

GRUPO I – CLASSE V – Plenário

TC 005.003/2014-8

Natureza(s): Relatório de Levantamento

Órgãos/Entidades: Controladoria -Geral da União; Empresa de Tecnologia e Informações da Previdência Social; Secretaria de Logística e Tecnologia da Informação - MP; Serviço Federal de Processamento de Dados

Interessado: Secretaria de Fiscalização de Tecnologia da Informação - Sefti

Advogado constituído nos autos: não há.

SUMÁRIO: RELATÓRIO DE LEVANTAMENTO. PUBLICAÇÃO DE DADOS ABERTOS NA ADMINISTRAÇÃO PÚBLICA FEDERAL. SOLUÇÕES DE COLETA, ARMAZENAMENTO E PROCESSAMENTO DE GRANDES VOLUMES DE DADOS (BIG DATA). TRANSPARÊNCIA. MODERNIZAÇÃO DA GESTÃO PÚBLICA.

Adoto como relatório o trecho a seguir da manifestação da equipe de auditoria (peça 63):

“1. Introdução

1.1 Deliberação

1. A fiscalização foi autorizada por meio do Despacho de 24/2/2014 do Ministro Benjamin Zymler, no âmbito do TC 033.814/2013-9, que trata de proposta de fiscalização formulada por esta unidade, sendo que as razões que motivaram a sua proposição encontram-se descritas a seguir.

2. O Governo Federal há anos vem realizando iniciativas referentes a governo aberto na Administração Pública Federal, tendo passado a integrar a Parceria para Governo Aberto (Open Government Partnership – OGP) em 2011, instituído a política de dados abertos em 2012, além de ter implementado diversos sítios de dados governamentais abertos.

3. A política de dados abertos possui potenciais diversos, como a melhoria da gestão pública, provimento da transparência, o fomento ao controle social, a promoção da participação social e o fomento à inovação. Por outro lado, há riscos associados, como, por exemplo, o comprometimento tanto do acesso às informações, quanto da segurança das informações disponibilizadas.

4. Verifica-se que há grande volume de dados públicos de interesse coletivo ou geral (dados passíveis de serem abertos) e que novas tecnologias surgem como ferramentas para coleta, armazenamento e processamento de dados em grandes volumes, velocidade e variedade, a exemplo do chamado Big Data, com potencial de geração de conhecimento e insumos para a tomada de decisão com objetivos estratégicos.

5. Por oportuno, o TCU manifestou intenção em realizar seminário internacional com vistas ao aprofundamento dos conhecimentos sobre tecnologias para coleta, armazenamento e processamento de grandes volumes de dados, a exemplo de Big Data, motivado, entre outros, pela deliberação do Plenário do TCU a partir de Comunicação do Ministro Aroldo Cedraz, de 7/8/2013. Em que pese o seminário internacional tenha sido

adiado para 2015, o levantamento é também uma oportunidade de o TCU conhecer melhor o assunto.

6. Considerando o exposto, esta unidade entendeu a importância de conhecer as iniciativas de publicação de dados abertos na Administração Pública Federal, bem como o uso atual de soluções de tecnologias para o tratamento de grandes volumes de dados pelo governo e sociedade como estratégias de transparência e modernização da gestão pública, para que, de posse desse conhecimento, este Tribunal possa planejar e realizar, futuramente, trabalhos de fiscalização sobre o tema.

1.2 Visão geral do objeto

7. Trata-se de fiscalização do tipo levantamento, conforme previsto no art. 238 do Regimento Interno do Tribunal de Contas da União (RI/TCU) e no art. 3º da Portaria-Segecex 15/2011.

1.3 Objetivo

8. O objetivo do trabalho é conhecer iniciativas de publicação de dados abertos na Administração Pública Federal (APF), considerando ainda o uso atual de soluções de coleta, armazenamento e processamento de grandes volumes de dados (Big Data) pelo governo e sociedade como estratégias de transparência e modernização da gestão pública.

1.4 Metodologia utilizada

9. Esta fiscalização foi realizada em consonância com os padrões de levantamento aprovados pela Portaria-Segecex 15/2011.

10. Inicialmente, a equipe de fiscalização buscou estabelecer o escopo e o não escopo do trabalho, bem como a profundidade dos procedimentos de levantamento.

11. Para a definição do escopo, foi considerada, inicialmente, a necessidade de entender o significado do tema dados abertos e sua relação com temas subjacentes, como governo aberto, governo eletrônico e Big Data, além de identificar os potenciais benefícios da utilização de dados governamentais abertos, tais como a possibilidade de servir de insumo para a utilização de soluções de Big Data, estimular novas oportunidades de mercado etc. Também foram abordados potenciais riscos associados à abertura dos dados governamentais.

12. Em paralelo, mediante pesquisa preliminar, a equipe tomou conhecimento de iniciativas de abertura de dados e soluções criadas a partir deles em diversos países e em alguns entes nacionais nas três esferas de governo. Desse modo, optou-se por descrever um panorama internacional e um panorama nacional sobre a disponibilização de dados governamentais abertos e, no caso brasileiro, foi incluído o cenário de utilização de Big Data. Observou-se que muitas das soluções criadas a partir das iniciativas identificadas inicialmente, em geral, relacionavam-se às áreas de educação, saúde, transporte e segurança.

13. Assim, na descrição do panorama nacional, elegeu-se como candidatos a serem entrevistados órgãos responsáveis pela produção e/ou custódia de dados das áreas da educação e da saúde, com vistas a verificar se órgãos responsáveis por informações de tal valor social já se encontram sensibilizados com o tema de dados abertos e as possibilidades de sua aplicação em proveito da gestão pública e da sociedade.

14. Considerando a abrangência do assunto e a opinião de alguns órgãos federais envolvidos com os temas de governo aberto e dados abertos, a equipe definiu o escopo do trabalho.

15. Quanto à profundidade dos procedimentos, foi elaborada matriz de planejamento, conforme o escopo definido, contendo as seguintes questões:

- a) Q1. Quais as características principais de governo aberto, dados abertos e Big Data, e como se relacionam com a transparência e a gestão pública?
- b) Q2. Qual o panorama internacional da disponibilização de dados governamentais em formatos abertos?
- c) Q3. Qual o panorama nacional da disponibilização de dados governamentais e da utilização de Big Data?

16. Nesse processo, também foi definido o não escopo deste trabalho, tendo sido definidos produtos que não seriam entregues por esse levantamento e que também estão relacionados a dados abertos, governo aberto e Big Data, a exemplo de:

- a) descrição de detalhes tecnológicos de soluções de Big Data ou de suporte à abertura de dados;
- b) descrição de eventuais projetos de Big Data com base em dados sigilosos e cujos resultados não poderiam ser publicados diretamente;
- c) proposição de métrica para avaliar o grau de abertura de dados dos entes da Administração Pública Federal;

17. Ressalta-se, ainda, que a pesquisa realizada para levantar soluções internacionais e nacionais que utilizam dados abertos não foi exaustiva, e que os exemplos trazidos neste relatório são aqueles que se destacaram dentre as soluções das quais a equipe teve conhecimento.

18. Durante a execução, a equipe visitou os seguintes órgãos/entidades: Departamento de Governo Eletrônico da Secretaria de Logística e Tecnologia da Informação do Ministério do Planejamento, Orçamento e Gestão (DGE/SLTI/MP), Secretaria de Transparência e Prevenção da Corrupção da Controladoria-Geral da União (STPC/CGU), Departamento de Participação Social da Secretaria Nacional de Articulação Social da Secretaria-Geral da Presidência da República (DPS/SNAS/SG/PR), Laboratório Hacker da Câmara dos Deputados, Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep), Departamento de Informática do SUS (Datusus), Dataprev e Serpro. Para conhecer a percepção da sociedade, além da visão dos gestores, foi realizada uma reunião com Alexandre Gomes, que foi representante da organização da sociedade civil Transparência Hacker e é estudioso e defensor do tema.

19. Registra-se que as considerações manifestadas por cada ente fiscalizado foram incorporadas ao tópico de descrição do panorama nacional no presente relatório de levantamento.

20. Cabe mencionar que os sítios da internet referenciados neste relatório foram acessados nos meses de junho e julho de 2014.

1.5 Limitações

21. Não houve fatores que limitassem a definição do escopo e da profundidade de análise considerados ideais para este trabalho.

1.6 Processos conexos

22. Não há.

2. Características Principais de Governo Aberto, Dados Aberto e Big Data, e como se relacionam com a transparência e a gestão pública

23. Este capítulo pretende apresentar algumas definições a partir de publicações e estudos acadêmicos sobre dados abertos, governo aberto, governo eletrônico, Big Data e

suas relações. Ademais, abordam-se os potenciais benefícios de dados abertos, a exemplo da possibilidade de sua utilização como instrumento de participação social e como fonte para a aplicação de soluções de Big Data. Por fim, destaca-se o registro de benefícios obtidos por países em decorrência da abertura de dados, bem como a enumeração de alguns riscos potenciais inerentes à tal.

2.1 Governo Aberto

24. Consoante a Parceria para Governo Aberto (Open Government Partnership – OGP), a expressão governo aberto refere-se a projetos e ações que visam à promoção da transparência, à luta contra a corrupção, ao incremento da participação social e ao desenvolvimento de novas tecnologias, de modo a tornar os governos mais abertos, responsáveis por suas ações e aptos a atender às necessidades dos cidadãos (opengovpartnership.org).

25. De acordo com a Organização para Cooperação e Desenvolvimento Econômico (OCDE), para a construção de um governo aberto existem três princípios-chave a serem levados em consideração:

a) Responsabilização e Prestação de Contas (Accountability) – É necessário que existam mecanismos que possibilitem aos cidadãos fiscalizarem o governo quanto ao desempenho de suas políticas e de seus serviços;

b) Transparência – Disponibilização de informações confiáveis, relevantes e tempestivas sobre as atividades do governo, de forma que os cidadãos compreendam as ações realizadas por ele;

c) Participação Social – O governo deve escutar os cidadãos e levar em consideração os seus anseios, tanto no desenho quanto na implementação das políticas públicas e dos serviços públicos prestados.

26. O Brasil é reconhecido como um protagonista no cenário internacional no que diz respeito a governo aberto, tendo sido convidado para compor, juntamente com outros países, a OGP, à qual foi o primeiro a presidir, juntamente com os Estados Unidos. Ao término da fase de execução deste levantamento, o Brasil compunha o comitê diretor da OGP e a parceria já contava com 64 países signatários. A OGP 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.

27. No lançamento da Parceria, os oito países fundadores (África do Sul, Brasil, Estados Unidos, Filipinas, Indonésia, México, Noruega e Reino Unido) assinaram a Declaração de Governo Aberto (peça 35), em que reconheceram a reivindicação da sociedade por mais abertura nos governos e se comprometeram com os princípios e com os objetivos de um governo aberto.

28. Segundo a Declaração de Governo Aberto (peça 35, p 1-3), para um governo ser considerado aberto, ele deve se comprometer a:

a) aumentar a disponibilidade de informações sobre atividades governamentais;

b) apoiar a participação social;

c) implementar os mais altos padrões de integridade profissional na Administração;

d) aumentar o acesso a novas tecnologias que promovam a transparência e a prestação de contas.

29. A OGP define que um compromisso por um governo aberto deve seguir quatro princípios, sendo três deles também declarados pela OCDE – transparência, participação

cidadã e responsabilização e prestação de contas – além do princípio tecnologia e inovação, por meio do qual o governo deve reconhecer a importância das novas tecnologias no fomento à inovação, provendo acesso à tecnologia e ampliando a capacidade da sociedade de utilizá-la (cgu.gov.br/governoaberto/a-ogp/o_que_e_Governo_Aberto.asp).

30. *Além disso, os compromissos da OGP devem ser estruturados em torno de um conjunto de cinco áreas de atuação denominadas Grandes Desafios pela Parceria. Por outro lado, a OGP reconhece que os países encontram-se em diferentes contextos e que, assim, cada país membro deve selecionar os desafios e compromissos mais relacionados à sua realidade (cgu.gov.br/governoaberto/a-ogp/como_Funciona.asp).*

31. *Os cinco grandes desafios propostos pela OGP são:*

- a) melhoria dos serviços públicos;*
- b) aumento da integridade pública;*
- c) gestão mais efetiva dos recursos públicos;*
- d) criação de comunidades mais seguras;*
- e) aumento da responsabilidade corporativa.*

32. *Para a OGP, a melhoria da qualidade da governança do país, bem como dos serviços públicos prestados aos cidadãos, são a base para a sustentabilidade de um governo mais transparente, mais receptivo à participação popular e mais disposto à prestação de contas e à responsabilização (opengovpartnership.org/about/mission-and-goals#sthash.uiuyIhXz.dpuf).*

33. *O governo aberto também é definido como uma doutrina de governança que estabelece o direito dos cidadãos a acessar documentos e procedimentos do governo, de forma a permitir supervisão por parte da sociedade (en.wikipedia.org/wiki/Open_government).*

34. *Como pode ser observado, os conceitos de governo aberto e transparência estão fortemente correlacionados. Governo aberto e, conseqüentemente, transparência, são frequentemente considerados indutores da responsabilização e da prestação de contas (accountability) e do combate à corrupção, à ineficiência e ao desperdício.*

35. *Defensores do governo aberto costumam frisar que não se obtém a transparência somente com a edição de normas legais, mas sim com a participação de toda a sociedade civil, por meio de uma imprensa livre, de organizações não governamentais e da atuação de instâncias fiscalizadoras dentro do próprio governo.*

36. *Nesse sentido, cabe citar que associações não governamentais tais como a Transparência Internacional (transparency.org), a Open Society Foundations (opensocietyfoundations.org) e a Open Knowledge Foundation (okfn.org) se tornaram referências mundiais que influenciam os debates e rumos adotados em diversos países no que diz respeito a governo aberto.*

37. *O comprometimento dos países perante a OGP e suas diretrizes se materializa na forma de um plano de ação nacional, contendo as ações concretas e mensuráveis para o fortalecimento da transparência e para a construção de um governo aberto no âmbito de cada país membro. Seguindo os preceitos da OGP, esse plano deve ser elaborado com a participação da sociedade civil.*

38. *À época dos trabalhos de campo, estava em vigor o 2º Plano de Ação Brasileiro, com 52 compromissos relacionados à temática do governo aberto (peça 60).*

39. *A evolução relativa à execução das ações estabelecidas no plano de ação é acompanhada por meio do Mecanismo de Avaliação Independente (Independent Report Mechanism) estabelecido pela OGP. Segundo esse mecanismo, cada país participante da OGP deve publicar seus relatórios independentes de avaliação bianuais contendo o estágio atual de cada ação, bem como recomendações técnicas quanto à implementação do plano. A avaliação do 2º Plano de Ação Brasileiro foi realizada e está disponível na página da CGU (cgu.gov.br/governoaberto/no_brasil/plano-brasileiro_segundo-plano/index.asp).*

2.2 Relação entre Governo Eletrônico e Governo Aberto

40. *O conceito de governo aberto muitas vezes aparece associado ao de governo eletrônico (e-gov), embora se refiram a objetivos diferentes. Observa-se que algumas publicações tentam delimitar o conceito de governo eletrônico.*

41. *A OCDE, no artigo intitulado “The e-Government Imperative”, define governo eletrônico como “o uso de tecnologias da informação e comunicação, e em particular a internet, como uma ferramenta para alcançar um governo melhor” (OCDE, 2003, p. 23), disponível em oecd-ilibrary.org/governance/oecd-e-government-studies_19901054.*

42. *A Organização das Nações Unidas (ONU) e a American Society for Public Administration (Aspa) realizaram, em 2002, um estudo em que definem governo eletrônico como “a utilização da internet e da world-wide-web para a entrega de informações e serviços do governo para cidadãos”, disponível em unpan1.un.org/intradoc/groups/public/documents/un/unpan021547.pdf.*

43. *Atualmente, as Nações Unidas (ONU) definem governo eletrônico como “a utilização das TIC e sua aplicação, por parte do governo, para a prestação de informações e serviços públicos à população”. A ONU classifica governo eletrônico em três grupos: Governo para Governo (G2G), que envolve o compartilhamento de dados entre atores de governos em todos os níveis federativos; Governo para Negócios (G2B), que envolve transações específicas, como pagamentos e compras públicas; e Governo para Cidadãos (G2C), que são as iniciativas voltadas para facilitar a interação entre as pessoas e o governo, o que inclui a entrega de serviços públicos e a participação do cidadão no processo de tomada de decisão, disponível em unpan1.un.org/intradoc/groups/public/documents/un/unpan021547.pdf, acessado em 4/6/2014.*

44. *Embora sejam conceitos distintos e que têm objetivos diferentes, existe uma clara sobreposição entre governo aberto e governo eletrônico no que diz respeito à transparência. Portanto, entende-se que governo aberto pode envolver ações de governo eletrônico, sendo, no entanto, mais abrangente, visto que inclui aspectos relacionados à responsabilização e ao incremento da participação social.*

45. *A Administração Pública Brasileira disponibiliza, hoje, inúmeros serviços de governo eletrônico, em diferentes estágios de maturidade, sendo que alguns deles já são amplamente utilizados pelos cidadãos, como, por exemplo, o serviço de declaração de imposto de renda e o serviço de consulta ao Cadastro de Pessoa Física (CPF) oferecidos pela Receita Federal do Brasil. Outros exemplos de ações de governo eletrônico no Brasil são os portais instituídos com o objetivo de prover acesso à população a informações sobre serviços públicos (servicos.gov.br) e a aplicativos para prestação desses serviços (aplicativos.gov.br) oferecidos pelo Governo Federal.*

2.3 Dados Abertos

46. Com o advento das tecnologias digitais, a prática de governo aberto no que diz respeito à informação e à transparência governamental passou gradativamente a se apropriar dos recursos de processamento e disseminação de dados eletrônicos em grande volume e alta velocidade. Atualmente, a transparência está fortemente associada à disponibilização dos chamados dados abertos, por meio dos quais é possível, a princípio, escrutinar o universo de ações governamentais de forma mais detalhada.

