as licenças recomendadas?

- *Open Database License (ODbL)*⁴⁹: Qualquer pessoa é livre para copiar, distribuir, usar, criar novas publicações a partir da abertura original, modificar, construir ou transformar os dados (ou o banco de dados). Usuários têm que atribuir a fonte, da forma que for definida na licença e garantir que reusos de seus trabalhos com os dados sigam a mesma regra de atribuição. Nos casos de usos dos dados para fins comerciais, deve ser disponibilizada cópia pública.
- *Database Content License (DbCL)*: mesmas orientações da ODbL, mas aplicável ao conteúdo dos bancos de dados - os dados em si;
- *Creative Commons Zero (CC0)*⁵⁰: aplicável a contextos em que o conjunto de dados caracteriza-se como de domínio público. Usada quando a análise do gestor aponta para a inexistência de direito autoral e passa a ser “natural” a renúncia permanente aos direitos de copyright e outros relacionados. Nenhuma atribuição à instituição gestora dos dados é esperada por parte dos usuários.

⁴⁹ Conteúdo integral da licença disponível em <https://opendatacommons.org/licenses/odbl/>

⁵⁰ Conteúdo integral da licença disponível em <https://creativecommons.org/share-your-work/public-domain/cc0/>.

- *Creative Commons Attribution 4.0*⁵¹: permite ao usuário (qualquer usuário) realizar livremente atividades de compartilhamento (cópia e redistribuição dos dados em qualquer formato ou mídia), de adaptação (incluindo remixagem, transformação e construção de novos produtos) para quaisquer propósitos inclusive comerciais. Ficam os usuários sujeitos às seguintes condições: (i) atribuir crédito apropriado, fornecendo URL da licença e indicando eventuais mudanças; (ii) ao atribuir crédito não é necessário seguir formato padrão, mas é exigido que não seja de forma que leve o leitor a supor que a instituição que licenciou o conteúdo endossa o usuário ou os trabalhos produzidos com os dados licenciados.

16. O *dataset* é o conjunto de dados?

Sim. *Dataset* é o termo em inglês para conjunto de dados. O *dataset* é o conjunto de dados gerado a partir da base de dados original.

17. O portal de dados abertos tem algum dado pessoal ou sigiloso?

Os dados que estiverem sujeitos a restrições de sigilo (art. 25) ou de privacidade (art. 31), conforme definidos na Lei de Acesso à Informação, precisam ser previamente classificados pelos seus responsáveis como tais, a partir de sua origem (art. 27) e por definição não são dados abertos. Assim, eles não passam pelo processo de abertura e não devem ser catalogados no Portal Brasileiro de Dados Abertos, podendo, em caso de descumprimento, o órgão ou entidade ser responsabilizado nos termos do art. 34.

18. O processo de publicação dos dados termina quando o *dataset* é publicado?

Não. O dado publicado se torna foco de controle de qualidade da instituição publicadora. Os *datasets* deverão ser mantidos atualizados e sincronizados com a origem primária das informações, com a menor periodicidade e o maior grau de granularidade viáveis. A atualização dos dados

⁵¹ Conteúdo integral da licença disponível em <https://creativecommons.org/licenses/by/4.0/>

se dará, preferencialmente, por meio de sincronização automática, estabelecendo-se um processo contínuo. Observe, entretanto, que manter atualizado não significa gerar um novo dataset e substituir o anterior. Significa gerar um novo dataset para publicação preservando o anterior. Isto origina uma série histórica que permite aumentar o leque de serviços possíveis pelo desenvolvimento de aplicativos ou pelas análises evolutivas do dataset.

19. O que é a Lei de Acesso à Informação?

A Lei nº 12.527/2011, conhecida como Lei de Acesso à Informação - LAI, regulamenta o direito, previsto na Constituição, de qualquer pessoa solicitar e receber dos órgãos e entidades públicos, de todos os entes e Poderes, informações públicas por eles produzidas ou custodiadas.

20. O que é essencial na priorização do dado a ser publicado?

A priorização dos conjuntos de dados que serão publicados, deve considerar:

- O grau de relevância das informações para o cidadão;
- os compromissos anteriormente assumidos pelo órgão;
- o alinhamento com os instrumentos de planejamento estratégico;
- as competências e
- os conjuntos de dados efetivamente mantidos pela instituição.

21. O que é Infraestrutura Nacional de Dados Abertos - INDA?

A Infraestrutura Nacional de Dados Abertos (INDA) é um conjunto de padrões, tecnologias, procedimentos e mecanismos de controle necessários para atender às condições de disseminação e compartilhamento de dados e informações públicas no modelo de Dados Abertos, em conformidade com o disposto na ePING. A INDA é a política do governo brasileiro para dados abertos.

A INDA é regulamentada através da Instrução Normativa da Infraestrutura Nacional de Dados Abertos.

22. O que é o Plano de Dados Abertos?

O Plano de Dados Abertos – PDA - tem por função registrar o planejamento de publicação dos dados no Portal de Dados Abertos da instituição. Sua elaboração vem ao encontro do disposto na Lei de Acesso à Informação, na Instrução Normativa SLTI nº 4, de 13 de abril de 2012, que institui a Infraestrutura Nacional de Dados Abertos, bem como os compromissos assumidos pelo Brasil no âmbito do Plano de Ação Nacional sobre Governo Aberto, entre outros normativos que abordam o tema de transparência.

23. O que é um componente semântico?

São palavras que dão sentido e contexto ao nome do elemento de dados.

24. O que é uma licença de uso?

Licença de software é o documento que define os limites de uso que um usuário pode ter em relação a um produto de terceiros. A origem do termo remonta ao uso de imagem e de som, como compreendido pelo Direito, e pode restringir a cópia, a distribuição e a adaptação da aplicação.

As licenças de software geralmente fornecem aos usuários finais o direito a uma ou mais cópias do software sem violar os direitos autorais.

25. O que é uma área temática?

Uma área temática reúne informações que se completam, que formam uma unidade coesa e seguem determinada política de produção, gestão e preservação de dados. São exemplos de áreas temáticas a educação, a saúde, o desenvolvimento humano e a administração. A relação completa de áreas temáticas com seus textos explicativos está no Anexo 1.

26. O que é uma informação sigilosa e que não pode ser aberta?

A informação submetida temporariamente à restrição de acesso público em razão de sua imprescindibilidade para a segurança da sociedade e do

Estado.

27. O que são as *tags*?

As *tags*, do inglês etiqueta ou rótulo, são palavras do contexto do dataset que tornam o acesso natural em pesquisas referentes ao seu conteúdo.

28. O que são conjuntos de dados e recursos?

Os dados catalogados no portal estão organizados utilizando as estruturas de conjuntos de dados e recursos. Os conjuntos de dados, *datasets*, são as entidades principais retornadas a partir das buscas. Cada conjunto de dados possui uma descrição, um ou mais recursos, e uma série de outros metadados, como periodicidade de atualização e órgão responsável. Exemplos de conjuntos de dados são uma coleção de tabelas relacionadas entre si, dados extraídos de um mesmo sistema de informações, ou ainda uma API de dados abertos. Um conjunto de dados deve possuir pelo menos um recurso que seja dado aberto.

Cada recurso compreende uma fonte de dados. Pode ser um arquivo como uma planilha, um método de webservice ou um documento.

A princípio cada recurso poderia ser catalogado como um conjunto de dados separado, porém é preferível que os recursos sejam agrupados, sempre que forem oriundos da mesma base de dados, ou possuírem metadados em comum, para facilitar a busca e o entendimento de seu conteúdo.

Simplificando, o conjunto de dados é o agregador dos recursos, que são os arquivos.

29. O que são Dados Abertos Governamentais?

De acordo com o Manual de Dados Abertos do Governo Federal: “são dados produzidos pelo governo e colocados à disposição das pessoas de forma a tornar possível não apenas sua leitura e acompanhamento, mas também sua reutilização em novos projetos, sítios e aplicativos; seu cruzamento com outros dados de diferentes fontes; e sua disposição em visualizações interessantes e

esclarecedoras”.

30. O que são Dados Abertos?

Dados Abertos são dados que estão livremente disponíveis para todos utilizarem e redistribuírem como desejarem, sem restrição de licenças, patentes ou mecanismos de controle. Na prática, a filosofia de dados abertos estipula algumas restrições tecnológicas para que os dados sejam legíveis por máquina.