47. Segundo a definição da Fundação do Conhecimento Aberto (Open Knowledge Foundation – OKF), “dados são abertos quando qualquer pessoa pode livremente usá-los, reutilizá-los e redistribuí-los, estando sujeito a, no máximo, a exigência de creditar a sua autoria e compartilhar pela mesma licença” (dados.gov.br/dados-abertos e versão original da OKF em okfn.org/opendata). Nesse caso, uma licença refere-se ao regramento de como os dados podem ser compartilhados.

48. A OKF é uma organização sem fins lucrativos, criada no Reino Unido em 2004, que busca promover o chamado “conhecimento aberto”, ou seja, qualquer conteúdo, informação ou dado que as pessoas possam usar livremente, reusar ou redistribuir sem qualquer restrição legal, tecnológica ou social (princípios do conhecimento aberto podem ser vistos em opendefinition.org/od). No capítulo seguinte, serão apresentados mais detalhes sobre a atuação da OKF.

49. O Portal Brasileiro de Dados Abertos (dados.gov.br), um ponto focal na disponibilização de dados abertos no Brasil, usa a definição de dados abertos proposta pela OKF. O portal toma como referência, ainda da OKF, as chamadas três leis (regras) e os oito princípios que devem reger os dados abertos.

50. As três leis foram propostas pelo especialista em políticas públicas e ativista dos dados abertos David Eaves e, embora tenham sido apresentadas para dados abertos governamentais, pode-se dizer que elas se aplicam aos dados abertos de forma geral. As três leis são:

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.

51. Em 2007, um grupo de trabalho da OKF definiu os oito princípios dos dados abertos governamentais:

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 detenha 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.

52. Na gênese dessas leis (regras) e princípios, existem razões e justificativas fortemente ligadas à realidade do mundo digital e aos princípios da própria transparência.

53. O mundo digital é fortemente conectado e a busca automática é um dos seus princípios. Os mecanismos de busca automática – dentre os quais o da Google (google.com) é atualmente o mais popular – servem como indexadores do conteúdo existente na internet. Soluções como essas não exigem, por exemplo, que se conheça o endereço de rede no qual um conteúdo é disponibilizado. Os indexadores permitem, por meio de palavras-chave ou frações de conteúdo, obter o endereço e então acessar o conteúdo integral procurado.

54. A primeira lei dos dados abertos proposta pela OKF reflete a realidade da internet. Um dado que não pode ser indexado não existe. Para que exista, o dado aberto precisa então ser disponibilizado em um local de rede livremente acessível pelos indexadores e pelos internautas interessados em acessá-lo. Os princípios 4 e 6 também refletem essa regra.

55. Outro pilar do mundo digital é o tratamento automatizado dos dados por meio de ferramentas que permitam a análise de grandes quantidades de registros sem intervenção manual. Mas, para que essa análise seja possível, é necessário que: a) o dado disponibilizado o seja já em formato digital, sendo possível replicá-lo em bloco no computador onde a análise ocorrerá; b) o dado respeite princípios básicos de estrutura tabular (organizados de modo que se possa associar diferentes itens e atribuir-lhes um valor, seja número ou texto); e c) o dado não esteja codificado em formato que exija transformação prévia, que por sua vez requeira ferramentas proprietárias (pagas ou de acesso restrito). Essas regras estão refletidas na segunda e terceira leis e nos princípios 5, 7 e 8 da OKF.

56. Com relação à transparência, é de senso comum que prover transparência a algo supõe informar, e que informar supõe fornecer informações completas e atuais. Daí derivam os princípios 1 e 3.

57. Quanto ao princípio 2, tem-se que dados primários, também chamados de dados brutos, derivam tipicamente de sistemas informatizados baseados em bancos de dados. Os bancos de dados armazenam as transações geradas durante o uso de um sistema informatizado. Uma transação está normalmente associada a um fato atômico, a exemplo da inclusão ou exclusão de um dado, a qual é uma operação que não pode ser subdividida, e por isso resulta no registro do dado de mais fina granularidade. Trazendo para um campo de aplicação, tal como a educação, podemos citar como exemplo de dado primário o registro da escolaridade de um aluno do ensino básico.

58. Em contraposição aos dados primários, utiliza-se a expressão “dados analíticos ou informacionais”, que seriam aqueles derivados dos dados primários por meio do uso de regras e ferramentas que transformam o dado primário em informações agregadas (grossa granularidade) e/ou traduzem o dado para formatos, codificações ou vocabulários mais facilmente compreensíveis por certo público alvo. Esses são tipicamente aqueles dados apresentados em sistemas de informações gerenciais e de apoio à decisão. Na área de educação, um exemplo de dado agregado seria o grau de escolaridade de todo os alunos do ensino básico em um determinado Estado.

59. *O que o princípio 2 busca atingir é a transparência absoluta, na qual todos os fatos informados em um sistema estejam acessíveis. A partir desses dados primários, o consumidor da informação poderá ele mesmo criar suas próprias análises ou agregações, de forma a tirar suas próprias conclusões, sem ser direcionado ou influenciado por algum intermediário que, ao aplicar regras de análise, agregação ou tradução, tenha criado algum viés ou interpretação equivocados.*

60. *As leis, princípios e conceitos propostos pela OKF foram adotados por diversos países, entre eles o Reino Unido, os EUA e o próprio Brasil. Entretanto, não é possível evidenciar uma padronização a nível mundial no que diz respeito aos dados abertos. Um levantamento feito pela Global Open Data Initiative (Godi; globalopendatainitiative.org/survey-and-interview-report) indica que, embora os padrões da OKF sejam os mais citados, diversos países até ignoram a existência desses padrões ou não consideram que a padronização seja relevante. O mesmo levantamento mostra bastante diversidade de implementações das políticas de dados abertos governamentais. Em muitos países, por exemplo, dados são obtidos por meio de demandas baseadas em leis de liberdade de informação ou “raspadas” de sítios governamentais para depois serem publicadas por organizações da sociedade civil ou institutos de estatística. Raspagem de dados, do inglês *data scraping*, é um jargão utilizado por desenvolvedores de aplicativos para a internet, significando a atividade de extrair dados de sítios e convertê-los a um formato legível para programas de computador, para que possam ser analisados e cruzados com mais facilidade.*

61. *Em relação ao caso brasileiro, não se pode deixar de citar interpretação dos dispositivos da Lei 12.527/2011, Lei de Acesso à Informação (LAI), acerca do conceito de dados abertos governamentais. Consoante o art. 8º da LAI, dados abertos são as informações de interesse coletivo ou geral que devem ser obrigatoriamente divulgadas pelos órgãos e entidades públicos em seus sítios oficiais e que devem atender, entre outros, aos seguintes requisitos: possibilitar o acesso automatizado por sistemas externos em formatos abertos; serem estruturados e legíveis por máquina; estarem acompanhados de detalhes sobre os formatos utilizados para estruturação da informação; serem autênticos, íntegros e atualizados.*

62. *Outros princípios que merecem ser citados são os adotados pelo G8, o grupo das oito maiores economias do mundo, na Declaração de Dados Abertos do G8 (G8 Open Data Charter; international.gc.ca/g8/open_data-donnees_ouvertes.aspx?lang=eng), os passos estabelecidos pelo governo Obama na chamada Open Government Directive (White House 2009, whitehouse.gov/open/documents/open-government-directive) para a criação de um governo aberto, e as recomendações do comitê de políticas públicas do capítulo americano da tradicional Association of Computing Machinery – USACM (acm.org/public-policy/open-government). Observa-se que, a despeito da especificidade ou ênfase adotada por cada uma dessas instituições, em síntese, todas elas harmonizam com os princípios de David Eaves, ao pregar que os dados: são públicos por princípio; devem ser abertos; e devem ter qualidade, utilidade, inteligibilidade e reusabilidade, no intuito de propiciar a inovação e a melhoria da gestão pública por meio da criação e estabelecimento de uma cultura de dados abertos.*

2.4 Big Data

63. *Inicialmente associada às pesquisas científicas em meteorologia, genômica e simulações físicas, que lidam com grandes quantidades de dados (nature.com/nature/journal/v455/n7209/edsumm/e080904-01.html), a expressão Big Data passou a ser associada à internet, com o processamento de dados associados ao*

comportamento dos usuários em mídias sociais, como no caso do Facebook ou em pesquisas da Google.

64. *Considerando-se que a expressão é recente e que existem inúmeras definições (opentracker.net/article/definitions-big-data), considera-se que Big Data ainda é um termo vago e com definição imprecisa.*

65. *De forma geral, há definições para Big Data que ressaltam aspectos mais técnicos, enquanto há outras que buscam definir o termo sem o emprego de referências técnicas, tais como a proposta por Victor Mayer Schönberger e Kenneth Cukier, no livro intitulado Big Data: A Revolution that Will Transform How We Live, Work and Think, que define Big Data como:*

Big Data se refere a análises de dados em larga escala que não são possíveis quando em baixa escala, para extrair novas ideias ou criar novas formas de valor, que alteram mercados, organizações, o relacionamento entre os cidadãos e os governos, e muito mais (tradução livre).

66. *A seguir, tem-se duas definições que consideram aspectos técnicos.*

67. *Segundo o Gartner, Big Data é a expressão empregada para descrever ativos de informação em grande volume de dados, velocidade de atualização ou consulta e variedade de formatos, que demandam formas de processamento da informação próprias para essas características, para melhor compreensão e tomada de decisão, de um modo que não estaria disponível em uma menor escala do conjunto de dados (gartner.com/it-glossary/big-data).*

68. *Desse modo, a expressão Big Data é frequentemente utilizada em sentido lato para designar quaisquer grandes bases de dados, cuja análise escapa aos métodos tradicionais em banco de dados relacionais e em estatística descritiva ou que requeiram grande capacidade computacional distribuída em numerosas máquinas para ser executada. A estatística descritiva é a parte que procura descrever e avaliar características de uma amostra da população sem generalizar conclusões ou inferências sobre a população como um todo.*

69. *Assim, tem-se as seguintes características tecnológicas associadas a soluções de Big Data:*

69.1. *complexidade do conjunto dos dados: dados em diversos formatos e pouco estruturados que não se adequam ao modelo de armazenamento, manipulação e recuperação de dados adotado em bancos de dados relacionais;*

69.2. *métodos próprios de análise: métodos baseados em estatística indutiva (a parte da estatística que, baseando-se em resultados obtidos da análise de uma amostra da população, procura inferir, induzir ou estimar as regras de comportamento da população como um todo) e em aprendizado automático (machine learning – o desenvolvimento de algoritmos e técnicas que permitam ao computador aperfeiçoar seu desempenho em alguma tarefa). Essa análise busca identificar padrões nos dados e correlações entre suas variáveis de forma a construir modelos que permitam realizar previsões mais confiáveis e, em última instância, que guiem ações mais eficientes;*

69.3. *uso de recursos computacionais distribuídos: devido às dimensões do conjunto de dados e à complexidade das análises estatísticas, seu processamento é tipicamente realizado de forma distribuída, sob numerosas máquinas independentes que atuam de forma coordenada e regidas por softwares próprios, tais como MapReduce (research.google.com/archive/mapreduce.html) e Hadoop (hadoop.apache.org).*

70. *Exemplo bem sucedido e conhecido de uso de Big Data, o serviço de tradução automática da Google (translate.google.com) consegue bons resultados utilizando*

correlações estatísticas entre textos que já foram previamente traduzidos por tradutores humanos, obtidos a partir de imensas bases de documentos escritos em várias línguas e de publicações da Organização das Nações Unidas (ONU).

71. *Por oportuno, cita-se que o governo dos EUA anunciou, em 2012, uma iniciativa de Big Data (whitehouse.gov/blog/2012/03/29/big-data-big-deal) que agrupa diversos projetos associados à tecnologia em áreas tão diversas quanto defesa, energia, meio ambiente, saúde, espaço e espionagem (whitehouse.gov/sites/default/files/microsites/ostp/big_data_fact_sheet_final_1.pdf). Essa iniciativa levou a um relatório de atividades, de maio de 2014, que aponta conclusões da experiência americana (whitehouse.gov/sites/default/files/docs/big_data_privacy_report_may_1_2014.pdf), das quais se destacam: o caráter inevitável da adoção de Big Data no governo daquele país; riscos relacionados à proteção da privacidade; a relevância dos dados abertos governamentais que podem servir de insumos para projetos de Big Data; e a necessidade de investimento de recursos para tirar o máximo proveito da tecnologia em benefício da população.*

72. *Diversas publicações propõem benefícios potenciais alcançáveis com o uso de Big Data. Dentre esses benefícios, incluem-se:*

a) personalização das respostas geradas por sistemas de informação, como nas recomendações de compra da Amazon (amazon.com), que sugerem automaticamente novos produtos em função de comparações de históricos de compras de sua base de clientes;

b) captação de informações dispersas em redes de sensores ou por meio de contribuições individuais voluntárias, como a iniciativa da cidade de Boston, nos Estados Unidos, que lançou um aplicativo para celulares que detecta a vibração causada pela passagem de um carro sobre um buraco na estrada, registra a posição geográfica em que isso ocorreu com o GPS e transmite esses dados pela internet para os servidores do departamento de estradas do município, que pode então alocar seus recursos em resposta à incidência de buracos e seu impacto sobre os motoristas que usam o aplicativo (cityofboston.gov/doi/apps/streetbump.asp).

c) detecção de fraudes: controle aduaneiro do trânsito de mercadorias (fcw.com/Articles/2013/09/25/big-data-transform-government.aspx?p=1);

d) resposta a incidentes e emergências: antecipação de problemas de trânsito; priorização de resposta a catástrofes (fcw.com/Articles/2013/09/25/big-data-transform-government.aspx?p=1);

e) democratização do acesso à informação: pelo uso de soluções de Big Data em conjunto com soluções baseadas em plataformas abertas, em que dados abertos podem suportar novas aplicações de governo eletrônico desenvolvidas por quaisquer pessoas. Como exemplo, citam-se as avaliações de qualidade e disponibilidade de hospitais e escolas, o que permitiria maior empoderamento da população e eventuais ganhos em participação social (fcw.com/Articles/2013/09/25/big-data-transform-government.aspx?p=1);

f) ganhos de eficiência e produtividade com a otimização do processo produtivo e redução de desperdícios (fcw.com/Articles/2013/09/25/big-data-transform-government.aspx?p=1);

g) novas oportunidades de negócio: como, por exemplo, o desenvolvimento do setor de mapeamento, georreferenciamento e rastreamento, após a abertura de bancos de imagens de satélites e acesso à rede GPS.

73. *Identifica-se um consenso de que entre os principais desafios para se obter os benefícios esperados com o uso do Big Data destacam-se questões organizacionais, como acesso aos dados brutos, melhora no uso de métodos estatísticos empregados na análise dos dados (ft.com/cms/s/2/21a6e7d8-b479-11e3-a09a-00144feabdc0.html#axzz2yyxPpCGB) e riscos à privacidade das pessoas (whitehouse.gov/sites/default/files/docs/big_data_privacy_report_may_1_2014.pdf), mais do que questões propriamente tecnológicas, em que há rápidas inovações e contínuo aperfeiçoamento de soluções.*

2.5 Benefícios potenciais de dados abertos

74. *As práticas de governo aberto, quando promovidas pelo governo, podem gerar benefícios tanto para a sociedade como para o próprio governo. Ao levar conhecimento e entendimento sobre a execução de suas políticas, ações, serviços e desempenho, o governo se permite fiscalizar, expondo junto à sociedade sucessos e insucessos.*

75. *Em artigo intitulado *Effective Open Government: Improving Public Access to Government Information*, publicado em 2005, a OCDE já registrava que a transparência e a abertura de dados figuram como aspectos essenciais de um governo moderno (papers.ssrn.com/sol3/papers.cfm?abstract_id=2069870), destacando como benefícios a participação democrática e compreensão do cidadão, a melhoria do processo de tomada de decisão, do gerenciamento dos dados governamentais, da eficiência da gestão, dos serviços oferecidos aos cidadãos, além da redução da corrupção.*

76. *A literatura lista outros potenciais ganhos decorrentes da abertura das informações públicas. A OKF (okfn.org/opendata) ressalta a participação do cidadão na gestão do governo, uma vez que a disponibilização dos dados governamentais não somente leva informação aos cidadãos, mas também a possibilidade de que eles se envolvam no processo de decisão e contribuam com as ações públicas. É o empoderamento do cidadão que passa ao papel de agente da transformação social por meio do acompanhamento e da fiscalização das ações governamentais, das políticas públicas, enfim do desempenho do governo. Por outro lado, com mais poder e conhecimento, o cidadão pode passar a ter mais confiança no poder público e a se sentir participante desse poder.*

77. *O aplicativo Política Aberta (politicaaberta.org), por exemplo, apresenta a relação das empresas maiores doadoras de campanhas eleitorais de 2012 e a relação das empresas mais contratadas do governo federal no mesmo ano, fornecendo informação para o exercício do controle social e acompanhamento do processo eleitoral (parágrafo 222). Outro exemplo é a solução QEdu, que constitui um portal aberto e gratuito com informações sobre a qualidade do aprendizado dos alunos do 5º e 9º anos em cada escola, município e estado do Brasil, permitindo o acompanhamento da evolução de determinadas políticas públicas educacionais (parágrafo 228).*

78. *Ao disponibilizar seu dados, o governo também favorece a criação de negócios e serviços inovadores com valor tanto comercial como social. Organizações, cidadãos, acadêmicos, empresas e até mesmo as instituições públicas têm a possibilidade de utilizar bases de dados públicos para a produção e o compartilhamento de novos conhecimentos e de novos serviços, numa concepção de coparticipação entre ente privado e governo na oferta de serviços públicos. Uma vez que o Estado não tem capacidade de prover a quantidade e a diversidade de serviços que a sociedade pode e quer consumir, a conjunção de esforços se mostra oportuna.*

79. *Em outro artigo, intitulado *Do More Transparent Governments Govern Better?* e publicado em 2003, o Banco Mundial ressalta que o acesso à informação pode, ainda,*

induzir a melhoria do gerenciamento de riscos, do desempenho econômico do governo e de sua eficiência burocrática (www.oecd.org/gov/46560184.pdf).

80. *Nota-se, portanto, a diversidade de potenciais benefícios oriundos de práticas de abertura de dados públicos, sendo que alguns deles já estão sendo mensurados, como é o caso dos exemplos a seguir.*

81. *A OKF publicou, em novembro de 2013, uma apresentação em que mostra os ganhos obtidos por países em decorrência da abertura de dados (peça 59, p. 56-65). Por exemplo, no Canadá, dados abertos expuseram um esquema de fraudes relacionadas a isenção de impostos sobre doações a instituições de caridade, que permitiu a contestação de US\$ 32.000.000.000,00 em impostos devidos. Na Espanha, um estudo realizado aponta que cerca de cinco mil empregos estão relacionados a empresas que criam aplicativos, produtos e serviços, utilizando informações do setor público. Em Uganda, a política de acesso à informação e a decorrente divulgação dos recursos públicos recebidos por cada escola contribuiu para a redução da corrupção que se apoderava desses valores, e as escolas, que antes se apropriavam de apenas 20% do valor a elas destinado, passaram a dispor de 90% desse valor. Ao divulgar a localização de trezentos mil pontos de ônibus, a comunidade do Reino Unido retificou dezoito mil deles, contribuindo para a melhoria da qualidade das informações públicas.*

2.6 Oportunidades de utilização de dados abertos em conjunto com Big Data

82. *Dados abertos governamentais frequentemente atingem dimensões que poderiam qualificar a utilização de soluções de Big Data, dado o volume elevado de informações, a velocidade em que os dados são atualizados e a variedade dos dados. Para exemplificar, tomam-se as informações acerca das compras governamentais no âmbito do governo federal que são volumosas e diversificadas, constituindo fontes de conhecimentos variados, profundos e inovadores, passíveis de serem extraídos com a utilização de ferramentas de Big Data, para diversos fins, tais como detecção de fraudes. A fiscalização realizada por esta secretaria, que analisou as informações registradas no sistema Comprasnet, especificamente quanto ao cumprimento do disposto na Lei 10.520/2002, art. 7º (TC 015.239/2012-8), foi realizada utilizando-se ferramentas de auditoria de dados (Audit Command Language). Para esse objeto, portanto, seria adequado o uso de ferramentas de Big Data, que poderiam trazer mais e mais diversas inferências sobre a relação entre gestores, empresas e licitações.*

83. *Além disso, dados abertos governamentais fornecem informações básicas que podem ser necessárias e úteis para projetos de construção de aplicações a partir dos preceitos do Big Data, como informações de georreferenciamento, bases de códigos postais, informações climáticas, dados históricos ou dados de levantamentos estatísticos (como os do IBGE).*

84. *Adicionalmente, repositórios de dados públicos que se tornam abertos podem estimular novos projetos de Big Data específicos para aproveitar seu conteúdo, como exemplifica o sistema de mapas interativos da Google (maps.google.com), construído, inicialmente, sobre a base de dados espaciais do instituto do censo dos EUA (theatlantic.com/technology/archive/2012/09/how-google-builds-its-maps-and-what-it-means-for-the-future-of-everything/261913).*

85. *Entende-se que as tecnologias de Big Data podem e devem fazer uso de dados abertos no âmbito governamental, seja para informar o cidadão, gerar conhecimento, promover o controle social ou para qualquer outro fim que favoreça o progresso da gestão Estado e dos serviços que por ele são prestados.*

2.7 Oportunidades de utilização de dados abertos com participação social

86. *A participação social é um dos pilares de governo aberto, estimulando que o cidadão conheça, verifique, fiscalize, se esclareça e acompanhe as ações públicas de seu interesse. Por outro lado, com as vias de comunicação estabelecidas, o cidadão pode, também, fornecer informações ao Estado, contribuindo para a avaliação e ajustes das políticas públicas e para a melhoria dos dados públicos, entre outros benefícios.*

87. *Essa relação entre Estado e sociedade já se faz presente em soluções desenvolvidas com base em dados abertos. A Open Knowledge Foundation (okfn.org) relaciona alguns exemplos em que a participação social propiciou a melhoria da qualidade dos dados (blog.okfn.org/category/open-government-data/page/5/). O Reino Unido, por exemplo, ao disponibilizar os dados referentes às trezentas mil paradas de ônibus, tornou possível a contribuição da comunidade Open Street Map (openstreetmap.org), que atuou na correção das informações de dezoito mil paradas de ônibus, colaborando para a melhoria das informações do transporte público daquele país (peça 59, p. 61).*

88. *No entanto, para que a sociedade realmente participe, a simples publicação dos dados governamentais não é suficiente. É necessário que a disponibilidade das informações públicas, com fácil acesso aos cidadãos, seja divulgada e tenha despertado o interesse para o seu uso, tanto para a criação de novos produtos e serviços, como para a construção de novas e diferentes visões das informações.*

89. *Para fomentar a participação social com o uso de dados abertos, parece necessário que seja criada uma cultura de dados abertos (mckinsey.com/insights/public_sector/how_government_can_promote_open_data). Além de levar ao público o conhecimento do acesso às suas informações, o governo deve promover eventos que estimulem o uso dessas informações pela sociedade civil (universidades, empresas, organizações etc.). Várias estratégias têm sido utilizadas para instigar a utilização de dados abertos em dezenas de países, incluindo Reino Unido, Estados Unidos, Austrália, Canadá, Nigéria, Singapura e Brasil. Uma delas é a realização de encontros presenciais, tais como hackathon, desconferência, barcamp, meetup, speedgeek, lightning talk, etc, que, grosso modo, permitem aos gestores públicos conhecer quem são os prováveis usuários dos dados públicos, quais suas demandas e como eles querem reutilizar os dados produzidos ou custodiados pelo Estado. Da mesma forma, esses encontros permitem aos membros da sociedade conhecer os dados já disponibilizados e potenciais dados a serem publicados.*

90. *O modelo de competições em torno de dados abertos tem sido bastante utilizado, em geral com o nome de Hackathon ou Hack Day. Nesse evento, um conjunto de dados públicos é disponibilizado aos desenvolvedores de aplicativos reunidos num mesmo local durante um período de tempo, de 48 horas a até algumas semanas, para que eles construam uma solução (software) utilizando os dados objeto do concurso. Ao término do tempo definido, eleger-se e premia-se a melhor solução desenvolvida. No Brasil, Hackathons já foram promovidos pela Câmara dos Deputados, Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep), MJ, além de vários órgãos das esferas estaduais e mesmo por entidades não governamentais como a Fiesp. Um dos produtos oriundos do Hackathon promovido pelo Inep foi a solução “Escola que Queremos” (parágrafo 226).*

91. *Apesar das ideias e protótipos de soluções inovadoras que surgem nesses eventos, muitos dos produtos gerados não encontraram um ecossistema que os fomente e mantenha como produto de mercado, ficando desatualizados e caindo em desuso. Uma*

questão que se discute atualmente nas comunidades de dados abertos é a sustentabilidade das soluções produzidas nesses eventos. É aconselhável que exista um ecossistema, nos quais os seus elementos (sociedade, universidades, estudantes, empresas e governo) se estruturam para prover e permitir continuidade aos produtos e ideias gerados que tenham sido considerados potencialmente úteis para agregar valor para o cidadão, disseminando-o na sociedade.

92. *Se por um lado aborda-se muito a temática de participação popular e acompanhamento para o envolvimento da população na política do país, tendo a tecnologia como valioso instrumento viabilizador das diversas formas de participação hoje existentes, por outro lado, pouca parece ser a evolução no sentido de mensurar como, se, e em que intensidade essa participação está efetivamente ocorrendo (peça 42, p. 13). Na busca realizada pela equipe de fiscalização, pouco se obteve de informação. Diante disso, cita-se que a OKF tem procurado coletar casos que mostrem os resultados da participação da sociedade na melhoria da qualidade das informações (blog.okfn.org/category/open-government-data/page/5/) e da gestão pública (peça 59, p. 56-64).*

2.8 Principais riscos inerentes à abertura dos dados

93. *A convergência e a ampla adoção dos conceitos e princípios relativos a dados abertos pelos governos das nações desenvolvidas demonstram sua solidez e um razoável consenso em torno da sua utilidade. Juntamente a esses conceitos e princípios, a necessidade de transparência governamental e a relação umbilical entre governo aberto, transparência e dados abertos são também consenso na maioria dos países estudados.*

94. *Entretanto, a disponibilização dos dados abertos por parte dos governos não deve ser considerada um fim em si mesma. Diversos profissionais e pesquisadores vêm alertando para uma certa “ilusão” de transparência que os dados abertos podem criar, além de jogar luz a uma série de questões adjacentes ao tema que não estão sendo discutidas na profundidade requerida.*

95. *Propõe-se, neste trabalho, expor um conjunto de riscos inerentes à adoção dos dados abertos, divididos em três grandes grupos:*

- a) custos da disponibilização continuada dos dados;*
- b) qualidade, utilidade, e usabilidade dos dados;*
- c) privacidade e proteção dos dados pessoais.*

2.8.1 Custos da disponibilização continuada dos dados

96. *Rob Kitchin, um especialista irlandês sobre dados abertos, no seu artigo *Four critiques of open data initiatives* (nuim.ie/progcity/2013/11/four-critiques-of-open-data-initiatives), destaca os custos e a sustentabilidade como um risco considerável à implantação das políticas de dados abertos. Segundo os princípios mais largamente aceitos a respeito de dados abertos, os custos da geração e disponibilização desses dados devem ser arcados pelo fornecedor, que nesse caso é o governo. Para justificar esse gasto público, o principal argumento econômico seria o de que o dado aberto estimula a inovação e, em consequência, a economia, abrindo novas possibilidades de negócios.*

97. *A questão é que, segundo o autor, esse retorno é dificilmente mensurável, imprevisível e mal distribuído. Enquanto algumas bases de dados públicas podem ser de grande valia, como as que disponibilizam mapas ou dados sobre transporte, outras tantas têm pouco ou nenhum valor econômico. Considerando-se que o dado aberto só é útil se for amplo, atual e detalhado, estima-se que os custos continuados em manter enormes*

quantidades de dados disponíveis pode vir a ser um peso considerável para a administração pública.

2.8.2 Qualidade, utilidade e usabilidade dos dados

98. Em um artigo publicado no sítio da W3C, o consórcio da World Wide Web, sob o título *A Realistic Look at Open Data* (w3.org/2012/06/pmod/pmod2012_submission_38.pdf), a pesquisadora Sharon S. Dawes, da Universidade de Albany, Nova York, faz uma análise crítica muito pertinente dos dados abertos e levanta uma série de questões importantes. O artigo começa destacando que, ironicamente, a informação é geralmente tratada como uma “caixa-preta” pelo movimento de dados abertos. Muita atenção é dada aos atores, ferramentas tecnológicas, e técnicas de análise. Mas a informação disponibilizada pelos governos é vista como entregue, utilizada sem espírito crítico e com o rótulo de confiável, mesmo antes de ser examinada. Contudo, como a autora alerta, os dados disponibilizados sob a bandeira dos dados abertos provenientes de sistemas administrativos foram criados ou coletados para outros propósitos, diferentes de sua divulgação ampla. Apesar de terem um inegável valor potencial, não se pode desconsiderar alguns riscos substanciais com relação à validade, relevância e confiança dos dados abertos.

99. Como ponto de partida na análise dos riscos inerentes aos dados abertos, a autora busca, ainda, diferenciar as fontes governamentais de dados tipicamente voltados à análise externa, ou seja projetadas para uso informativo, como é o caso das informações dos institutos de pesquisa, dos censos e das séries estatísticas, daquelas que disponibilizam dados administrativos, provenientes da automatização das atividades governamentais, voltadas apenas para o seu registro.