Todo dado público tem vocação para ser dado aberto. Como praticamente todo dado governamental é público, é fundamental que os governos implementem políticas para disponibilizá-los. Para saber mais, conheça as leis e os princípios dos dados abertos

Segundo a definição da Open Knowledge Internacional, dados são abertos quando qualquer pessoa pode livremente acessá-los, utilizá-los, modificá-los e compartilhá-los para qualquer finalidade, estando sujeito a, no máximo, a exigências que visem preservar sua proveniência e sua abertura. Isso geralmente é satisfeito pela publicação dos dados em formato aberto e sob uma licença de uso aberta. Os dados abertos também são pautados por três leis e oito princípios.

O especialista em políticas públicas e ativista dos dados abertos David Eaves propôs as seguintes “leis”:

1. Se o dado não pode ser encontrado e indexado na Web ele não existe;
2. Se não estiver aberto e disponível em formato compreensível por máquina ele não pode ser reaproveitado; e
3. Se algum dispositivo legal não permitir sua replicação ele não é útil.

As leis foram propostas para os Dados Abertos Governamentais, mas pode-se dizer que elas se aplicam aos Dados Abertos de forma geral. Em 2007 um grupo de trabalho de 30 pessoas reuniu-se na Califórnia, Estados Unidos da América, para definir os princípios dos Dados Abertos Governamentais.

Chegaram a um consenso sobre os 8 princípios para o dado ser considerado aberto:

1. Completos. Todos os dados públicos são disponibilizados. Dados são informações eletronicamente gravadas, incluindo, mas não se limitando a, documentos, bancos de dados, transcrições e gravações audiovisuais. Dados públicos são dados que não estão sujeitos a limitações válidas de privacidade, segurança ou controle de acesso, reguladas por estatutos.
2. Primários. Os dados são publicados na forma coletada na fonte, com a mais fina granularidade possível, e não de forma agregada ou transformada.
3. Atuais. Os dados são disponibilizados o quanto rapidamente seja necessário para preservar o seu valor.
4. Acessíveis. Os dados são disponibilizados para o público mais amplo possível e para os propósitos mais variados possíveis.
5. Processáveis por máquina. Os dados são razoavelmente estruturados para possibilitar o seu processamento automatizado.
6. Acesso não discriminatório. Os dados estão disponíveis a todos, sem que seja necessária identificação ou registro.
7. Formatos não proprietários. Os dados estão disponíveis em um formato sobre o qual nenhum ente tenha controle exclusivo.
8. Livres de licenças. Os dados não estão 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.

31. O que são informações?

De acordo com o art. 4º, inciso I, da Lei nº 12.527/2011, informações

são dados, processados ou não, que podem ser utilizados para produção e transmissão de conhecimento, registrados em qualquer suporte ou formato.

32. O que são licenças? Como posso disponibilizar os meus dados com uma licença aberta?

Uma licença, como o nome sugere, é um conjunto de autorizações que o proprietário dos dados proporciona explicitamente a quem a eles a cede.

Para especificar uma licença para os dados, basta indicar de forma visível qual a licença e um link para a mesma.

33. O que são metadados?

Metadados são geralmente descritos como dados sobre dados. Metadados são essenciais na organização e classificação de datasets. Todo conjunto de dados estruturado faz uso de metadado para descrever a estrutura e os tipos de dados contidos no conjunto.

34. O que se pode fazer com os dados abertos?

Há inúmeras áreas nas quais podemos usar dados abertos para melhorar a vida das pessoas. Podemos encontrar vários casos de sucesso no envolvimento da sociedade civil no processo democrático, na colaboração e participação cívica, na melhoria de serviços públicos, na criação de produtos e serviços inovadores, na investigação científica ou na constituição de modelos de negócio eficaz. Os dados abertos podem constituir uma preciosa matéria-prima para novos projetos por parte da sociedade civil, de iniciativas comunitárias, do setor público ou do setor privado.

35. O que significa dizer que se irá preparar os dados para abertura?

Preparar os dados para abertura implica em identificá-los, formatá-los e obter permissão dos patrocinadores no órgão para a publicação.

36. O que significa dizer que se irá preparar os dados para abertura?

Preparar os dados para abertura implica em identificá-los, formatá-los e obter permissão dos patrocinadores no órgão para a publicação.

37. Onde o PDA é publicado?

No Portal de Dados Abertos da instituição publicadora, no Portal de Dados Abertos do Governo Federal e em outros portais que se deseje.

38. Por que abrir os dados?

A publicação “5 Motivos Para a Abertura de Dados na Administração Pública” (TCU, 2015), elaborada pelo Tribunal de Contas da União, apresenta razões para que as organizações públicas invistam em iniciativas de abertura de dados governamentais. Os cinco motivos para a abertura dos dados são:

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

39. Por que existe tanta informação fechada?

Apesar de todo este potencial, as fontes de dados abertos são ainda escassas. Podemos apontar duas razões principais:

- O desconhecimento da existência de licenças abertas, e o uso de termos de utilização padrão que, por vezes sem intenção por parte das entidades, colocam restrições que tornam os seus dados efetivamente fechados.
- A falta de incentivo à abertura por parte das entidades detentoras da informação. Para resolver este problema, é fundamental a implementação de legislação e medidas claras que consagrem a abertura de dados públicos como procedimento normal e necessário.

40. Quais os formatos recomendados para dados abertos?

JSON, CSV, XML, ODS e RDF.

41. Quais os formatos recomendados para dados abertos?

JSON, CSV, XML, ODS e RDF.

42. Quais são as condições que um conjunto de dados deve cumprir para estar apto a ser publicado?

O conjunto de dados devem cumprir as seguintes condições básicas:

1. Os dados devem estar em seu formato mais bruto possível.
2. Os dados devem estar em formato aberto, não proprietário, estável e de amplo uso.
3. Não deve existir nenhum instrumento jurídico que impeça sua reutilização e redistribuição por qualquer parte da sociedade.
4. Para os dados que são estruturados ou estão em planilhas na sua fonte, deve-se preservar ao máximo a estrutura original.
5. É recomendável a disponibilização dos dados em JSON, XML, CSV, ODS, RDF, SVG, GML HTML/RDFA.
6. Cada conjunto de dados deve possuir um identificador único e persistente, seguindo uma padronização de URL.

43. Quais são os tipos de componentes semânticos?

Componentes de identificação, referência, qualificação e sentença.

44. Quais são os tipos de dados sigilosos?

Ultrassegredo, restrito por até 25 anos, Segredo, restrito por até 15 anos e Reservado, restrito por até 5 anos.

45. Qual a diferença dos portais de dados abertos para os portais de transparência?

A Lei Complementar 131, de 27 de maio de 2009, que alterou a redação da Lei de Responsabilidade Fiscal no que se refere à transparência da gestão fiscal, determina a disponibilização de informações pormenorizadas sobre a execução orçamentária e financeira da União, dos Estados, do Distrito Federal

e dos Municípios. Em outras palavras, os portais de transparências têm o objetivo de aumentar o controle das despesas e receitas no governo.

O portal de dados abertos tem o objetivo de ser o referencial para a busca e o acesso à dados públicos. Ele é um serviço simplificado que organiza e padroniza o acesso aos dados públicos, primando pelo reuso dos dados e o uso de tecnologias modernas. Os portais de dados abertos disponibilizam acesso às bases de dados do governo, sem filtros, completas e seriadas.

46. Qual é a estrutura necessária para publicação do conjunto de dados?

A publicação dos dados no Portal de Dados Abertos da PBH necessita dos seguintes artefatos:

1. Indexadores.
2. Dataset.
3. Dicionário de dados.

Metadados.

47. Qual é o fluxo para publicação do PDA?

Para se publicar o PDA deve-se:

1. Nivelar o conhecimento dos participantes.
2. Estruturar a equipe de trabalho.
3. Estabelecer as metas do PDA.
4. Elaborar o PDA
5. Publicar o PDA

48. Qual é o principal desafio na elaboração do PDA?

O principal desafio na elaboração do PDA é a priorização dos dados a serem publicados. Entretanto, a publicação dos dados pressupõe o enfrentamento de diversos limites e dificuldades inerentes a este processo. Dentre eles pode-se citar:

- Capacitar os envolvidos naquilo que for necessário para sua participação na elaboração e publicação do PDA.
- Executar um inventário dos dados publicáveis na instituição.
- Analisar quais dados são estáveis, confiáveis.
- Realizar uma priorização para que seja possível propor os planos de ação para publicação.
- Analisar os dados quanto a serem dados sensíveis ou sigilosos.
- Estimar o espaço em disco necessário para armazenamento dos dados no Portal de Dados Abertos da instituição.
- Estimar a capacidade de transmissão dos dados na rede da instituição.
- Garantir que os recursos de software e hardware estejam disponíveis para a publicação.