100. Por terem finalidade informativa, os dados voltados à análise externa são de mais fácil interpretação se comparados aos dados administrativos. Frisa-se, todavia, que ambos devem incorporar o fornecimento de informações voltadas especificamente à inteligibilidade dos dados (metadados), de forma a evitar entendimentos equivocados por parte da sociedade.

101. Outro trabalho interessante foi publicado pela Association of Government Accountants (AGA), associação americana de profissionais financeiros no governo, intitulado *What Do Citizens Want From Public Sector Reporting? Findings From Four Focus Groups* (peça 61). O artigo inicia destacando duas “escolas de pensamento” em transparência. A primeira pensa dados no seu formato bruto e em forma granular, disponibilizados na web, de forma que os usuários possam criar seus próprios relatórios. Eles acreditam que dados disponibilizados de forma agregada contém o viés da entidade governamental que realizou as agregações. A segunda escola acredita que os dados deveriam ser agregados e apresentados de forma amigável e compreensível, além de serem distribuídos por diversos meios e mídias, inclusive a internet. Esse grupo acredita que o “dado” não organizado não é genuinamente “informação”. Para a AGA, os dois caminhos são válidos e deveriam existir esforços nos dois sentidos por parte do governo: o de fornecer dados primários e também dados analíticos.

102. O mesmo artigo apresenta os achados de grupos de trabalho que contaram com a participação dos membros da associação, diretamente ligados à difusão e uso de dados governamentais. Desses achados, destacamos:

a) entidades estão publicando informações orçamentárias aos cidadãos, outras estão colocando detalhes de pagamentos na internet, e com isso dizem estar sendo transparentes. Entretanto, a informação não está sendo lida ou utilizada pelos cidadãos

e/ou pagadores de impostos, porque é muito difícil de ser assimilada e compreendida em seu contexto;

b) as pessoas estariam mais interessadas em relatórios financeiros se elas pudessem entender como essas informações estão associadas aos resultados da entidade;

c) antes que haja progresso (no que diz respeito ao uso dos dados abertos) é preciso haver uma mudança cultural.

103. *O levantamento sobre dados abertos realizado pela Global Open Data Initiative (Godi), já citado anteriormente, reforça o primeiro desses achados acima. A elevada complexidade técnica das bases de dados disponibilizadas, que inibe seus uso e sua exploração, foi citada como um dos principais desafios às políticas de dados abertos. Esse mesmo levantamento destacou também problemas de qualidade dos dados disponibilizados, dados incompletos e desatualizados como desafios a serem vencidos.*

104. *Outra questão que merece reflexão com relação à utilidade e à usabilidade dos dados abertos diz respeito ao tratamento a posteriori dos dados abertos, uma vez disponibilizados pelo governo. Andrea Di Maio, do Gartner Group, em seu artigo *Open Data and the New Divide* (blogs.gartner.com/andrea_dimaio/2012/05/14/open-data-and-the-new-divide) discorre sobre o “dilúvio dos dados abertos”, e o risco de “afogamento” do usuário devido ao volume avassalador dos dados brutos disponíveis ou simplesmente devido a sua completa falta de sentido quando apresentados sem a preparação adequada. A questão altamente relevante que Di Maio coloca é: quem deve processar esses dados para que eles façam sentido? O próprio governo? Mas isso contrariaria o princípio 2 proposto pela OKF e suas razões subjacentes (parágrafos 57 a 59). Seria, então, a sociedade organizada, por meio dos seus ativistas, militantes, hackers e suas associações não governamentais? Di Maio pondera que, por mais bem intencionados que a maioria desses atores sejam, o risco de interpretações errôneas, casuísticas e/ou direcionadas ao convencimento de certo público com relação a certos pontos de vista ainda existe. No que o autor chama de “lado negro dos dados abertos”, o cenário que se desenha após uma primeira “primavera da transparência” pode ser o de uma forte divisão entre quem tem recursos e habilidades para manipular os dados abertos e quem não tem, e essa divisão pode ser ainda mais insidiosa que a divisão entre quem tem, e quem não, acesso às tecnologias em geral e, em particular, à internet.*

105. *No mesmo sentido, a crítica já citada de Rob Kitchin (parágrafo 2.8.1) destaca o potencial efeito dos dados abertos no sentido de municiar os poderosos: embora a intenção intrínseca à transparência seja o fortalecimento da democracia e o acesso igual e indiscriminado aos dados públicos, a força de grupos e empresas, já preparadas e dotadas de recursos tecnológicos para a análise e exploração de grandes quantidades de dados, faz com que haja necessariamente diferenças com relação ao potencial de aproveitamento e obtenção de resultados a partir dos dados abertos, favorecendo certos segmentos.*

2.8.3 Privacidade e proteção dos dados pessoais

106. *Da mesma forma que os dados abertos vêm sendo impulsionados pelos ativistas do mundo digital, a luta pela proteção dos direitos dos cidadãos – referentes à privacidade e contra o uso não autorizado ou danoso de informações pessoais mantidas pelo Estado – também começa a despontar entre comunidades transnacionais na internet.*

107. *A questão da privacidade e do uso danoso de dados pessoais na internet é um tema recorrente e amplamente debatido, principalmente a partir do advento das redes sociais, onde pessoas tendem a se expor em demasia e acabam sofrendo com o uso não autorizado ou não previsto de informações que lhes dizem respeito. O termo “direitos digitais” (pt.wikipedia.org/wiki/Direitos_digitais) refere-se aos direitos humanos*

associados às tecnologias digitais, e está particularmente relacionado à proteção e realização dos direitos existentes, tais como o direito à privacidade ou à liberdade de expressão.

108. *A privacidade e a proteção dos dados pessoais foi destacada pela ONU, na declaração de princípios da Sociedade da Informação resultante do World Summit on the Information Society (WSIS; itu.int/dms_pub/itu-s/md/03/wsis/doc/S03-WSIS-DOC-0004!!PDF-E.pdf), em 2003, entre os princípios chave de uma sociedade da informação para todos. Diversos grupos militam em torno desse tema. Um deles, o Open Rights Group (ORG; openrightsgroup.org/campaigns/opensource/open-data-privacy), aborda especificamente o tema da privacidade em dados abertos. A ORG apoia os dados abertos, mas afirma que privacidade é primordial. Para eles, os governos devem admitir e tratar os riscos de “reidentificação” de cunho pessoal nos dados abertos. A reidentificação é o processo de identificação de dados pessoais a partir de informações ditas “anonimizadas” (smartplanet.com/blog/pure-genius/privacy-reidentification-a-growing-risk). As bases de dados com informações anonimizadas contém, normalmente, dados pessoais, mas buscam impedir a identificação das pessoas ali referenciadas ao excluir ou mascarar, total ou parcialmente, atributos chave tais como nome, número de telefone ou outros documentos de identificação dos indivíduos. A reidentificação é feita através da chamada “triangulação de informações”, onde diversas bases, anonimizadas ou não, são pesquisadas até que se encontre coincidências tão restritivas que acabem isolando as informações de um indivíduo em meio aos demais.*

109. *Para a ORG, deve haver um debate público apropriado antes de serem lançadas iniciativas amplas de disponibilização de bases de dados pessoais anonimizadas pelos governos, assim como uma análise criteriosa das tecnologias de anonimização utilizadas.*

110. *Kieron O’Hara, da Universidade Southampton – Reino Unido, em seu relatório Transparent Government, Not Transparent Citizens: A Report on Privacy and Transparency for the Cabinet Office, preparado por encomenda do gabinete do Primeiro Ministro Britânico, chegou às seguintes conclusões com relação à divulgação de dados pessoais anonimizados:*

- a) privacidade é extremamente importante para a transparência. A legitimidade política de um programa de transparência dependerá crucialmente de sua habilidade em reter a confiança do público. Privacidade deveria, então, ser parte integrante de qualquer programa de transparência e não somente uma preocupação subsequente;*
- b) privacidade e transparência são compatíveis, desde que essa última seja cuidadosamente protegida e considerada a cada estágio;*
- c) discussões sobre anonimização têm sido conduzidas largamente por demandas legais, com um conseqüente negligenciamento das considerações feitas pela comunidade técnica;*
- d) inexistência de consenso técnico ou legal para a questão da anonimização. Devemos continuar a anonimizar dados sensíveis, mas nos manter cautelosos em disponibilizá-los livremente enquanto continuamos a avançar no sentido de administrar e pesquisar os riscos da anonimização;*
- e) a transparência deve ser debatida com transparência – metatransparência.*

111. *O mesmo relatório faz, ainda, uma série de recomendações, entre as quais destacam-se:*

- a) incluir a proteção da privacidade já na definição das iniciativas de abertura de dados governamentais, e não após sua instalação;*

- b) quando a disponibilização de uma base de dados possa vir a ameaçar a privacidade das pessoas ali retratadas, e caso seja possível criar mecanismos de controle de acesso, então esses controles devem ser considerados;*
- c) incluir o paradigma técnico na discussão da transparência;*
- d) atuar em direção a uma transparência sob demanda: disponibilizar preferencialmente bases para as quais existe demanda comprovada, deixando aos próprios setores demandantes a demonstração de que ela não trará nenhuma violação de privacidade;*
- e) criar um catálogo de bases de dados governamentais, descrevendo, entre outras características, o nível de confidencialidade e as restrições de acesso das informações nelas armazenadas. Mesmo bases de dados não disponibilizadas devem constar no catálogo, incluindo a razão da sua não disponibilização. Dessa forma, setores da sociedade poderão solicitar mais facilmente bases de seu interesse e até questionar as razões da não disponibilização de algumas informações;*
- f) criar painéis setoriais para discutir a transparência sobre questões relacionadas à disponibilização de bases e privacidade;*
- g) definir procedimentos de análise prévia de dados a serem abertos para garantir o respeito à privacidade;*
- h) avançar nas pesquisas no tema da privacidade e reidentificação.*

112. *Uma questão interessante a respeito de dados pessoais foi levantada pela OKF no texto intitulado Open Data & My Data (blog.okfn.org/2013/02/22/open-data-my-data), que vai além da questão da privacidade. Segundo o artigo, é bastante provável que cada cidadão seja referenciado em inúmeras bases de dados, algumas delas governamentais e outras mantidas por empresas privadas. Mas é muito pouco provável que um cidadão tenha conhecimento sobre os dados que lhe dizem respeito ou em que bases ele é referenciado, com que propósito e como essa base está sendo usada. Em um cenário onde cada cidadão pudesse questionar o governo e obter os dados por ele custodiados que lhe dizem respeito, a relação entre dados pessoais e dados abertos tomaria outra dimensão. O cidadão poderia não só ser agente de controle de sua privacidade, mas, em alguns casos, escolher autorizar e compartilhar certos dados pessoais a seu respeito, se desejasse. Um paciente de câncer, por exemplo, poderia voluntariamente abrir seus dados médicos se isso pudesse contribuir para pesquisas.*

113. *As políticas públicas que tratam da abertura de dados podem se beneficiar da experiência obtida diante das dificuldades enfrentadas por outros países, de forma que sua implementação considere as principais armadilhas associadas a essas iniciativas.*

114. *Por outro lado, considerando que programas de dados abertos tendem a contribuir para que o Estado atue com mais transparência (princípio básico da governança corporativa) e que, em certa medida, ainda estimula a participação e colaboração da sociedade na melhoria da gestão pública, entende-se que a execução de uma política de abertura de dados, considerando os riscos relacionados, deve ser objeto de futuros trabalhos do TCU sobre o tema.*

3. Panorama Internacional de Disponibilização de Dados Governamentais em Formatos abertos

115. *Esse capítulo tem o objetivo de apresentar o cenário da política de abertura dados governamentais no âmbito internacional, trazendo algumas das iniciativas que os países têm realizado e o avanço por eles obtido. São relatadas aqui, também, as avaliações publicadas por organizações sobre o nível de abertura de dados dos países.*

116. Durante a etapa de execução da fiscalização, foram coletados exemplos de soluções construídas sobre dados governamentais, sendo listadas nesse capítulo as mais interessantes na avaliação da equipe.

117. Cabe mencionar o documento do qual a equipe de fiscalização teve conhecimento, que consiste no primeiro relatório de tendências de dados abertos realizado pelo Tribunal de Contas da Holanda, em março de 2014. O relatório, à semelhança desta fiscalização, abordou conceitos de dados abertos, políticas e legislação correlata, além do cenário nacional (peça 43) e evidencia a preocupação em relação à avaliação dos controles relacionados ao tema, da qual este Tribunal compartilha.

3.1 Países em que se observam iniciativas relevantes de abertura de dados

118. Há anos organizações de vários países se empenham para compartilhar princípios e práticas de governo aberto e disseminá-los entre os governos. Apesar disso, o assunto dados abertos ainda constitui um domínio em evolução, sendo inclusive novidade para muitos, tanto no exterior como no Brasil. É crescente o interesse em tornar os dados governamentais abertos e difundir o seu uso. Espalhadas pelo mundo, essas organizações pregam que, dentre outros efeitos, o acesso à informação empodera o cidadão e lhe dá liberdade, conhecimento e possibilidade de participação nas ações governamentais. Open Government Partnership (OGP), Open Data Institute (ODI), W3C, W3C Brasil, Open Knowledge Foundation (OKF), OKF Brasil, Open Institute, Fundar, Sunlight Foundation e Global Open Data Initiative são algumas das organizações hoje atuantes nesse tema.

119. O governo britânico foi o pioneiro na publicação de dados abertos com o objetivo inicial de disponibilizar a maior quantidade de dados possível, criar uma pequena infraestrutura e fazer alguns experimentos (w3c.br/Noticias/MissaoTecnologicaSobreDadosAbertos), e, desde então, se mantém na liderança em relação a práticas de governo aberto (parágrafo 126). Todavia, em que pese a constatação de que ainda há muito a evoluir, os esforços para a consolidação da cultura de dados abertos permanece na agenda do governo, como relatou Francis Maude, Ministro do Governo Britânico:

The UK is a world-leader in opening up data because we know that it creates a more accountable, efficient and effective government. Open Data is a raw material for economic growth, supporting the creation of new markets, business and jobs and helping us compete in the global race. To ensure this agenda continues to thrive, we are supporting a number of projects which will drive forward this culture of openness.

120. Hoje, dezenas de países já adotam a abertura como política pública e abrem seus dados a cidadãos e empreendedores.

121. Sessenta e três (63) países integram hoje a OGP, desenvolvendo ações para a implementação de um governo aberto. As ações são estabelecidas pelos próprios países nos seus planos de ação (parágrafo 37), de acordo com sua realidade. As estratégias adotadas variam de país para país, o que, dentre outros fatores, torna desafiadora a comparação da evolução das ações entre os países.

122. Com o intuito de estimular a inovação e ampliar a transparência e a responsabilização pelos gastos públicos, nações e organizações ao redor do mundo estão buscando conhecer como os dados abertos agregam valor. Para isso, precisam medir o nível de abertura dos países, bem como o impacto social e econômico das práticas associadas. A medição do nível de abertura, contudo, é complexa e difícil, ante a prematuridade do tema. Muitas definições ainda estão por serem estabelecidas e padronizadas; as estratégias e políticas relacionadas a dados governamentais abertos ainda encontram-se em evolução; a gestão desses dados é, muitas vezes, delegada a

governos locais ou regionais; isso tudo dificulta a definição de indicadores próprios a essa medição.

123. Um dos indicadores utilizados é o número de conjuntos de dados (*datasets*) disponibilizados pelo país. Entretanto, a simples comparação desse número pode induzir a conclusões equivocadas, pois o essencial é a informação que está sendo disponibilizada (seu conteúdo) e seu potencial impacto. Se o cidadão tem acesso a centenas de *datasets*, mas esses dados não trazem a informação que ele deseja, estão desatualizados ou tampouco são capazes de produzir os benefícios esperados (item 2.5 do presente relatório), considerá-los em uma métrica é desaconselhável (peça 40, p. 15). Além disso, há outras variáveis que diferenciam conjuntos de dados disponibilizados, como, por exemplo: a utilidade dos dados; o modelo de publicação adotado, isto é, se é utilizado um portal centralizado, de mais fácil acesso, ou vários portais regionais; e a adequação do *dataset* aos requisitos estabelecidos no conceito de dados abertos, entre outras.