49. Quando a Lei de Acesso à Informação entrou em vigor?

A LAI foi publicada em 18 de novembro de 2011, mas só entrou em vigor 180 (cento e oitenta) dias após essa data, ou seja, em 16 de maio de 2012.

50. Que bons exemplos existem de reutilização de dados abertos?

A Informação pública, na forma de dados abertos, pode ajudar o cidadão a tomar decisões na sua vida pessoal, ou possibilitar-lhe ser mais ativamente interveniente na sociedade.

- Na Dinamarca, o Findtoilet.dk mapeou todos os sanitários públicos dinamarqueses, para que pessoas com problemas de incontinência pudessem ter confiança para sair de casa mais frequentemente.
- O *Mapumental* no Reino Unido e o *Mapnificent* na Alemanha permitem encontrar locais para comprar ou alugar casa, tendo em

conta fatores como a proximidade de transportes públicos ou a duração da deslocação para o emprego.

No campo da economia, os dados abertos também potencializam o aparecimento de novas abordagens e produtos por parte de empresas e organizações.

- Na Dinamarca, o husetsweb.dk ajuda os cidadãos a encontrar meios de melhorar a eficiência energética da sua casa, incluindo o planeamento financeiro e ajudar proprietários a implementar melhorias que levem à poupança de eletricidade. Este projeto baseia-se na reutilização de informações cadastrais e informação relativa a subsídios governamentais, assim como o cadastro local de empresas.
- Ferramentas de tradução online como o *Linguee* ou o *Google Translate* recorrem ao enorme volume de documentos da União Europeia, traduzidos em todos os idiomas europeus e publicados sob termos abertos, para afinar e melhorar os seus algoritmos de tradução.

Os dados abertos também têm valor para o próprio governo e instituições públicas para, entre outras benesses, facilitar a prestação de serviço à sociedade e aos cidadãos.

- O Ministério da Educação holandês publicou na internet os seus dados relacionados com a educação, o que fez diminuir o número de perguntas recebidas, reduziu a carga de trabalho e os custos associados a dar respostas; agora, os funcionários podem também responder às perguntas mais rapidamente, uma vez que a informação está centralizada num portal estruturado.

- No Reino Unido, o *Where Does My Money Go?* mostra como o dinheiro dos impostos é distribuído pelo governo, permitindo um debate mais ampliado e consequente sobre a despesa pública.
- No Canadá, o *OpenParliament* publica versões acessíveis dos debates parlamentares, permitindo aos cidadãos uma visão muito mais profunda e significativa do processo legislativo.

Todos estes exemplos existem porque os governos e instituições públicas desses países disponibilizam a informação necessária usando licenças abertas.

51. Que informações os órgãos e entidades do Poder Executivo federal são obrigados a disponibilizar proativamente em seus sites?

O art. 8º da LAI definiu como um dever dos órgãos e entidades públicos publicar na internet informações públicas de interesse coletivo ou geral. De acordo com o Decreto nº 7.724/2012, os órgãos e entidades do Poder Executivo Federal deverão publicar o seguinte rol mínimo de informações nos seus sítios eletrônicos:

- Estrutura organizacional e competências dos órgãos, além dos endereços e telefones de suas unidades e horários de atendimento ao público;
- programas, projetos, ações, obras e atividades, indicando a unidade responsável, principais metas e resultados e indicadores (se existirem);
- repasses ou transferências de recursos financeiros;
- execução orçamentária e financeira detalhada;
- procedimentos licitatórios, com os contratos celebrados e notas de empenho emitidas;
- remuneração recebida por servidores e empregados públicos de maneira individualizada;
- respostas a perguntas mais frequentes da sociedade;
- contato da autoridade de monitoramento da LAI na instituição e informações

sobre o Serviço de Informações ao Cidadão, e

- informações classificadas e desclassificadas, nos termos do art. 45, I e II do Decreto 7.724/2012.

52. Que tipo de aplicações digitais podem ser feitas com esses dados?

As mais variadas. A sociedade poderá, por exemplo, desenvolver aplicações que auxiliem a análise das propostas de leis em tramitação na Câmara dos Deputados; aplicações que forneçam os resultados das votações de uma comissão; ou que comparem, em tabelas interativas, como os parlamentares usam sua verba indenizatória; ou ainda, que permitem conhecer todos os deputados envolvidos em um determinado tema – direitos das crianças, por exemplo.

53. Quem é responsável pelos dados catalogados no portal?

O Portal de Dados Abertos organiza os dados abertos em um catálogo para fácil localização pelo cidadão. Entretanto, os responsáveis pelos dados são as organizações públicas que os publicam. Por isso cada uma dessas organizações responde pelos seus próprios dados.

Para mais informações ou esclarecimentos sobre um conjunto de dados específico, procure na própria página do conjunto de dados a seção “Informações Adicionais”. Nela há um campo “Autor que se pode contatar.

Exemplo:

Dados Consumidor.gov.br - Janeiro/2015

Defesa do consumidor Justiça

Informações Adicionais

Campo	Valor
Fonte	https://www.consumidor.gov.br/pages/dadosabertos/externo/
Autor	Secretaria Nacional do Consumidor - Senacon/MJ
Mantenedor	Secretaria Nacional do Consumidor - Senacon/MJ
Estado	active
Cobertura geográfica	Brasil
Cobertura temporal	2014 e 2015
Periodicidade	Mensal

Comunidade Manuais Saiba mais

- Comunidades
- Cartilha de publicação de dados
- O que são dados abertos?

Manter os dados *on-line* também é responsabilidade do órgão que o publica. Em alguns casos, essa pode ser uma instituição diferente daquela responsável pelo conteúdo dos dados. Caso um conjunto de dados esteja *off-line* ou apresentando problemas técnicos, procure na própria página do conjunto de dados a seção “Informações Adicionais”. Nela há um campo “Mantenedor”. Procure o nome do órgão que consta no valor desse campo para entrar em contato.

Exemplo:

Dados Consumidor.gov.br - Janeiro/2015

Defesa do consumidor Justiça

Informações Adicionais

Campo	Valor
Fonte	https://www.consumidor.gov.br/pages/dadosabertos/externo/
Autor	Secretaria Nacional do Consumidor - Senacon/MJ
Mantenedor	Secretaria Nacional do Consumidor - Senacon/MJ
Estado	active
Cobertura geográfica	Brasil
Cobertura temporal	2014 e 2015
Periodicidade	Mensal

Comunidade Manuais Saiba mais

- Comunidades
- Cartilha de publicação de dados
- O que são dados abertos?

No caso de não constar e-mail de contato nesses campos, sugere-se procurar o site do próprio órgão e localizar as opções “Institucional”, “Quem é quem” e/ou “Contato”.

13 Bibliografia

- AGUNE, R. M.; GREGORIO FILHO, A. S.; BOLLIGER, S. P. 2010. Governo aberto SP: disponibilização de bases de dados e informações em formato aberto. In: CONGRESSO CONSAD DE GESTÃO PÚBLICA, III, Brasília.
- AKUTSU, L. Portais de governo no Brasil: accountability e democracia delegativa. 2005. In: Congresso Internacional Del Clad Sobre la Reforma Del Estado Y De La Administración Pública., X, Santiago, Chile, 18 - 21 oct.
- ALBANO, C.S.; ARAUJO, M.H. DE. 2013. Inteligência competitiva e monitoramento ambiental utilizando informações disponíveis na internet: a viabilidade de utilizar dados governamentais abertos. Anais do Encontro de Administração da Informação (EnADI), Bento Gonçalves, RS, Brasil.
- ALCANTARA, W. et al. 2015. Desafios no uso de dados abertos conectados na educação brasileira. In: WORKSHOP DE DESAFIOS DA COMPUTAÇÃO APLICADA À EDUCAÇÃO, 4., 2015, Recife, Brasil. ~ Anais ... Recife: CSBC.
- ALFORD, M., LAWSON, J., 1979. Software Requirements Engineering Methodology (Development). RADC-TR-79-168, U. S. Air Force Rome - Air Development Center, Griffiss, AFB, NY (Jun), (DDc-AD-A073132).
- ALVES, C., 2001. “Seleção de Produtos de Software Utilizando uma Abordagem Baseada em Engenharia de Requisitos”. Centro de Informática, Universidade Federal de Pernambuco, Tese de Mestrado, Março/2001.
- ANGÉLICO, F. 2012. Lei de Acesso à Informação Pública e seus possíveis desdobramentos para a accountability democrática no Brasil. 133 f. Dissertação (Mestrado) – Escola de Administração de Empresas de São Paulo, Fundação Getúlio Vargas, São Paulo, 2012.
- ANTON, A. I., McCracken, W. M., Potts, C., 1994. "Goal Decomposition and Scenario Analysis in Business Process Reengineering". In: Proceedings of the 6th International Conference on Advanced Information Systems Engineering (CAiSE'94), Springer, Utrecht, NL (Jun), pp. 94-104.
- ARANGO, G., 1988. Domain Engineering for Software Reuse. Ph.D. Thesis, Department of Computer Science, University of California, Irvine.
- ARAUJO, E. N. de. 2006. Curso de Direito Administrativo. São Paulo: Saraiva.
- ARAUJO, E. N. de. 2006. Curso de Direito Administrativo. São Paulo: Saraiva.

- ÁVILA, Thiago J. T. 2015. Uma proposta de modelo de processo para publicação de Dados Abertos Conectados Governamentais. Disponível em: http://www.consultaesic.cgu.gov.br/busca/dados/Lists/Pedido/Attachments/571437/RESPONSTA_PEDIDO_Thiago%20Avila%20-%20Dissertao%20-%20PPGMCC.pdf. Acesso em 12 dez.2019.
- BANDEIRA, J. et al. 2014. Dados abertos conectados. In: Anais da Jornada de Atualização em Informática na Educação. Maceio: SBTI. Disponível em: <https://www.br-ie.org/pub/index.php/pie/article/view/3551/2937>. Acesso em 11 dez. 2019.
- BARBALHO, F.A. 2014. Emergência de um campo de ação estratégica: o caso de política pública sobre dados abertos. Tese (Doutorado). Universidade de Brasília, Faculdade de Economia, Administração e Contabilidade, Programa de Pós-Graduação em Administração, Brasília, DF.
- BARNAGHI. 2012. Payam et al. Semantics for the Internet of Things: early progress and back to the future. International Journal on Semantic Web and Information Systems (IJSWIS), v. 8, n. 1, p. 1-21..
- BARRETO, A.A. 2000. O Mercado da Informação no Brasil. Inf. Londrina. Paraná. V.5, n.1,p-25-34, jan/jun.
- BERNEERS-LEE, Tim. 2001. The Semantic Web: A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities. Scientific American, v.284, n. 5, p. 34-43.
- BERNERS-LEE, T. 2006. Linked data. Disponível em: <https://www.w3.org/DesignIssues/LinkedData.html>. Acesso em: 18 out. 2019.
- BOBBIO, N. 1987. Estado, governo, sociedade; para uma teoria geral da política. Rio de Janeiro: Paz e Terra.
- BRASIL. 2003. Perspectivas e desafios para a ampliação do governo eletrônico nos municípios brasileiros. In: SEMINÁRIO INTERNACIONAL GOVERNO ELETRÔNICO E GOVERNOS LOCAIS, Brasília.
- BRASIL. 2011. Lei 12.527, de 18 de novembro de 2011. Regula o acesso a informações. Disponível em: http://www.planalto.gov.br/ccivil_03/ato2011-2014/2011/lei/12527.htm. Acesso em 12 jun. 2019.
- BRASIL. 2012. Acesso à Informação. Brasília, DF. Disponível em: <http://www.acessoainformacao.gov.br/assuntos/pedidos/excecoes/excecoes>. Acesso em 19 nov. 2109.
- BRASIL. 2012b. Ministério do Planejamento, Orçamento e Gestão. Secretaria de Logística e Tecnologia da Informação. Instrução Normativa nº 4, de 12 de abril de

2012. Disponível em <http://dados.gov.br/pagina/instrucao-normativa-da-inda>. Acesso em 21 out. 2019.

BRASIL. 2012c. Governo Federal. Cartilha Técnica para Publicações de Dados Abertos no Brasil v.1.0. Brasília, DF. Disponível em: <http://dados.gov.br/pagina/cartilha-publicacao-dados-abertos>. Acesso em: 13 nov. 2019.

BRASIL. 2012d. Governo Federal. TERMO DE USO DOS DADOS ABERTOS. Brasília, DF. Disponível em: [http:// https://dados.gov.br/termo-de-uso.html](http://https://dados.gov.br/termo-de-uso.html). Acesso em: 22 nov. 2019.

BRASIL. 2014. Manual para elaboração de Plano de Dados Abertos. Ministério do Planejamento, Orçamento e Gestão (MPOG). Disponível em: http://www.planejamento.gov.br/secretarias/upload/arquivo/governo-aberto/copy_of_manual_elaboracao_plano_dados_abertos.pdf/@@download/file/manual_elaboracao_plano_dados_abertos.pdf. Acesso em 18 de nov. de 2019.

BRASIL. 2014b. Ministério do Planejamento, Orçamento e Gestão. Secretaria de Logística e Tecnologia da Informação VCGE Vocabulário de Governo Eletrônico / Ministério do Planejamento, Orçamento e Gestão, Secretaria de Logística e Tecnologia da Informação - Brasília: MP, SLTI. Disponível em: https://www.gov.br/governodigital/pt-br/governanca-de-dados/VCGE_2_1_0.pdf. Acesso em 16 dez. 2019.

BRASIL. 2016. Aplicativos e serviços que utilizam dados abertos. Portal Brasileiro de Dados Abertos. Ministério do Planejamento, Orçamento e Gestão. Disponível em: <http://dados.gov.br/aplicativos>. Acessado em 12 jul. 2019.

CAMPOS, A. M. 1990. Accountability: quando poderemos traduzi-la para o português? Revista de Administração Pública, Rio de Janeiro, fev.-abr.

CKAN. About. 2015a. Disponível em: <<http://ckan.org/about/>>. Acesso em: 08 jun. 2019.

DATA REVOLUTION GROUP. 2014. A World That Counts: Mobilising the Data Revolution for Sustainable Development. Disponível em: <https://www.undatarevolution.org/wp-content/uploads/2014/11/A-World-That-Counts.pdf>. Acesso em: 01 ago. 2019.

DAVIES, T. 2010. Open data, democracy and public sector reform. Dissertação – University of Oxford, Oxford. Disponível em: <http://www.opendataimpacts.net/report/>. Acesso em 29 nov. 2019.

DI PIETRO, M. S. Z. 2006. Direito Administrativo. São Paulo: Atlas.

DING, L. et al. 2011. TWC LOGD: A portal for linked open government data ecosystems. Web Semantics: Science, Services and Agents on the World Wide Web, Semantic Web Dynamics Semantic Web Challenge, 2010. v. 9, n. 3, p. 325–333. Disponível

em: <http://tw.rpi.edu/media/latest/twclogdaportalforlinkedopen.pdf>. Acesso em 22 out. 2019.

DINIZ, V. 2010. Como conseguir dados governamentais abertos. In: CONGRESSO CONSAD DE GESTÃO PÚBLICA, III, Brasília. Disponível em: https://i3gov.planejamento.gov.br/como_conseguir_dados_governamentais_abertos.pdf. Acessado em 2 nov. 2019.

DIRECTIVA 2003/98/EC do Parlamento Europeu e do Conselho, relativa a à reutilização de informações do setor público. Jornal Oficial da União Europeia, L 345/90., 17 nov. 2003. Disponível em: <https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2003:345:0090:0096:pt:PDF>. Acesso em: 14 nov. 2019.

EAVES, D. 2009. The three laws of open government. Disponível em: <http://eaves.ca/>. Acesso em: 8 jul. 2018.