124. Portanto, o número total de *datasets* pode ser considerado na avaliação do grau de abertura de um país ou estado, desde que conjuntamente com outros indicadores, e, na medida do possível, estabelecendo-se critérios para que o *dataset* possa ser considerado.

125. Metodologias e parâmetros de medição do nível de abertura dos países têm sido desenvolvidos e testados por organizações. Em pesquisa, a equipe de auditoria teve conhecimento de algumas análises já realizadas, descritas a seguir.

Open Knowledge Foundation (OKF)

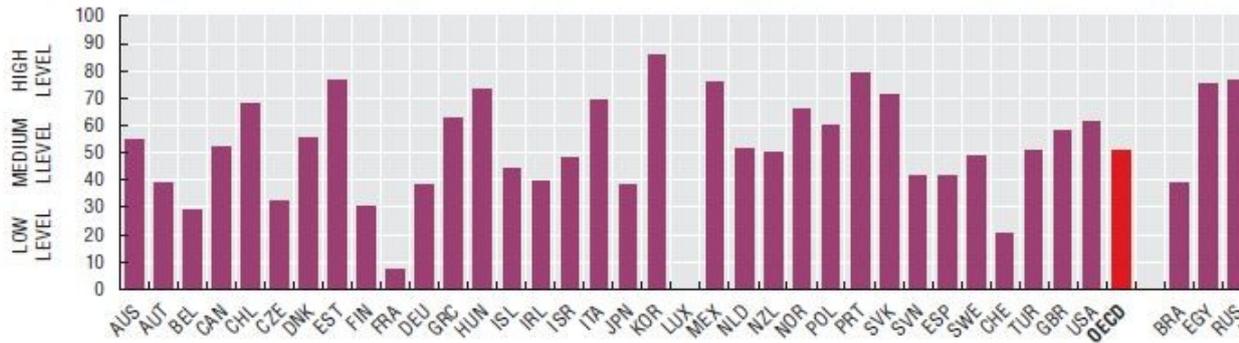
126. A OKF criou um *Open Data Census*, do qual participam, voluntariamente, setenta países, com informações sobre as suas bases de dados disponibilizadas. Com essas informações, publicou, em 2013, um índice que mede a disponibilização de dados abertos: o *Open Data Index* (index.okfn.org/country). O índice é calculado a partir de dez temas governamentais: tabelas de horário do transporte, orçamento governamental, gastos governamentais, resultados de eleições, registros de empresas, mapa nacional, estatísticas nacionais, legislação, códigos postais e emissão de poluentes. No âmbito de cada tema, é avaliada a existência dos dados e seus atributos (disponível, atualizado, formato digital, acessível por máquina, divulgado, completo, bruto, gratuito e licenciado). No ranking obtido por meio desse índice, o Brasil aparece em trigésimo sexto lugar com 410 pontos de um total de mil pontos (index.okfn.org/country/overview/Brazil). No topo desse ranking, está o Reino Unido com 940 pontos.

Organisation for Economic Co-operation and Development (OECD)

127. A OECD (Organização para Cooperação e Desenvolvimento Econômico) também tem buscado desenvolver indicadores para avaliar a abertura dos governos, de forma que sejam complementares aos indicadores já utilizados para avaliar os mecanismos legais e institucionais na intenção de promover a transparência, a acessibilidade e a capacidade de resposta dos governos. Esses indicadores são utilizados na análise de comparação internacional sobre o desempenho do setor público, divulgados no documento *Government at a Glance*, e apresentados no capítulo *Open and Inclusive Government* (peça 41). A avaliação da OECD é realizada a partir de informações de seus 34 países membros (peça 41, p. 13 e 14). Alguns dos indicadores apresentados na avaliação abarcam, também, informações dos países que não são membros da OECD, mas participam do seu Comitê de Governança Pública (Brasil, Egito, África do Sul e Ucrânia). Para esses quatro países, as informações são apresentadas separadamente, ao final das tabelas e figuras.

128. O capítulo *Open and Inclusive Government* aborda quatro aspectos: dados governamentais abertos; conflito de interesse e divulgação de recursos; transparência orçamentária; e participação no processo legislativo, classificando a abertura de dados em três níveis: baixo, médio e alto. O Brasil figura no ranking somente quanto ao aspecto de divulgação dos bens e recursos dos agentes públicos, encontrando-se em posição intermediária em relação aos demais países.

8.4. Asset disclosure: Level of disclosure of private interests and public availability of information (2013)



Source: 2012 OECD Survey on Managing Conflict of Interest.

StatLink  <http://dx.doi.org/10.1787/88893294>

Figura 1 – OECD – *Government at a Glance* – Nível de divulgação das informações referentes ao patrimônio e recursos dos agentes públicos (peça 41, p. 142 e 143)

129. Em termos de dados abertos governamentais, a avaliação de 2013, com dados de 2012, apresenta o Canadá, a Holanda, Eslovênia e os Estados Unidos como os países da OECD com o maior número de datasets disponibilizados no seu portal centralizado de dados abertos governamentais, excluindo-se da avaliação os países que não possuíam, à época, um portal centralizado (peça 41, p. 140 e 141). Nessa avaliação, realizada em 2013, o Canadá tinha 188.921 datasets. No dia 1º/8/2014, momento em que foi consultado o Portal Brasileiro de Dados Abertos (parágrafo 282), o Brasil tinha 239 conjuntos de dados.

World Wide Web Foundation

130. A World Wide Web Foundation criou o Web Index (thewebindex.org) para avaliar a contribuição da internet para o desenvolvimento e o fortalecimento dos direitos do cidadão em 81 países, com indicadores que abrangem quatro dimensões: Acesso Universal, Liberdade e Nível de Abertura, Conteúdo Relevante e Empoderamento.

131. A classificação apresentada no Web Index de 2013 (peça 42, p. 6) traz, nas cinco primeiras posições, os seguintes países: Suécia, Noruega, Reino Unido, Estados Unidos e Nova Zelândia. Na dimensão Liberdade e Nível de Abertura, o Brasil ocupa a posição 41, sendo que os cinco primeiros países foram: Noruega, Finlândia, Islândia, Holanda e França. Na classificação geral, de 81 países, o Brasil figura na posição 33. Dentre os países emergentes abrangidos pela pesquisa, o México figura em primeiro lugar, seguido da Colômbia, Brasil, Costa Rica e África do Sul.

132. O documento aponta que os governos avaliados ainda não se desenvolveram convenientemente quanto à disponibilização de dados governamentais brutos. Nos países pesquisados, menos de dez por cento dos datasets considerados prioritários (key datasets) são disponibilizados online em formato aberto, permitindo sua reutilização (peça 42, p. 4 e 14).

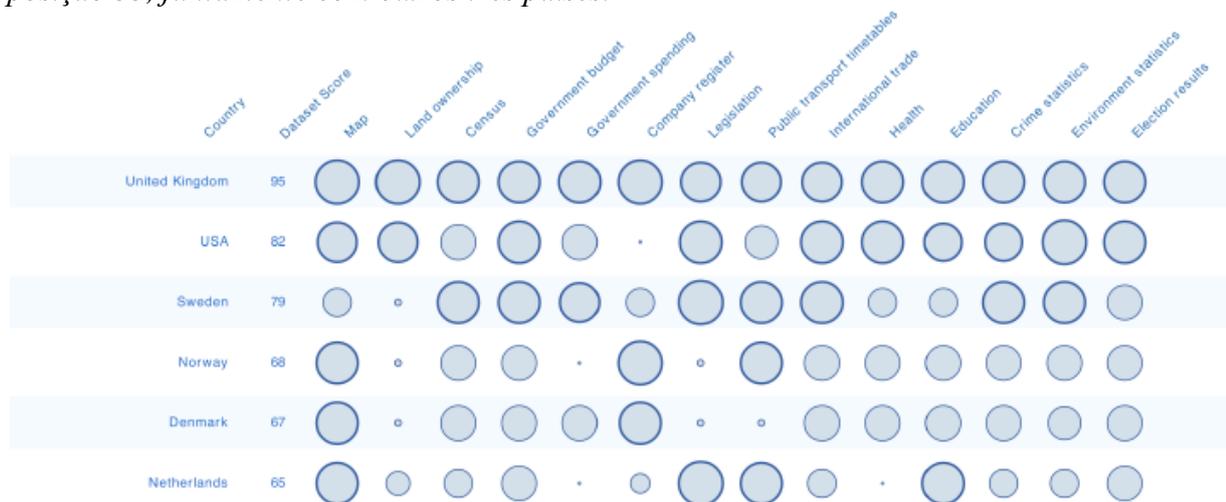
Open Data Barometer (ODB)

133. O Open Data Barometer (opendatabarometer.org) é um projeto do Open Data Institute e da World Wide Web Foundation que busca desenvolver métodos de avaliação para dados abertos. O objetivo do projeto é desvendar o impacto das iniciativas de dados abertos, lançando um olhar para a disseminação da política governamental de dados abertos e sua prática em todo o mundo. O projeto realiza a análise das tendências globais e classifica os países por meio de uma metodologia que pondera a prontidão para garantir os benefícios de dados abertos, os níveis reais de implementação e o impacto das iniciativas de dados abertos.

134. Em 2013 foi realizado um estudo abrangendo 77 países e publicado no relatório Open Data Barometer Global Report (peça 40). Esse trabalho apontou o Reino Unido como o país mais avançado em termos de dados abertos governamentais, seguido dos Estados Unidos, Suécia, Nova Zelândia, Dinamarca e Noruega, estando os dois últimos empatados em quinto lugar.

135. Nessa análise, houve a preocupação para que os dados abertos fossem avaliados não apenas quantitativamente, mas buscou-se ponderar sua relevância, formato, localidade e todos os demais atributos que compõem o conceito de dados abertos (parágrafos 50 e 51). No total, foram coletadas informações de 1.078 conjuntos de dados de catorze categorias selecionadas como prioritárias (peça 42, p. 15-17).

136. Para a análise específica dos conjuntos de dados (datasets) realizada pelo estudo, foram selecionadas catorze categorias: mapa, registro de imóveis, censo, orçamento, gastos, registro de empresas, legislação, transporte público, comércio internacional, saúde, educação, estatística de crimes, estatísticas ambientais e resultados de eleições. Os resultados são representados por um círculo, cujo tamanho representa quatro níveis de abertura, ou seja, quanto maior o círculo, maior a abertura dos dados. Se o dataset enquadra-se na definição de dado aberto, o círculo apresenta uma borda mais grossa. Abaixo, são trazidas as análises dos datasets de alguns países. O Brasil ocupa a posição 33, juntamente com outros três países.



(...)

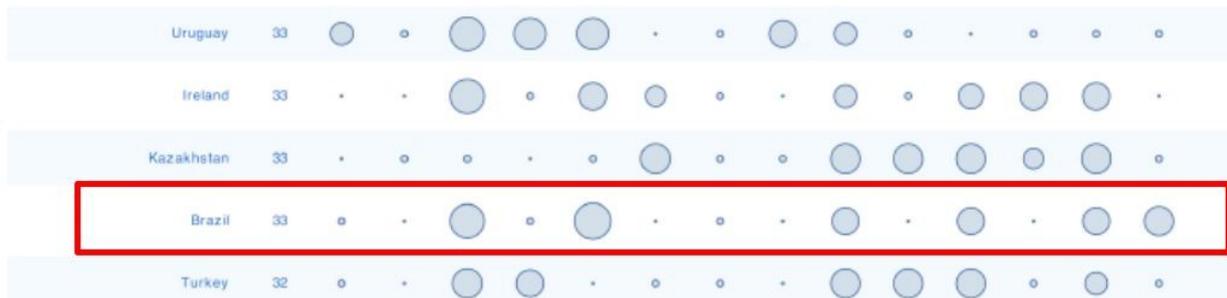


Figura 2 - Análises dos conjuntos de dados (datasets)

(Fonte: Open Data Barometer 2013 Global Report, peça 40)

137. Segundo o estudo, verifica-se que o Reino Unido disponibiliza, em formato aberto, todos os conjuntos de dados cujos conteúdos são considerados de maior relevância para que os benefícios de dados abertos possam ser usufruídos pelo país. Nas categorias analisadas, o estudo conclui que o Brasil não apresenta nenhum dataset que apresente todos os atributos de dados abertos. Na área de educação, apresenta o nível 3 de abertura, e na saúde apresenta o nível 1.

138. O relatório sustenta que o sucesso das iniciativas requer mais do que apenas a disponibilização de conjunto de dados. É necessária a ação de agentes, chamados de intermediários, capazes de transformar os dados governamentais em plataformas e produtos com valor social e econômico, equipados para trabalhar com os dados de diferentes maneiras. Nesse sentido, o Open Data Barometer reconhece que, para se alcançar os benefícios dos dados abertos e assegurar sua sustentabilidade, é necessário cuidar de um número diverso de variáveis, visando construir e manter um ecossistema em torno da infraestrutura dos dados centrais de um programa de dados abertos governamentais.

139. O estudo traz as seguintes considerações (peça 40, p. 6-7):

- a) políticas de dados governamentais abertos têm sido difundidas e publicadas pelos países nos últimos anos, embora a disponibilização de dados abertos, no real sentido de abertura, permanece baixa, com menos de 7% dos conjuntos de dados pesquisados publicados em formato compreensível por máquina e sob uma licença aberta;
- b) países mais avançados em dados abertos (líderes do ranking) estão investindo em infraestruturas de dados abertos nacionais como base para inovação e eficiência público e privada. Todavia, nenhum país pode ser declarado aberto por padrão, antes da difusão das práticas de dados abertos nos órgãos e entidades governamentais;
- c) países intermediários na abertura têm implementado iniciativas de dados abertos, tais como a criação de um portal de dados abertos e a promoção de eventos para induzir o uso desses dados. Todavia, têm falhado na disponibilização de informações relevantes e na implementação dos fundamentos que possibilitam uma efetiva reutilização dos dados. A ausência de lei de acesso à informação, em países específicos, impede o uso dos dados abertos no controle social. A falta de treinamento e incentivo à utilização dos dados compromete a geração de benefícios sociais e econômicos;
- d) países imaturos em dados abertos (últimos do ranking) ainda não começaram a trabalhar com a disponibilização dos dados, e muitos dos países em desenvolvimento não possuem os fundamentos da política, tais como dados gerenciados e digitalizados;
- e) diferentes países e regiões enfrentam diferentes desafios no processo de abertura de dados. Não há uma receita única para a política de dados governamentais abertos;

f) conjuntos de dados relevantes, como o registro de imóveis e registro de empresas, são os menos disponibilizados dentre os conjuntos de dados listados como prioritários. Isso indica que as iniciativas de dados abertos governamentais ainda não garantem a disponibilização de dados politicamente relevantes, que contribuem para a responsabilização de governos e empresas;

g) na maioria dos países, conjuntos de dados voltados ao empreendedorismo e para a melhoria das políticas públicas não são, em geral, disponibilizados ou são disponibilizados em formatos não padronizados;

h) categorias de dados gerenciados por autoridades estatísticas são as mais prováveis de serem acessíveis online, mas, muitas vezes, são disponibilizadas apenas de forma agregada e com licenças restritivas;

i) ainda há carência de evidências dos impactos da utilização dos dados governamentais. Poucos programas de dados abertos são avaliados.

140. O Brasil figura na 28ª posição da análise do Open Data Barometer (dentre 77 países), com uma pontuação de 36,83, numa escala de zero a cem, sendo que a pontuação máxima, cem, pertence ao Reino Unido. Destaca-se o fato de, para o Brasil, a análise apresentar como nulo o impacto econômico oriundo das iniciativas de dados governamentais abertos (peça 40, p. 26).