FREITAS, José Antônio de Carvalho. 2016. CONTRIBUIÇÕES AO ECOSISTEMA DE DADOS ABERTOS DO GOVERNO FEDERAL COM ENFOQUE EM TECNOLOGIAS CÍVICAS. Universidade Católica de Brasília. Disponível em: https://repositorio.cgu.gov.br/bitstream/1/33201/4/4_Dissertacao_Jose_Antonio.pdf. Acesso em 18 nov. 2019.

GAMA, K.; LÓSCIO, B. F. 2014. Towards Ecosystems based on Open Data as a Service. In: 16TH INTERNATIONAL CONFERENCE ON ENTERPRISE INFORMATION SYSTEMS. Disponível em: <http://www.scitepress.org/DigitalLibrary/Link.aspx?doi=10.5220/0004974506590664>. Acesso em 16 dez. 2019.

HELBIG, N., CRESSWELL, A.M., BURKE, B.G, LUNA-REYES, LUNA-REYES, L. 2012. The Dynamics of Opening Government Data. Disponível em <http://www.ctg.albany.edu/media/pubs/pdfs/opendata.pdf>. Acesso novembro 2019.

HITZLER, Pascal; JANOWICZ, Krzysztof. 2013. Linked Data, Big Data, and the 4th Paradigm. Semantic Web, v. 4, n. 3, p. 233- 235.

HYLAND, B.; WOOD, D. 2011. The joy of data - a cookbook for publishing linked government data on the Web. Disponível em: https://link.springer.com/chapter/10.1007/978-1-4614-1767-5_1. Acesso em: 2 nov. 2019

ISA. 2014. How Linked Data is transforming eGovernment. [S.l.]. Disponível em: https://joinup.ec.europa.eu/sites/default/files/inline-files/D4.3.2_Case_Study_Linked_Data_eGov.pdf. Acesso em: 10 nov. 2019.

- JANSSEN, M.; CHARALABIDIS, Y.; ZUIDERWIJK, A. 2012. Benefits, adoption barriers and myths of open data and open government. *Information Systems Management*, Taylor & Francis, v. 29, n. 4, p. 258–268. Disponível em: https://www.researchgate.net/publication/262276711_Benefits_Adoption_Barriers_and_Myths_of_Open_Data_and_Open_Government. Acesso em 13 nov. 2019.
- KALAMPOKIS, E.; TAMBOURIS, E.; TARABANIS, K.A. 2011. A Classification Scheme for Open Government Data: Towards Linking Decentralized Data. *International Journal of Web Engineering and Technology* 6(3), 266–285.
- KURBALIJA, Jovan. 2016. Uma Introdução à Governança da Internet. São Paulo: Comitê Gestor da Internet no Brasil. Disponível em: https://cgi.br/media/docs/publicacoes/1/CadernoCGIbr_Uma_Introducao_a_Governanca_da_Internet.pdf. Acesso em 3 set. 2019.
- LAPORTE, T.; DEMCHAK, C., FRIIS, C. 2000. Webbing governance: national differences in constructing the face of public organizations. In: GARSON, G. *Handbook of public information systems*. New York: Marcel Dekker Publishers.
- MANYIKA, J.; et al. 2013. Open data: Unlocking innovation and performance with liquid information. Mackinsey & Company, Londres. Disponível em: http://www.mckinsey.com/insights/business_technology/open_data_unlocking_innovation_and_performance_with_liquid_information. Acesso em 6 nov. 2018.
- MARTINS JUNIOR, W. P. 2004. *Transparência administrativa: publicidade, motivação e participação popular*. São Paulo: Saraiva.
- MARTINS JUNIOR, W. P. 2004. *Transparência administrativa: publicidade, motivação e participação popular*. São Paulo: Saraiva.
- MEIJER, Albert J.; CURTIN, Deirdre; HILLEBRANDT, Maarten. 2012. Open government: connecting vision and voice. *International Review of Administrative Sciences*, v. 78, n. 1, p. 10-29.
- MELLO, C. A. B. de. 2008. *Curso de Direito Administrativo*. São Paulo: Malheiros.
- OPEN DEFINITION. 2016. The Open Definition. Disponível em: <http://opendefinition.org/od/2.1/en/>. Acesso em: 18 nov.2019.
- OPEN GOV DATA. 2007. Eight principles of open government data. Disponível em: https://public.resource.org/8_principles.html. Acesso em: 5 jun. 2018.
- OPEN GOV DATA. 2010. Eight principles of open government data. Disponível em: http://resource.org/8_principles.html. Acesso em: 8 jul. 2019.

- OPEN KNOWLEDGE BRASIL. 2013. Maturidade em Dados Abertos. Entenda as 5 estrelas. Disponível em <https://br.okfn.org/2013/01/17/maturidade-em-dados-abertos-entenda-as-5-estrelas/>. Acesso em 8 jun.2019.
- OPEN KNOWLEDGE FOUNDATION. 2010. Open Data Handbook. Disponível em: <http://opendatahandbook.org/guide/en/>. Acesso em: 14 set.2019.
- PEREIRA, Luiz Carlos; CUNILL GRAU, Nuria. 1999. O público não-estatal e a reforma do Estado. Rio de Janeiro: Editora Fundação Getúlio Vargas.
- PINHO, J. A. G. 2008. Investigando portais de governo eletrônico de estados no Brasil: muita tecnologia, pouca democracia. Revista de Administração Pública (RAP), v. 42, n. 3, p. 471-93, mai.-jun.
- PISCITELLO, David. 2016. O que são metadados? Disponível em: <http://https://www.icann.org/news/blog/parte-i-o-que-sao-metadados>. Acessado em 15 jan. 2020.
- PRADO, O.; LOUREIRO, M. R. 2006. Governo eletrônico e transparência: avaliação da publicização das contas públicas das capitais brasileiras. Revista Alcance, vol.13, n.3, p. 355-372, set./dez. Disponível em: <https://www6.univali.br/seer/index.php/ra/article/view/126>. Acessado em 2 mai. 2019.
- RIBEIRO, M. M. 2009. Transparência nos portais do governo federal: os casos do COMPRASNET e do Portal da Transparência. (Trabalho de Conclusão de Curso), Universidade de São Paulo, São Paulo.
- SILVA, D. B. 2010. Transparência na esfera pública interconectada. (Dissertação de Mestrado), Faculdade Cásper Líbero, São Paulo. Disponível em <https://casperlibero.edu.br/wp-content/uploads/2014/04/Transpar%C3%A2ncia-na-Esfera-P%C3%BAblica-Interconectada.pdf>. Acesso em 29 nov. 2019.
- SOLAR, M. et al. 2014. A Model to Guide the Open Government Data Implementation in Public Agencies. Journal of Universal Computer Science, v.20,n.11, p.1564-158.
- UBALDI, B. 2013. Open Government Data: Towards empirical analysis of open government data initiatives. OECD Working Papers on Public Governance, nº 22. Disponível em: <http://dx.doi.org/10.1787/5k46bj4f03s7-en>. Acesso em: 06 mar. 2019.
- VAZ, J. C. 2002. Administração pública e governança eletrônica: possibilidades para a tecnologia da informação. In: Governo eletrônico - os desafios da participação cidadã. Fortaleza: Fundação Konrad Adenauer, Série Debates n. 24.
- W3C. 2014. Best Practices for Publishing Linked Data. World Wide Web Consortium Disponível em: <https://www.w3.org/TR/ld-bp/>. Acessado em 6 dez. 2019.

W3C. 2016. Dados Abertos Governamentais. World Wide Web Consortium Escritório Brasil. Disponível em: <http://www.w3c.br/pub/Materiais/PublicacoesW3C/dados-abertos-governamentais.pdf>. Acessado em 6 abril 2019.

WEISS, Marcos Cesar; BERNARDES, Roberto Carlos; CONSONI, Flavia Luciane. Cidades inteligentes: a aplicação das tecnologias de informação e comunicação para a gestão de centros urbanos. Revista Tecnologia e Sociedade, [s.l.], v. 9, n. 18, p.1-18, 13 dez. 2013. Universidade Tecnológica Federal do Paraná (UTFPR). <http://dx.doi.org/10.3895/rt.s.v9n18>. Disponível em: <https://periodicos.utfpr.edu.br/rt/s/article/view/2634>. Acesso em: 17 set. 2019

WELP, 2007. Y. Democracia y fractura digital en América Latina. In: CONFERENCIA INTERNACIONAL DEMOCRACIA DIRECTA EN AMÉRICA LATINA, 14-15 de marzo de 2007, Buenos Aires, Argentina. Disponível em: www.dd-la.ch/download/Welp_Yanina.pdf . Acesso em: 18. set. 2009.

WIKIPEDIA. Experiência do usuário. Disponível em: https://pt.wikipedia.org/wiki/Experi%C3%Aancia_do_usu%C3%A1rio. Acessado em 18 dez. 2019.

WIKIPEDIA. Linked Data. Disponível em: https://pt.wikipedia.org/wiki/Linked_data. Acessado em 12 dez. 2019.

WIKIPEDIA. Open Knowledge. Disponível em: https://pt.wikipedia.org/wiki/Open_Knowledge. Acessado em 2 dez. 2019.

YANNOUKAKOU, A.; ARAKA, I. 2014. Access to government information: Right to information and open government data synergy. Procedia-Social and Behavioral Sciences, v. 147, p. 332-340.

YILDIZ, Mete. 2007. Decision-Making Models Used in E-Governmente Projects: Evidence from Turkey. In: Handbook of Decision-Making, Marcel Dekker Publications, pp. 395-416 Disponível em: https://www.researchgate.net/publication/313241983_Yildiz_Mete_2007_Decision-Making_in_E-government_Projects_The_Case_of_Turkey_Goktug_Morcol_Ed_Handbook_of_Decision-Making_Marcel_Dekker_Publications_pp_395-416. Acessado em 12 ou. 2019.

14 Anexo 1 – Áreas Temáticas de assunto

Este quadro apresenta as áreas temáticas da Prefeitura Municipal de Belo Horizonte utilizadas nos exemplos e orientação deste texto.

Núm	Nome da área temática	Descrição
1	Administração	Trata das informações de controle dos aspectos patrimoniais, contratos, SCM - suprimentos e logística (transporte, frota, estoque e compras), acervos documentais, solicitações e execução de serviços gerais.
2	Arrecadação Tributária	Reúne informações de lançamento, arrecadação tributária, fiscalização de tributos mobiliários e imobiliários, autônomos, autorização de notas fiscais e livros fiscais, certidões, impostos e taxas, revisão de lançamentos, segundas vias, isenções, cadastros para abertura de comércio, indústrias e outros serviços.
3	Comunicação	Informações de apoio às atividades de comunicação e informações relacionadas ao planejamento, recebimento/expedição, controle, gestão de contratos, de toda comunicação, escrita ou não, da administração municipal.
4	Desenvolvimento humano	Reúne informações relativas à promoção do potencial das pessoas, do aumento de suas possibilidades e o desfrute da liberdade de viver. Trata ainda de informações relacionadas ao desenvolvimento e atividades de ação cultural, programas, projetos e atividades de casas de espetáculos, museus, bibliotecas, arquivos municipais, centros culturais patrocinados pelo Município. Agrupa ainda as informações relacionadas com esporte, recreação e lazer para a população do município, atividades de práticas esportivas, recreativas e de educação física para a população, equipamentos esportivos, produção artesanal local, feiras, exposições e eventos turísticos. Reúne os temas cultura, turismo, lazer e esporte.
5	Educação	Informações de gestão dos dados de alunos, escolas, professores, atividades vinculadas aos processos educacionais e de bolsa-escola, informações relacionadas à política educacional, à formação escolar e profissional, ensino pré-escolar, fundamental, educação de adultos e educação especial.
6	Estratégico	Reúne dados consolidados das funções da PBH que têm potencial

Dados Abertos: transparência governamental
Perspectivas, cenários e planejamento

Núm	Nome da área temática	Descrição
		de facilitar a tomada de decisão dos gestores.
7	Estruturante	Trata de informações de sistemas com características de processo de trabalho da PBH que permeiam vários departamentos/unidades/secretarias e de sistemas que objetivam concentrar/prover cadastros básicos/comuns que interessam à maioria dos setores da PBH.
8	Habitação	Trata de informações alfanuméricas e espaciais sobre a habitação de interesse social, informações relacionadas à política de moradia, coordenação e planejamento das atividades e políticas de habitação, elaboração de diagnósticos ambientais e de áreas de risco, intervenção urbanística em vilas, favelas e conjuntos habitacionais de interesse social, fiscalização para fins de obras e/ou implantação de conjuntos habitacionais de interesse social. Trata também de indicadores sociais, análises espaciais e mapas para o planejamento e o monitoramento da Política Habitacional no município.
9	Inclusão Digital	Reúne informações que objetivam a melhoria dos serviços públicos prestados ao cidadão e a democratização do acesso à informação e à informática.
10	Legislação	Abrange informações que propiciam agilidade e dinamismo aos trabalhos da Câmara Municipal de Belo Horizonte, garantindo o acompanhamento do processo legislativo pelo público interno e externo da Casa.
11	Limpeza Urbana	Trata de informações relativas à programação e controle da execução de limpeza urbana através do registro de rotas, equipamentos e pessoal disponível, volume e tipos de lixo. Engloba também informações sobre tratamento do lixo, informações relacionadas à varrição e coleta de resíduos sólidos e sua destinação final, acompanhamento e controle de coleta de resíduos sólidos, acompanhamento e controle de aterramento de resíduos e obras de terraplanagem e drenagem complementar. Ampliação e manutenção da coleta seletiva e reciclagem de lixo.
12	Meio Ambiente	Trata de informações sobre espécies vegetais, plantas e sementes produzidas, hortas comunitárias, parques e jardins públicos, fontes e índices de poluição.
13	Mobilidade	Reúne informações da gestão e acompanhamento de atividades e

Dados Abertos: transparência governamental
Perspectivas, cenários e planejamento

Núm	Nome da área temática	Descrição
	Urbana	de ocorrências relacionadas ao transporte público, trânsito de veículos e de pedestres, tais como: infrações de trânsito, acidentes, monitoramento de equipamentos de fiscalização, estacionamento rotativo e sinalização, transporte coletivo, táxi, escolar, suplementar, concessão de benefícios e estatísticas de transporte. Trata ainda de informações relacionadas à melhoria da circulação viária, implantação, manutenção, mudança das placas de regulamentação, veículos envolvidos em colisões, multas de trânsito ou por estacionar em lugares proibidos e informação de áreas de circulação restrita.
14	Orçamentário	Reúne informações de gestão orçamentária e financeira, execução orçamentária, financeira e contábil.
15	Planejamento	Trata de informações da coordenação e do planejamento das políticas públicas municipais.
16	Políticas Sociais	Agrupar informações relacionadas à coordenação e execução de projetos, programas e atividades que visam à erradicação da pobreza, ao desenvolvimento social e à garantia dos direitos sociais de portadores de necessidades especiais, dos direitos das crianças, adolescentes e idosos que se encontram em situações de risco pessoal e social, implantação de centros de juventude, inclusão das crianças e adolescentes portadoras de deficiência nas instituições sociais. Abordagem de rua para identificação das crianças e adolescentes em situação de risco pessoal e social, atendimento às crianças e adolescentes com trajetória e vida nas ruas. Agrupa ainda informações relacionadas aos programas de abastecimento, controle de ações de credenciamento, fiscalização e gerenciamento dos equipamentos de abastecimento e programas que integram o sistema municipal de abastecimento.
17	Recursos Humanos	Trata de informações da gestão integrada de Recursos Humanos, informações relacionadas à manutenção do registro dos aprovados em concurso público, coordenação, orientação e controle do programa de estágio de estudantes de nível médio e superior
18	Regulação Urbana	Agrupar as informações relacionadas à política de controle urbano, licenciamento de edificações, análise, aprovação, execução e acompanhamento de projetos de edificações, acompanhamento de obras, alvará de construção, início de construção, baixa e habite-se,

Dados Abertos: transparência governamental
 Perspectivas, cenários e planejamento

Núm	Nome da área temática	Descrição
		licenciamento de atividades econômicas, certidão de mudança de numeração, certidão de baixa de construção e habite-se bem como alvará de construção, de localização e funcionamento. Licenciamento de execução de parcelamento, de ocupação e uso do solo. Licenciamento das atividades de indústria, comércio e serviços no Município. Acompanhamento da execução das obras de parcelamento do solo licenciadas. Fiscalização de controle urbano do município: parcelamento, ocupação e uso do solo e de posturas municipais. Fiscalização de edificações, de comércio, indústria, prestação de serviços e demais atividades urbanas no que se refere às posturas municipais. Identificação e proteção dos bens do patrimônio histórico urbano e manutenção de registros de edificações para a preservação da memória dos mesmos. Coordenação do programa de iluminação pública
19	Saúde	Abrange informações de gestão e acompanhamento das atividades de ambulatório, postos de saúde e atendimento médico, informações relacionadas à saúde a nível ambulatorial e hospitalar, acompanhando e avaliando a prestação de serviços básicos na rede de suporte de especialidades ambulatoriais e diagnósticos e de assistência hospitalar, coordenando ações de vigilância epidemiológica, vigilância sanitária e controle de zoonoses.
20	Segurança Pública	Trata de informações da segurança pública da cidade.
21	Tecnologia e Informação	Trata de informações para apoio e gestão dos serviços, programas e projetos relativos à tecnologia da informação.