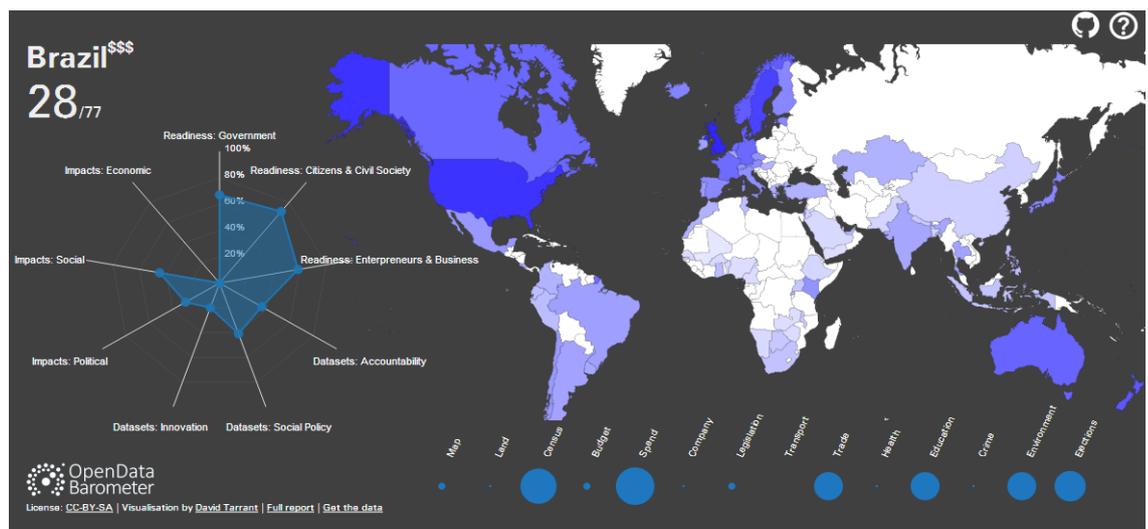


Figura 3. ODB – Dados do Brasil

(Fonte: www.opendatabarometer.org Explore the data: Interactive Charts)

141. Na tentativa de padronizar uma classificação dos dados abertos, Tim Berners-Lee, o inventor da Web, sugeriu um modelo progressivo para a publicação de dados abertos, atribuindo o nome de Five Stars of Linked Data (Figura 4). O modelo busca induzir os governos a darem o primeiro passo para a abertura das informações, publicando seus dados online, no formato em que se encontrarem, e, progressivamente, irem aperfeiçoando seu nível de abertura, tornando-os legíveis por máquina, complementando com dados que contribuam para a inteligibilidade da informação, publicando-os em formatos não proprietários, e, posteriormente, identificando unicamente tais dados e finalmente relacionando esses dados a outros para prover um contexto sobre a informação a que se refere.

142. Essa evolução, em última análise, faria com que todos os requisitos de um dado aberto fossem paulatinamente cumpridos (5stardata.info). Embora citada no relatório do

Open Data Barometer, e também em reunião entre a equipe de fiscalização e representantes da Dataprev (item 4.5 do presente relatório), a equipe de auditoria não evidenciou a utilização desse modelo pelos gestores da política de dados abertos no Brasil.

- ★ make your stuff available on the Web (whatever format) under an open license¹
- ★★ make it available as structured data (e.g., Excel instead of image scan of a table)²
- ★★★ use non-proprietary formats (e.g., CSV instead of Excel)³
- ★★★★ use URIs to denote things, so that people can point at your stuff⁴
- ★★★★★ link your data to other data to provide context⁵

Figura 4 - Exemplos para cada nível do modelo de classificação de abertura de dados 5 star Open Data

(Fonte: <http://5stardata.info>)

National Health Service (NHS)

143. *Métricas são elementos críticos para a avaliação do sucesso ou do fracasso de um projeto ou atividade. Além da preocupação em se mensurar o nível de abertura de uma nação, a preocupação com a avaliação dos impactos reais da abertura dos dados tem crescido na comunidade internacional.*

144. *A consciência de que há pouco entendimento de se e como as iniciativas de dados abertos realmente conduzem aos resultados prometidos (item 2.5 do presente relatório) é unânime. Diante disso, o Sistema de Saúde da Inglaterra, National Health Service (NHS), publicou recente artigo, elaborado pelo GovLab (thegovlab.org), com a proposta de uma estrutura conceitual destinada a auxiliar na mensuração dos impactos de abertura de dados na área de saúde e na avaliação de circunstâncias específicas que os potencialize (peça 44, p. 6).*

145. *A intenção dos autores é que esse modelo possa ser usado por pesquisadores e gestores de programas no momento da definição de iniciativas de dados abertos na área de saúde, estabelecendo prioridades e forma de mensuração de seu impacto, não somente ao término do processo de implementação das iniciativas, mas também ao longo dele (peça 44, p. 7). Cabe citar que o GovLab é um centro de pesquisa da Universidade de Nova York, com atuação global, que busca novos caminhos para solucionar problemas públicos utilizando os avanços da tecnologia e da ciência. Seu objetivo é avançar no conhecimento de como o comprometimento do cidadão pode contribuir para uma governança mais legítima e efetiva (peça 44, p. 81).*

146. *Com a ambição de que esse se torne um modelo mundial para a abertura de dados da saúde, o NHS define seis proposições de valor decorrentes do uso de dados abertos da saúde:*

- a) prestação de contas;*
- b) escolha;*
- c) eficiência;*
- d) serviços prestados e satisfação do paciente;*

- e) crescimento econômico e inovação;
- f) resultados.

147. Para que os dados abertos agreguem valor, o modelo propõe, ainda, a quebra do processo de abertura em sete etapas: inputs, usuários, atividades, produtos, resultados, impactos e metodologia (peça 44, p. 64).

148. A relevância do estudo, que, com as adaptações necessárias, pode ser estendido a todas as demais áreas da administração pública, encontra-se no estabelecimento da mensuração dos resultados em um momento inicial, no qual se define a iniciativa de abertura de dados. Isso reflete a devida importância que deve ter a avaliação dos resultados potenciais e reais relativos ao processo de abertura de dados.

3.2 Exemplos internacionais de iniciativas de abertura de dados

149. Com a abertura dos dados governamentais, organizações, cidadãos, acadêmicos, empresas e instituições públicas têm construído soluções para a produção e o compartilhamento da informação pública, levando conhecimento e estabelecendo um canal de comunicação direto do governo com o cidadão. Muitas dessas soluções incluem a interação com a sociedade, de forma que ela possa contribuir com informações específicas do seu próprio ambiente, estabelecendo, assim, um rico fluxo bidirecional de informação.

150. As sociedades ao redor do mundo já contam com soluções para algumas de suas demandas. Também chamadas de aplicações cívicas, as soluções podem ser voltadas para agilizar os serviços públicos oferecidos aos cidadãos, promover a cooperação na execução dos serviços públicos, levar conhecimento, criar novas oportunidades de negócios e empregos, guiar o governo na elaboração ou condução das políticas públicas ou empoderar o cidadão. Durante a fase de execução deste levantamento, a equipe de fiscalização realizou pesquisa e selecionou algumas soluções internacionais com o intuito de ilustrar essa realidade.

3.2.1 England's best places to live – Inglaterra – <http://illustreets.co.uk/>

151. O serviço proposto por esse sítio auxilia o cidadão na escolha da sua moradia na Inglaterra, apresentando indicadores da região selecionada. A solução apresenta um mapa interativo, colorido de acordo com o padrão de vida, no qual o cidadão seleciona o ponto (região) para a pesquisa. São, então, apresentados inúmeros indicadores da região selecionada, tais como: padrão de vida (ONS 2010), índice de criminalidade, preços médios de aluguéis, características da população da região, taxa de desemprego, distribuição da utilização (parques, prédios, residências, estradas etc.), escolas da região e da redondeza, e transporte público (illustreets.co.uk/about-illustreets). A solução oferece, ainda, o detalhamento dessas informações. Ao clicar no menu referente a escolas, por exemplo, são listadas as cinco escolas mais próximas e apresentados os indicadores.

152. Os dados abertos utilizados na solução são oriundos de várias fontes governamentais, incluindo gov.uk, police.uk, departamento de transporte e departamento de educação (illustreets.co.uk/about-illustreets).

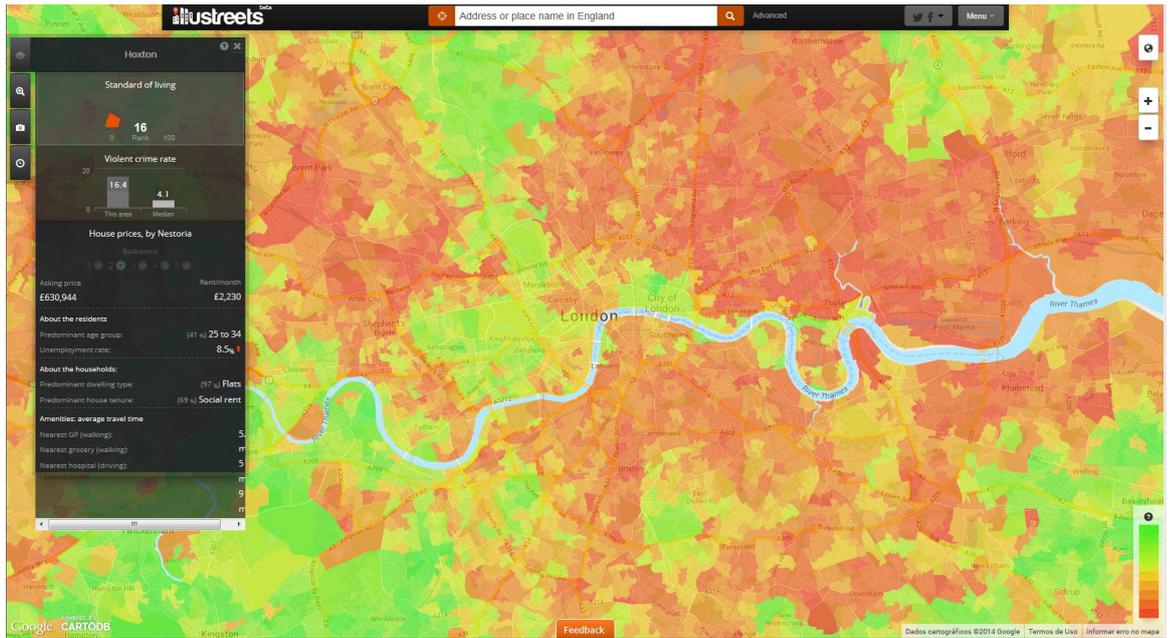


Figura 5 - Illustreets - Padrão de vida por regiões

(Fonte: <http://illustreets.co.uk/explore-england>)

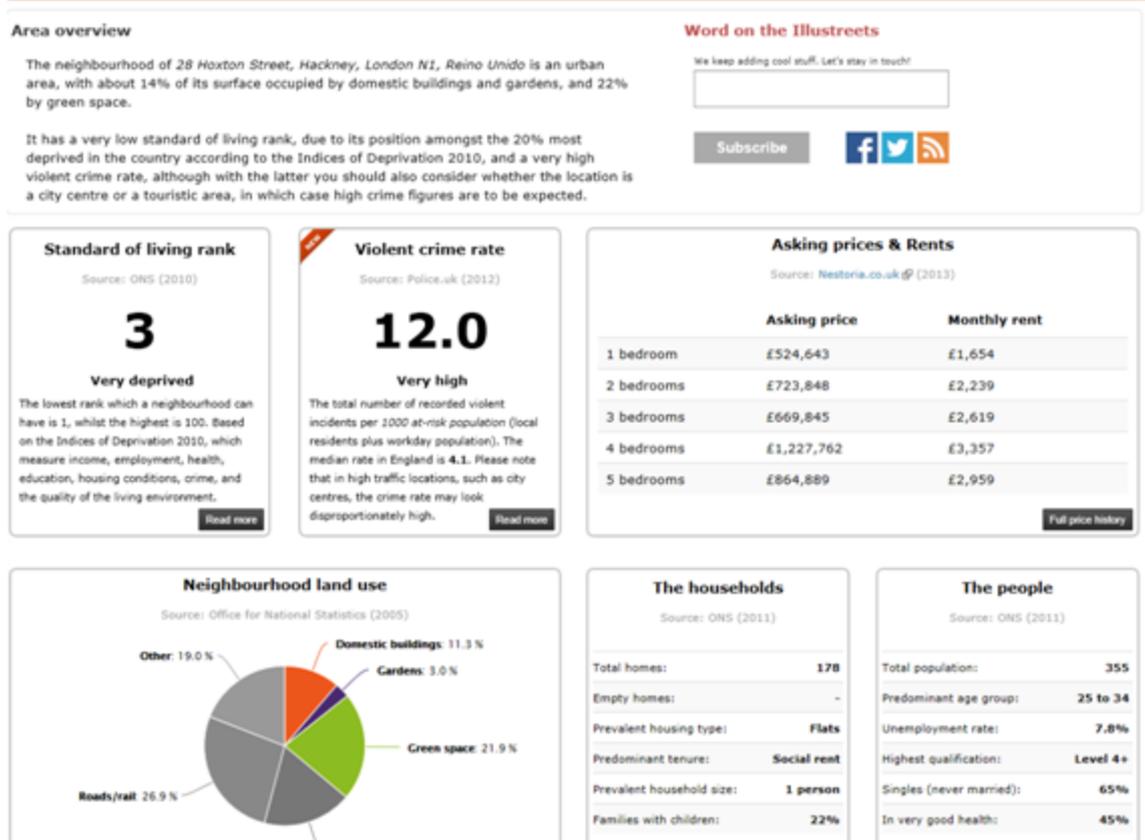


Figura 6 - Detalhamento dos indicadores de 28 Hoxton Street, Londres

(Fonte: <http://illustreets.co.uk/explore-england/28-hoxton-street-londres/overview>)

3.2.2 [iTriage – EUA – https://www.itriagehealth.com](https://www.itriagehealth.com)

153. A solução americana iTriage oferece ao usuário informações sobre clínicas, hospitais, medicamentos, procedimentos, médicos, apresentando inclusive o currículo resumido dos médicos e os links dos telefones para agendamento de consultas. Traz, ainda, uma funcionalidade que realiza um pré-diagnóstico, indicando as possíveis doenças

relacionadas aos sintomas informados pelo usuário. Na construção, são utilizadas várias bases de dados, inclusive dados governamentais abertos (peça 56). As informações de médicos e hospitais são georreferenciadas, com um mapa de visualização de localização, permitindo a subseleção por especialidade.

154. Segundo artigo da revista *ComputerWorld*, publicado em 9 de julho de 2013, aproximadamente 9,5 milhões de pessoas haviam feito o download do aplicativo na época (peça 56).

3.2.3 [School Comparison – EUA – http://dashboard.dcpsb.org/comparison](http://dashboard.dcpsb.org/comparison)

155. Essa solução permite comparações entre escolas do distrito de Columbia, EUA, com a visualização de resultados em diversas métricas, tais como: evolução dos estudantes em leitura e matemática no tempo, percentuais de estudantes que obtiveram nível avançado ou de proficiência em leitura e matemática; resultados em disciplinas chave para o sucesso educacional, bem como apresenta indicadores sobre a frequência dos estudantes, percentual de estudantes que renovam suas matrículas a cada ano etc. O sítio apresenta a funcionalidade de simulador, em que diretores podem, a partir de possíveis melhoras em algumas dimensões onde pode influir ou que pode priorizar, testar os efeitos sobre diversas métricas de avaliação da sua escola.

156. A solução é gerenciada pelo Public Charter School Board (PCSB) do Distrito de Columbia e os dados utilizados na solução estão disponibilizados no portal data.dcpsb.org.

3.2.4 [US Hospital Finder – EUA – http://www.ushospitalfinder.com](http://www.ushospitalfinder.com)

157. A solução US Hospital Finder, oferecida em uma página na internet e também como aplicativo móvel (itunes.apple.com/us/app/ushospfinder/id412630399), identifica os hospitais próximos a uma localidade dos Estados Unidos especificada pelo usuário, permitindo que sejam obtidas informações detalhadas dos hospitais que o usuário selecionar. No detalhamento, são apresentados os serviços oferecidos pelo hospital e alguns indicadores de qualidade, como taxas de mortalidade comparada com números nacionais, e também os índices de satisfação dos usuários do hospital.

158. A solução declara utilizar dados governamentais abertos (ushospitalfinder.com/abouts/#terms).

Location:

If you like this site, please try [US Nursing Home Finder](#), [US Home Care Finder](#), [Find Good School](#) and [US F](#)

Results for:"10013"

Found 155 hospital(s) within 30 miles

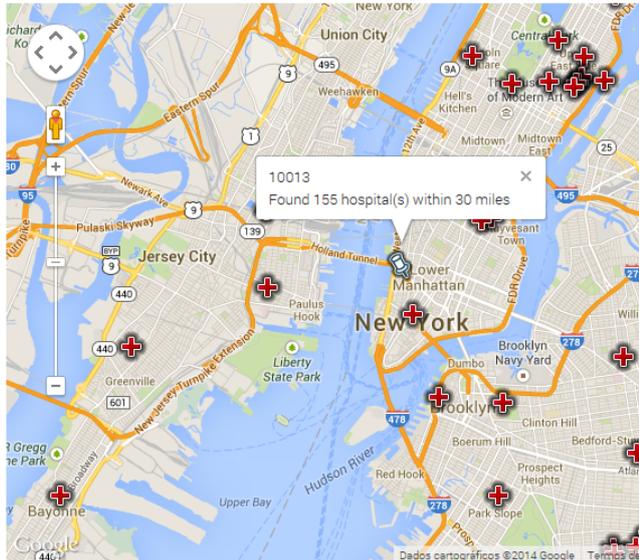


Figura 8 - Detalhamento de um hospital na solução US Hospital Finder (Fonte: <http://www.ushospitalfinder.com>)

Figura 7 - Resultado de pesquisa do código 10013 na solução US Hospital Finder (Fonte: <http://www.ushospitalfinder.com>)

3.2.5 [Sistema de Medición de la Calidad de la Educación \(SIMCE\) – Chile –](http://www.simce.cl/mapas/)

159. Apresenta o mapa de todo o Chile com um identificador apontando para o local onde estão os estabelecimentos de ensino públicos e privados, permitindo ao usuário selecionar uma escola por localização e também pelo nome do estabelecimento de ensino. Para a escola selecionada, são apresentados os resultados das avaliações daquela escola em cada nível oferecido, bem como os resultados referentes às avaliações das disciplinas abarcadas pelo sistema de medição educacional nacional.

3.2.6 [Citizen Dashb@ard – Canadá –](https://dashboard.edmonton.ca/)

160. Constitui-se em um painel que mostra os dados referentes ao desempenho dos serviços providos pela cidade de Edmonton, no estado de Alberta, Canadá, em seis diferentes domínios: transporte público, qualidade de vida, meio-ambiente, questões urbanas, economia e finanças. Ao selecionar um serviço de um domínio específico, o cidadão tem acesso a dados mais detalhados. Essas informações são apresentadas graficamente e também disponibilizadas para download em diversos formatos, como pode ser visto em <https://dashboard.edmonton.ca/dataset/Security-Incidents-Citizen-Dashboard-Data/aax4-5dzb>.

3.2.7 [Crime Incidents – EUA –](https://data.sfgov.org/Public-Safety/Map-Crime-Incidents-Previous-Three-Months/gxxq-x39z)

161. Um dos aplicativos mais populares do sítio de dados abertos da cidade de São Francisco, na Califórnia, EUA (data.sfgov.org), é o que localiza no mapa da cidade os crimes ocorridos nos últimos três meses. O usuário pode ter a visão da cidade inteira considerando o número total de ocorrências, como também ir ampliando o mapa e ir restringindo a região pesquisada até chegar a uma quadra específica. Nesse caso, podem ser identificados os endereços onde houve o crime e outros detalhes, como categoria,

horário, data, número da ocorrência e a solução dada ao caso. A página disponibiliza essas informações em vários formatos, inclusive em tabelas.

3.2.8 [Chicago Lobbyists – EUA – http://www.chicagolobbyists.org](http://www.chicagolobbyists.org)

162. Trata-se de um projeto destinado a aperfeiçoar a transparência das interações entre a cidade de Chicago, os lobistas e seus clientes. Foi criado por um grupo de desenvolvedores e projetistas que constroem aplicativos cívicos utilizando dados abertos. As informações utilizadas nesse projeto são oriundas do portal de dados da cidade de Chicago (City of Chicago Data Portal), já que norma municipal estabelece que todos os lobistas são registrados e devem informar, semestralmente, sobre suas atividades (chicagolobbyists.org/about).

163. A solução apresenta o número total de lobistas da cidade, de firmas de lobby, de clientes e o valor total pago aos lobistas. Também são apresentados os nomes dos lobistas e das firmas que mais faturam, dos clientes mais ativos, bem como os departamentos e agências do governo municipal que mais trataram com lobistas.

164. Os desenvolvedores destacam que, além de levar conhecimento a cidadãos, jornalistas e o próprio governo, o aplicativo pode ser usado como fonte de informação tanto para firmas que querem contratar lobistas competentes, com base no seu desempenho, custo e experiência, quanto para os próprios lobistas, permitindo a comparação de seu desempenho perante a classe.

165. Com base nesse aplicativo, foi criado o aplicativo brasileiro Política Aberta, que é detalhado no parágrafo 222.

3.2.9 [Open Counter – EUA – http://opencounter.us](http://opencounter.us)

166. O sítio Open Counter foi criado pela parceria entre a organização não governamental Code for America e a cidade de Santa Cruz, na Califórnia, com o objetivo de agilizar o processo de criação de novos negócios. A solução provê aos donos de pequenos negócios uma interface simples, fácil de navegar e pela qual são executados todos os passos necessários para que sejam obtidas as licenças para o funcionamento do novo empreendimento.

167. Essas são algumas das soluções dentre as milhares hoje existentes no mundo que utilizam dados abertos governamentais. De forma complementar, o leitor que tiver interesse em conhecer além pode consultar: data.gov.uk/apps, data.gov/applications e data.gov.au/related.

3.3 [Exemplos internacionais relevantes](#)

168. Uma vez que o tema dados abertos pode abranger todas as áreas finalísticas de governo, decidiu-se selecionar duas áreas para as quais seria voltado um olhar mais atento neste trabalho. Considerando a criticidade e essencialidade das áreas de educação e saúde, bem como o grande volume de informações que elas produzem, essas foram as áreas escolhidas.

169. Neste capítulo, serão apresentadas soluções internacionais de destaque para cada uma dessas áreas. No capítulo seguinte, Panorama Nacional, será realizada uma comparação entre as iniciativas brasileiras na área de educação e saúde frente às soluções aqui descritas.

3.3.1 [Exemplo internacional relevante na área de educação](#)

170. Em relação a soluções internacionais que utilizam dados públicos educacionais, verificou-se que há uma predominância de soluções dedicadas a fornecer informações

sobre as características das escolas. Dessa maneira, a equipe de fiscalização selecionou um desses aplicativos para detalhamento.

171. A solução *FindTheBest Primary Schools* (<http://primary-schools.findthebest.co.uk/d/a/London>), que utiliza dados abertos do governo do Reino Unido, se destina a auxiliar os pais na escolha da escola de seu filho, fornecendo informações sobre o desempenho das escolas de educação básica. Sua base de dados contém mais de dezenove mil escolas, que são apresentadas dispostas no mapa do Reino Unido ou em uma lista classificada conforme parâmetros definidos pelo próprio usuário. Além disso, a solução permite que o usuário selecione as escolas candidatas para uma futura comparação e apresentados os desempenhos e atributos das escolas selecionadas.

172. Para ajudar no processo de escolha, é apresentado, também, um guia sobre como escolher a escola certa, com informações do *Department for Education* e do *Office for Standards in Education*, ambas unidades federais do Reino Unido.

173. Para exemplificar a utilização da solução, foram listadas as escolas no bairro de Harrow, subúrbio de Londres.

DISTANCE	SCHOOL NAME	SMART RATING	TOWN	TOTAL PUPILS	AGE RANGE	SCHOOL TYPE
1.2 km	Roxeth Primary School COMPARE SEE DETAILS >	79	Harrow	431 PUPILS	3-11 YEARS	Community School
1.6 km	Grange Primary School 5.0 1 user rating COMPARE SEE DETAILS >	61	Harrow	495 PUPILS	3-11 YEARS	Community School
1.9 km	St George's Primary School COMPARE SEE DETAILS >	87	Harrow	467 PUPILS	4-11 YEARS	Voluntary Aided School
1.9 km	Whitefriars Community School COMPARE SEE DETAILS >	80	Harrow	431 PUPILS	3-11 YEARS	Community School
2.0 km	Longfield Primary School COMPARE SEE DETAILS >	77	Harrow	698 PUPILS	3-11 YEARS	Community School
2.1 km	Avanti House School COMPARE SEE DETAILS >	NOT YET RATED Unrated	Harrow		4-18 YEARS	Free School

MY PRIMARY SCHOOLS (1) Select at least 2 listings to compare Clear all

Figura 9 - Escolas de Harrow, Londres

(Fonte: <http://primary-schools.findthebest.co.uk/d/a/London>)

174. Foram, então, selecionadas três escolas e discriminadas as suas informações. Observou-se que foram apresentados diversos indicadores, tais como o desempenho global da escola, o desempenho e o progresso em cada uma das disciplinas, o tamanho da escola, número de professores, a média de alunos por professor, a distribuição por gênero, a religião dominante, a receita da escola, a distribuição dos gastos da escola etc.

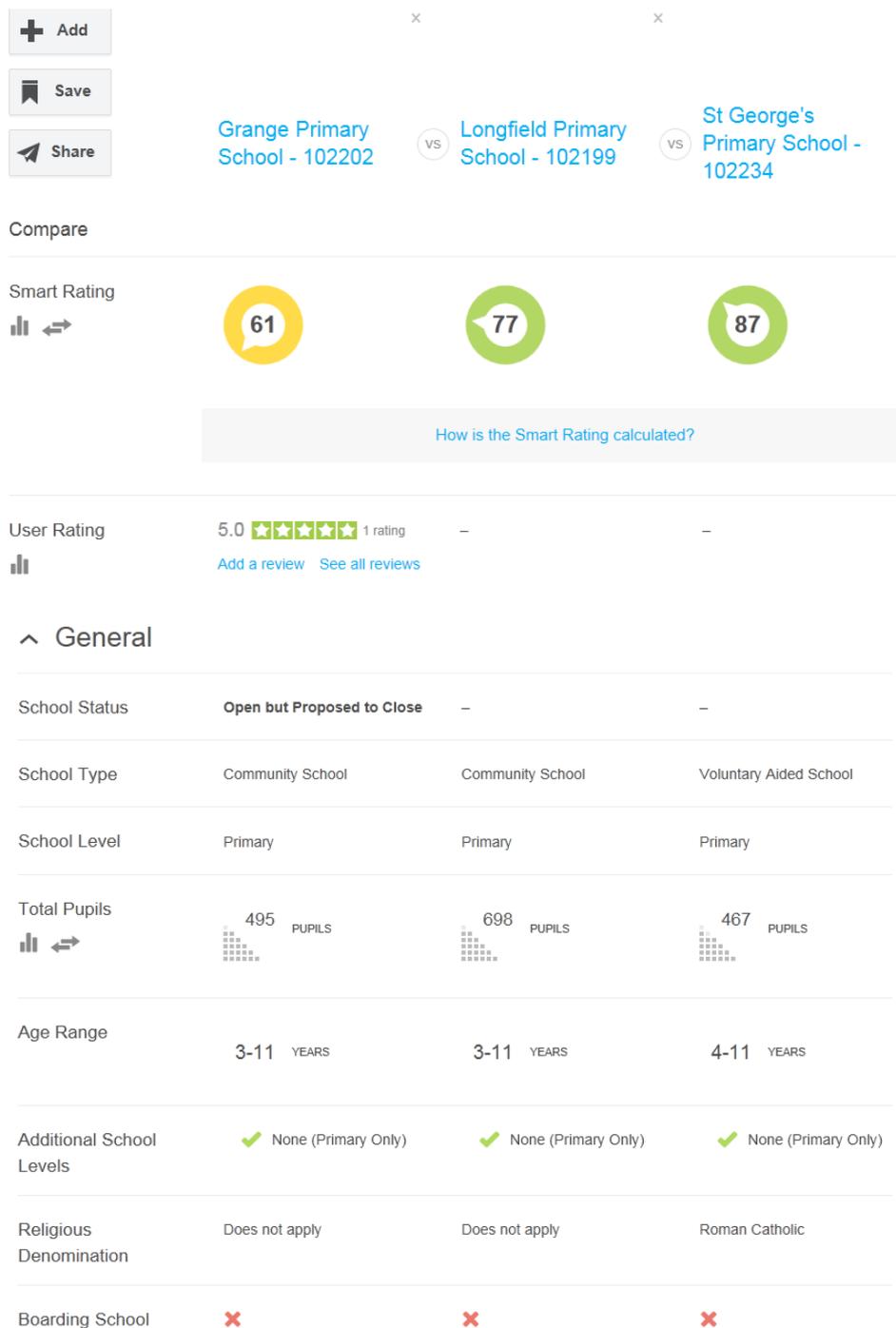


Figura 10 - Comparação de escolas de Harrow
 (Fonte: <http://primary-schools.findthebest.co.uk/d/a/London>)

Support Staff (Full-time Equivalent)	7.2 SUPPORT STAFF	11.4 SUPPORT STAFF	7.9 SUPPORT STAFF
--------------------------------------	-------------------	--------------------	-------------------



Income

Total Income	£4,617 PER PUPIL	£4,077 PER PUPIL	£4,826 PER PUPIL
--------------	------------------	------------------	------------------



Grant funding	£4,490 PER PUPIL	£3,985 PER PUPIL	£3,917 PER PUPIL
---------------	------------------	------------------	------------------

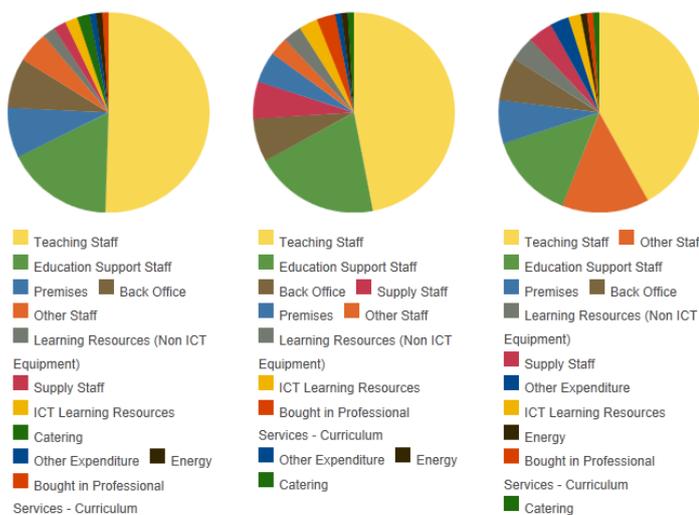


Self Generated Income	£126 PER PUPIL	£92 PER PUPIL	£909 PER PUPIL
-----------------------	----------------	---------------	----------------



Expenditures

Expenditure Breakdown



Total Expenditures	£4,322 PER PUPIL	£3,242 PER PUPIL	£4,749 PER PUPIL
--------------------	------------------	------------------	------------------



Figura 11 - Comparação de escolas de Harrow

(Fonte: <http://primary-schools.findthebest.co.uk/d/a/London>)

175. Trata-se, portanto, de uma solução simples e informativa para o cidadão. A maioria dos dados apresentados são dados brutos, ou seja, sem tratamento, que foram apresentados de forma ilustrativa e de fácil entendimento, que contribui para o processo de decisão do cidadão.