15 Anexo 2 – Exemplos de Planos de Ação

Este quadro apresenta exemplos de Planos de Ação que poderiam ser propostos para a abertura de dados da Prefeitura Municipal de Belo Horizonte utilizada nos exemplos e orientação deste texto.

Assunto	Atividades	Meta/prazo	Validade/ expiração	Responsável
Área temática: Educação	Órgão: Secretaria Municipal de Educação			
Conjunto de dados sobre unidades de ensino de todos os níveis da PPBH	1. Estruturar abertura de dados	04/2018	Anual. Segue planejamen to do ano letivo	
	2. Construção dos dados e dicionário de dados	04/2018		
	3. Automação de ambiente para leitura e geração dos arquivos	04/2018		
	4. Publicação dos dados no Portal de Dados Abertos da PBH	06/2018		
Censo de alunos de todos os níveis de ensino da PBH	1. Estruturar abertura de dados	06/2018	Trimestral a partir do início do ano letivo	
	2. Construção dos dados e dicionário de dados	06/2018		
	3. Automação de ambiente para leitura e geração dos arquivos	08/2018		
	4. Publicação dos dados no Portal de Dados Abertos da PBH	09/2018		
Censo de profissionais do magistério da PBH com seu vínculo profissional ⁵² .	1. Estruturar abertura de dados	10/2018	Trimestral a partir do início do ano letivo	
	2. Construção dos dados e dicionário de dados	10/2018		
	3. Automação de ambiente para leitura e geração dos arquivos	11/2018		
	4. Publicação dos dados no Portal de Dados Abertos da PBH	11/2018		
Conjunto de dados sobre movimentação ⁵³	1. Estruturar abertura de dados	10/2018	Anual ao final do período de	
	2. Construção dos dados e dicionário de dados	10/2018		

⁵² O vínculo trata da relação profissional entre o docente e a PBH. Pode ser efetivo, contratado e temporário quanto ao tipo de vínculo e cedido ou em exercício quanto ao local de atividade.

⁵³ O aluno muda de escola por ter sido aprovado e a escola não ter o nível de ensino/ano de

Dados Abertos: transparência governamental
Perspectivas, cenários e planejamento

de alunos na rede municipal	3. Automação de ambiente para leitura e geração dos arquivos	11/2018	matriculas	
	4. Publicação dos dados no Portal de Dados Abertos da PBH	11/2018		
Taxas de distorção ⁵⁴ idade-série escolar na Educação Básica	1. Estruturar abertura de dados		Anual ao final do período de matrículas	
	2. Construção dos dados e dicionário de dados			
	3. Automação de ambiente para leitura e geração dos arquivos			
	4. Publicação dos dados no Portal de Dados Abertos da PBH			
Número médio de Horas-Aula Diária ⁵⁵ na Creche, Pré-Escola, Ensino Fundamental e Ensino Médio.	1. Estruturar abertura de dados		Anual ao final do período de matrículas	
	2. Construção dos dados e dicionário de dados			
	3. Automação de ambiente para leitura e geração dos arquivos			
	4. Publicação dos dados no Portal de Dados Abertos da PBH			
Conjunto de dados sobre Bibliotecas de todos os níveis da PPBH	1. Estruturar abertura de dados		Anual. Segue planejamento do ano letivo	
	2. Construção dos dados e dicionário de dados			
	3. Automação de ambiente para leitura e geração dos arquivos			
	4. Publicação dos dados no Portal de Dados Abertos da PBH			
Censo de	1. Estruturar abertura de dados		Anual.	

escolarização ou por ter alteração de endereço.

⁵⁴ Em um sistema educacional seriado, existe uma adequação teórica entre a série e a idade do aluno. No caso brasileiro, considera-se a idade de 7 anos como a idade adequada para ingresso no ensino fundamental, cuja duração, normalmente, é de 8 anos. Seguindo este raciocínio é possível identificar a idade adequada para cada série. Este indicador permite avaliar o percentual de alunos, em cada série, com idade superior à idade recomendada.

⁵⁵ Este indicador expressa o tempo médio de permanência dos alunos na escola. Trata-se de uma média aritmética ponderada, cujo fator de ponderação é a matrícula na data de referência do Censo Escolar, por série, grupos de séries e nível de ensino.

Dados Abertos: transparência governamental
Perspectivas, cenários e planejamento

profissionais de magistério cedidos, em licença médica e com aposentadoria próxima de todos os níveis da PPBH	2. Construção dos dados e dicionário de dados		Segue planejamen to do ano letivo	
	3. Automação de ambiente para leitura e geração dos arquivos			
	4. Publicação dos dados no Portal de Dados Abertos da PBH			

Área temática: Saúde	Órgão: Secretaria Municipal de Saúde			
Assunto	Atividades	Meta/prazo	Validade/ expiração	Responsável
Conjunto de dados sobre unidades de saúde da Grande BH.	1. Estruturar abertura de dados	06/2018	Anual. Segue ano fiscal	
	2. Construção dos dados e dicionário de dados	06/2018		
	3. Automação de ambiente para leitura e geração dos arquivos	08/2018		
	4. Publicação dos dados no Portal de Dados Abertos da PBH	09/2018		
Censo de profissionais em saúde X local de atuação de todos os níveis da PBH	1. Estruturar abertura de dados	06/2018	Anual. Segue ano fiscal	
	2. Construção dos dados e dicionário de dados	06/2018		
	3. Automação de ambiente para leitura e geração dos arquivos	08/2018		
	4. Publicação dos dados no Portal de Dados Abertos da PBH	09/2018		
Estatística do custo de atendimento por paciente	1. Estruturar abertura de dados		Anual. Segue ano fiscal	
	2. Construção dos dados e dicionário de dados			
	3. Automação de ambiente para leitura e geração dos arquivos			
	4. Publicação dos dados no Portal de Dados Abertos da PBH			
Conjunto de dados sobre demanda por	1. Estruturar abertura de dados		Anual. Segue ano fiscal	
	2. Construção dos dados e dicionário de dados			

Dados Abertos: transparência governamental
Perspectivas, cenários e planejamento

atendimento na rede municipal de saúde	3. Automação de ambiente para leitura e geração dos arquivos			
	4. Publicação dos dados no Portal de Dados Abertos da PBH			
Taxas de mortalidade	1. Estruturar abertura de dados		Anual. Segue ano fiscal	
	2. Construção dos dados e dicionário de dados			
	3. Automação de ambiente para leitura e geração dos arquivos			
	4. Publicação dos dados no Portal de Dados Abertos da PBH			
Taxas de nascidos mortos	1. Estruturar abertura de dados		Anual. Segue ano fiscal	
	2. Construção dos dados e dicionário de dados			
	3. Automação de ambiente para leitura e geração dos arquivos			
	4. Publicação dos dados no Portal de Dados Abertos da PBH			
Conjunto de dados sobre demandas por atendimento nas farmácias populares	1. Estruturar abertura de dados		Anual. Segue ano fiscal	
	2. Construção dos dados e dicionário de dados			
	3. Automação de ambiente para leitura e geração dos arquivos			
	4. Publicação dos dados no Portal de Dados Abertos da PBH			
Censo de ocupação de leito por características de pacientes	1. Estruturar abertura de dados		Anual. Segue planejamento do ano letivo	
	2. Construção dos dados e dicionário de dados			
	3. Automação de ambiente para leitura e geração dos arquivos			
	4. Publicação dos dados no Portal de Dados Abertos da PBH			

16 Anexo 3 – Termo de uso do Portal de Dados Abertos⁵⁶

Este Termo de Uso do Portal de Dados Abertos estabelece condições para que cidadãos possam dar diferentes usos aos recursos disponíveis no Portal de Dados Abertos da Prefeitura de Belo Horizonte (PBH). Os usos incluem, mas não se limitam a: atividades lucrativas, reprodução total ou parcial, tradução, adaptação, reordenação ou qualquer outra modificação, distribuição do original ou do resultado de ressignificação, no limite das restrições legais e da respectiva regulamentação.

O entendimento das condições apresentadas a seguir é, portanto, essencial para que suas interações com o Portal sejam feitas em segurança. Dúvidas sobre algum aspecto podem ser encaminhadas pelo espaço “Contato” do Portal.

A referência geral que rege as condições de uso dos dados publicados no Portal, definida pela PBH, é a autorização de uso dos dados para qualquer finalidade, desde que os mesmos não possuam termos adicionais, conforme tratado a seguir.

Ao conjunto de dados que, após análise da PBH, se mostrar específico quanto às condições de uso serão atribuídas “licenças abertas”, cujas regras são complementos às conceitualmente definidas. Para verificar se determinado conjunto de dados está associado a uma licença aberta, você deve verificar no metadado presente na documentação do arquivo. Quando o conjunto de dados de seu interesse estiver associado a uma licença, você deve observar atentamente os termos da licença e respeitá-los ao realizar qualquer uso.

As licenças empregadas nessas situações foram propostas por organizações internacionais especialmente voltadas ao assunto condições de uso de dados. Foram redigidas em inglês, mas em síntese estabelecem o seguinte:

- **Open Database License (ODbL):** Qualquer pessoa é livre para copiar, distribuir, usar, criar novas publicações a partir da abertura original, modificar, construir ou transformar os dados (ou o banco de dados). Usuários têm que atribuir a fonte, da forma que for definida na licença e garantir que reusos de seus trabalhos com os dados sigam a mesma regra de atribuição. Nos casos de usos dos dados para fins comerciais, deve ser disponibilizada cópia pública. Para acessar o conteúdo integral da licença (em inglês): <https://opendatacommons.org/licenses/odbl/>
- **Creative Commons Zero (CC0):** aplicável a contextos em que o conjunto de dados caracteriza-se como de domínio público. Usada quando a análise do gestor aponta para a inexistência de direito autoral e passa a ser “natural” a renúncia permanente aos direitos de *copyright* e outros relacionados. Nenhuma atribuição à instituição gestora dos dados é esperada por parte dos usuários. Para acessar o conteúdo integral da licença (em inglês): <https://creativecommons.org/share-your-work/public-domain/cc0/>

⁵⁶ Disponível no Portal de Dados Abertos da Prefeitura de Belo Horizonte. Disponível em <https://dados.pbh.gov.br/termo-de-uso.html>. Acessado em 15 jan. 2020.

- **Creative Commons Attribution 4.0:** permite ao usuário (qualquer usuário) realizar livremente atividades de compartilhamento (cópia e redistribuição dos dados em qualquer formato ou mídia), de adaptação (incluindo remixagem, transformação e construção de novos produtos) para quaisquer propósitos inclusive comerciais. Ficam os usuários sujeitos às seguintes condições: (i) atribuir crédito apropriado, fornecendo URL da licença e indicando eventuais mudanças; (ii) Ao atribuir crédito não é necessário seguir formato padrão, mas é exigido que não seja de forma que leve o leitor a supor que a instituição que licenciou o conteúdo endossa o usuário ou os trabalhos produzidos com os dados licenciados. Para acessar o conteúdo integral da licença (em inglês): <https://creativecommons.org/licenses/by/4.0/>

É essencial que os cidadãos interessados em usar dados do Portal estejam cientes que as condições que regem cada conjunto de dados podem ser atualizadas, corrigidas e/ou substituídas a qualquer tempo, a critério da PBH.

Outras condições gerais demandam atenção para uso dos dados disponibilizados no Portal:

- Feito o acesso e/ou manipulação dos dados pelo usuário, o mesmo passa a responder integralmente pelas informações e serviços secundários que oferece, bem como por quaisquer outras atividades empreendidas com base nos dados fornecidos e exime a PBH de qualquer responsabilidade relativa às atividades, informações e serviços por ele desenvolvidos.
- Os usuários dos dados do Portal não devem apresentar os dados obtidos de forma incorreta ou enganosa, nem os utilizar para amparar ou sustentar atividades criminosas ou ilícitas ou de favorecimento ostensivo de atividades de igual caráter.
- A PBH exigirá que o usuário cancele a utilização, distribuição e/ou serviço de todo e qualquer conjunto de dados por ele utilizado se for verificada má-fé; sem prejuízo da adoção das demais providências legais cabíveis.
- Os usuários dos dados do Portal ao gerarem produtos/serviços com os respectivos dados, deverão se responsabilizar pelos conteúdos gerados e prestarem as devidas informações sobre os aplicativos, sistemas, estatísticas ou qualquer informação sobre o que foi produzido.

Importante observar que os gestores dos dados abertos têm responsabilidades limitadas:

- A PBH não garante fornecimento continuado dos dados nem a sua entrega tempestiva aos interessados, salvo nos casos previstos na legislação.
- Em linha com o conceito de Dados Abertos, os dados são disponibilizados “como estão”, e a PBH exime-se de qualquer garantia e obrigação em relação aos dados e à utilização dos mesmos.
- A instituição pública produtora dos dados não será responsável pela utilização feita por pessoas e/ou instituições que transformem e/ou usem direta ou indiretamente os dados e que porventura venham a gerar prejuízos, danos econômicos ou materiais advindos desse uso e/ou transformação dos dados.

Para cumprir os propósitos deste documento, é importante registrar ainda as condições de uso e responsabilidades da Administração para outros conteúdos do Portal:

- Conteúdos disponíveis no Portal Municipal de Dados Abertos de Belo Horizonte incluem *hyperlinks* ou redirecionamento para aplicativos e produtos externos criados e mantidos por terceiros. Essa inclusão é feita para que a navegação seja conveniente para o usuário. Nesses casos, em que é feita a seleção de qualquer *link* ou serviço externo, o usuário estará deixando o Portal e, por extensão, sujeito às políticas de privacidade e uso dos dados dos provedores da página ou serviço referenciados por tal *link*.
- A PBH não patrocina sítios e aplicativos de terceiros disponibilizados neste Portal, de modo que interpretações expressas por produtos e serviços ofertados por terceiros não correspondem, necessariamente, à opinião desta instituição.
- A PBH não garante a autenticidade, relevância, qualidade, atualidade ou completude de dados e informações de terceiros publicadas no Portal, bem como de dados e informações disponibilizados em página ou aplicativos de terceiros, respondendo esses integralmente pelas informações e serviços secundários que oferecerem, bem como por quaisquer outras atividades empreendidas com base nos dados fornecidos, eximindo estas instituições de qualquer responsabilidade relativa a essas atividades, informações, serviços e suas consequências.
- O criador de aplicativo, de software ou outros produtos que utilizem as bases de dados do Portal Municipal de Dados Abertos de Belo Horizonte deverá creditar a devida fonte, bem como informar à PBH de tal fato, pela seção “Contato”, para que o aplicativo seja divulgado.
- Será facultado aos desenvolvedores enviar seu trabalho para avaliação quanto à sua publicação neste Portal, podendo retirá-lo da plataforma a qualquer tempo. O meio para solicitar a publicação de um produto ou serviço, como o descrito acima, deverá ser feito aos mantenedores do Portal por meio da seção “Contato” disponível no Portal.

A validade dos Termos de Uso tem início a partir da data de publicação da Portaria no Diário Oficial do Município, podendo ser alterados, corrigidos ou substituídos a qualquer tempo, visando sempre aprimorar os serviços prestados e assegurar o amplo uso dos dados.

Referência:

Resolução nº 2, de 24 de março de 2017 do Ministério do Planejamento, Desenvolvimento e Gestão, Secretaria de Tecnologia da Informação,
<http://wiki.dados.gov.br/GetFile.aspx?File=%2fComiteGestor%2fResolu%C3%A7%C3%B5es%2fResolucao%20n%C2%BA%2002%20cginda%20-%20assinada.pdf>