176. No capítulo seguinte deste relatório (item 4.6), serão descritas algumas das soluções brasileiras que utilizam dados do Sistema Nacional de Educação, para que possa ser realizado a comparação proposta no planejamento deste trabalho.

3.3.2 Exemplos internacionais relevantes na área de Saúde

177. Pesquisa realizada pela equipe de fiscalização mostrou que vários países têm buscado a melhoria nas políticas de saúde pública por meio da abertura dos dados de saúde. Ao se consultar sítios governamentais internacionais que disponibilizam esses dados, pode-se ter uma noção do nível de abertura e conhecer as ações realizadas no sentido de tornar públicas as informações e estatísticas relacionadas ao bem-estar do cidadão (a exemplo de healthdata.gov, cdc.gov/sortablestats, aihw.gov.au e caredata.co.uk).

178. O National Health Service (NHS), serviço de saúde pública da Inglaterra, por exemplo, tem feito significantes investimentos para a disponibilização de seus dados e nos últimos anos realizou inúmeras iniciativas nesse sentido (peça 44, p. 6). Atualmente, ele busca ser instituição modelo na abertura dos dados da saúde e, dentre as ações implementadas, lançou o programa *Better Information Means Better Care* (melhor informação implica em melhor saúde – tradução livre). Esse programa defende que o compartilhamento de informações, dentre outros benefícios, contribui para que a qualidade e segurança dos serviços sejam uniformemente distribuídos no país, e também que os investimentos possam ser direcionados para as áreas e enfermidades que realmente necessitam (www.england.nhs.uk/ourwork/tsd/care-data).

179. O Estado de Illinois, nos Estados Unidos, também tem fomentado a abertura e o uso dos dados da saúde. Para isso, foi criado o consórcio Health Data Consortium entre governo, organizações sem fins lucrativos e organizações do setor privado para trabalharem juntos na intenção de que o uso dos dados conduza a inovações que aperfeiçoem o sistema de saúde (www.smartchicagocollaborative.org/making-public-health-data-work-in-illinois).

180. Quanto à utilização dos dados da saúde, observou-se uma grande variedade de soluções internacionais utilizando dados abertos. Por exemplo, existem soluções para: i) identificar hospitais próximos a uma localidade (US Hospital Finder); ii) fornecer informações sobre hospitais, clínicas, medicamentos, procedimentos, médicos, telefones para agendamento e realização de pré-diagnóstico (iTriage); iii) auxiliar na escolha de postos de atendimento (hospitals.findthebest.com), iv) comparar preços dos procedimentos mais comuns (data.cms.gov/Medicare/Inpatient-Prospective-Payment-System-IPPS-Provider/97k6-zzx3); v) prover informações sobre o tratamento de doenças e avaliar os medicamentos prescritos pelos médicos (prescribinganalytics.com); vi) comparar os gastos e resultados entre hospitais de modo a promover a prestação de contas; vii) mapear doenças no país, entre tantos, incluindo-se soluções que fornecem o acesso dos pacientes aos seus próprios registros médicos.

181. Dentre as soluções pesquisadas, uma solução simples e de enorme utilidade é a identificação do melhor e mais próximo hospital ou posto de atendimento ao cidadão que necessita de atendimento e encontra-se em uma determinada localidade. Há uma grande oferta desse tipo de solução em alguns países. Para exemplificar, cita-se, aqui, a solução *FindTheBestHospitals* (hospitals.findthebest.co.uk), uma aplicação similar àquela trazida na área da educação: a solução *FindTheBest Primary Schools* (primary-schools.findthebest.co.uk/d/a/London).

182. A solução *FindTheBestHospitals* identifica hospitais próximos a uma localidade em Londres e os compara com base na avaliação do paciente, corpo clínico, estatísticas das cirurgias, taxa de mortalidade e tempo de atendimento, dentre outros.

183. Para exemplificar a utilização da solução, foram selecionados dois hospitais e discriminadas as suas informações: Torbay Hospital e Freeman Hospital. Ambos hospitais apresentam alto índice de aceitação 83% e 86%, respectivamente. Embora com maior índice de aceitação, o hospital Freeman tem uma taxa de mortalidade 1,63% maior que o hospital Torbay. Por outro lado, quanto à melhoria na saúde relatada pelo paciente operado, o hospital Torbay apresenta o índice de 76%, superando em 4% o hospital Freeman, com 72%.

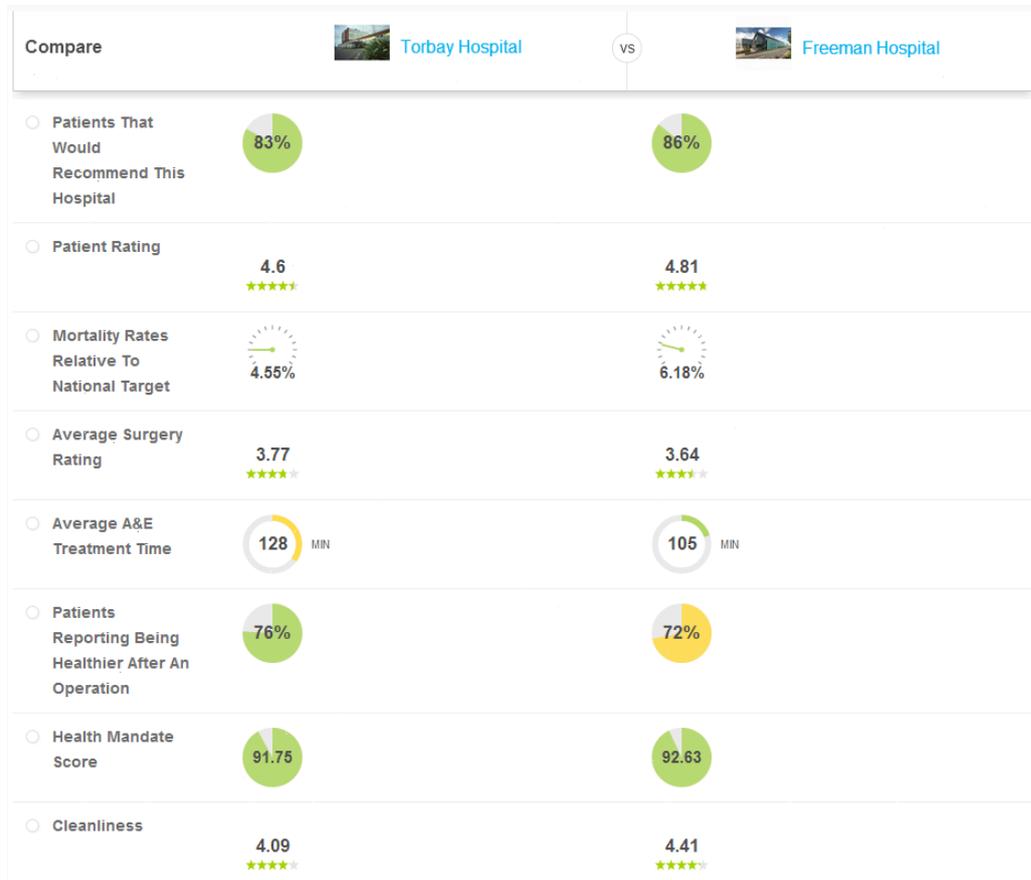


Figura 12 - Indicadores de hospitais do Reino Unido

(Fonte: <http://hospitals.findthebest.co.uk/compare/429-949/Torbay-Hospital-vs-Freeman-Hospital>)

184. Observa-se, do estudo, que organizações de vários países têm se empenhado em compartilhar princípios e práticas de governo aberto e disseminá-los entre os governos, mas o tema dados abertos ainda constitui um domínio em evolução. Estudos apontam que os governos avaliados ainda não se desenvolveram convenientemente quanto à disponibilização de dados governamentais brutos. Além da preocupação em se mensurar o nível de abertura de uma nação, a preocupação com a avaliação dos impactos reais da abertura dos dados, ou seja do valor agregado, tem crescido na comunidade internacional. Nesse sentido, metodologias e parâmetros de medição têm sido desenvolvidos e testados por organizações.

185. A partir dos exemplos listados nesta seção, conclui-se pela existência de extenso conjunto de soluções tecnológicas internacionais que utilizam dados abertos e se voltam à resolução de problemas diversos da sociedade, tendo sido construídas a partir de iniciativas privadas e governamentais de âmbito nacional, estadual e municipal.

4. Panorama Nacional de Disponibilização de Dados Governamentais em Formatos Abertos

186. *Esse capítulo tem o objetivo de apresentar o cenário de dados governamentais abertos no âmbito nacional, trazendo algumas iniciativas de entes dos Poderes Executivo, Legislativo e Judiciário e nas três esferas governamentais.*

187. *Em seguida, são trazidos aspectos gerais da legislação brasileira correlata a acesso à informação, dados abertos, governo aberto e participação social, bem como comentários relativos às competências e à atuação dos órgãos nacionais gestores desses assuntos.*

188. *Durante a etapa de execução do levantamento, foram coletados exemplos de soluções construídas a partir de dados governamentais, sendo listadas, neste capítulo, as mais interessantes na avaliação da equipe.*

189. *Ademais, são abordados alguns aspectos do Portal Brasileiro de Dados Abertos, bem como descrita a atuação das empresas públicas prestadoras de serviços de TI em dados abertos e Big Data.*

190. *Considerando a relevância e impacto das ações de governo nas funções de saúde e educação, o capítulo apresenta o cenário da abertura de dados observado em dois entes públicos responsáveis por custodiar dados públicos nas áreas de educação e saúde, a saber, o Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep) e o Departamento de Informática do SUS (Datasus).*

191. *A abertura dos dados governamentais deve ser capaz de, entre outros efeitos, induzir que organizações, cidadãos, acadêmicos, empresas e instituições públicas construam soluções para a produção e o compartilhamento da informação pública, levando conhecimento ao cidadão e estabelecendo um canal de comunicação direto do governo com o cidadão. Isso já ocorre no cenário internacional e, de forma ainda acanhada, também no Brasil.*

192. *Quanto a esse processo de abertura, as dificuldades relatadas pelos órgãos entrevistados foram: sustentabilidade das iniciativas geradas, em especial, das soluções construídas pela sociedade ou de forma colaborativa; falta de pessoal e de governança na gestão do processo; e falta de padronização.*

4.1 Iniciativas nacionais de destaque

193. *A pesquisa sobre as iniciativas de dados abertos na esfera federal mostrou heterogeneidade do nível de abertura entre órgãos e entidades da APF. As entrevistas realizadas com alguns órgãos e também com as unidades gestoras das ações relacionadas a dados governamentais aberto apontaram que, em geral, as iniciativas e o avanço na abertura de dados evidenciado em alguns órgãos se devem em grande parte ao interesse pessoal de gestores, os quais, com a crença de se alcançar os resultados prometidos, encampam o tema e trabalham pela abertura das informações.*

194. *Dentre as iniciativas de dados abertos pesquisadas, destacam-se aquelas realizadas pelas seguintes instituições: Câmara dos Deputados, Ministério da Justiça, Tribunal de Contas do Estado de Pernambuco, além do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep) e do Departamento de Informática do SUS (Datasus), esses últimos descritos no item 4.6 deste relatório.*

4.1.1 Iniciativas de dados abertos na Câmara dos Deputados

195. *A Câmara dos Deputados se destaca entre os órgãos da APF por possuir estrutura física e de pessoal específicas para implementar ações de dados abertos, além do*

apoio das instâncias superiores da Casa. Com o objetivo de promover ações colaborativas para o aprimoramento da transparência legislativa e da compreensão do processo legislativo, em 2013, a Câmara dos Deputados realizou um Hackathon, no qual foram selecionados e premiados três aplicativos: Meu Congresso, Monitora Brasil e Deliberatório, criados a partir de dados gerados pela Casa. Segundo os gestores, o evento contribuiu para a divulgação das informações legislativas, bem como para a evolução e para o aperfeiçoamento das informações, na medida em que os desenvolvedores das soluções contribuíram ao apontar problemas e oportunidades de melhoria nas bases de dados da instituição (peça 55).

196. *Em janeiro de 2014, a Câmara dos Deputados inaugurou em sua estrutura organizacional o Laboratório Hacker. Criado por meio da Resolução 49, de dezembro de 2013 (peça 46), e localizado fisicamente na Câmara dos Deputados, conta com espaço físico de acesso e uso livres para qualquer cidadão, especialmente programadores e desenvolvedores, que podem utilizar os dados públicos para construir ações de cidadania. Além da infraestrutura, o laboratório possui nove servidores públicos dedicados exclusivamente às iniciativas de divulgação das informações do órgão e promoção da participação social. A equipe de fiscalização visitou o local, onde também foi realizada reunião com os gestores. Na ocasião, observou-se a presença de um dos desenvolvedores que participou do Hackathon, externo ao quadro da Câmara, utilizando a infraestrutura do laboratório.*

197. *Em junho de 2014, foi lançado o projeto-piloto Dados Abertos 2.0, a ser desenvolvido com o objetivo de experimentar, testar e desenvolver processos, tecnologias, ferramentas, documentação e governança para fornecimento de dados abertos pela Câmara dos Deputados.*

198. *Vale comentar, dada a ligação umbilical entre política de dados abertos e governo aberto, sobre uma outra ferramenta criada pela Câmara, o Portal e-Democracia, que tem o objetivo de incentivar a participação social. O e-Democracia é dividido em dois grandes espaços de participação: as comunidades legislativas e o espaço livre. Uma comunidade legislativa é um espaço de participação em que o cidadão pode participar de debates de temas específicos, normalmente relacionados a projetos de lei já existentes. Essas comunidades oferecem diferentes instrumentos de participação e, ainda, orientações quanto ao andamento da matéria no Congresso Nacional. Já no espaço livre, a sociedade pode definir o tema da discussão e ser a grande motivadora dele. O debate é acompanhado pela equipe e-Democracia (edemocracia.camara.gov.br/o-que-e#U8IRLPldWN0). Uma dessas comunidades, por exemplo, refere-se ao Laboratório Hacker, e se constitui em um espaço de discussão sobre o projeto Dados Abertos 2.0 entre a equipe da Câmara, representantes de outros órgãos de governo e da sociedade.*

4.1.2 Iniciativas de dados abertos no Ministério da Justiça

199. *O Ministério da Justiça é outro órgão que tem desenvolvido ações para a promoção da abertura de seus dados, estando, inclusive, desenvolvendo seu Plano de Dados Abertos e Espaciais, que se encontrava sob consulta pública durante a elaboração deste relatório (peça 47 e participa.br/dadosabertos/dados-abertos-no-ministerio-da-justica). O Ministério já realizou dois Hackathons em 2013, e possui em torno de uma dezena de conjunto de dados (datasets) catalogados no Portal Brasileiro de Dados Abertos (dados.gov.br/tag/MJ).*

4.1.3 Iniciativas de dados abertos em Institutos de Pesquisa

200. *As entidades brasileiras que são provedoras de dados e informações do país também disponibilizam parte do grande volume de informações que possuem ou custodiam.*

201. *O Instituto de Pesquisa Econômica Aplicada (Ipea), por exemplo, por meio de sua página Ipeadata.gov.br, permite o acesso aos dados de séries estatísticas por ele geradas ou custodiadas, indicando a fonte da informação: nacional ou internacional. As informações abrangem dados econômicos, financeiros, demográficos, geográficos e indicadores sociais, podendo ser consultadas por temas, fontes dos dados e outros filtros. As séries mais usadas também podem ser acessadas na página Ipeadata.gov.br, sem, contudo, informar o quantitativo de acesso a essas séries. Os dados são disponibilizados no formato csv (Comma-separated values), em planilha Excel e também são apresentados graficamente. Arquivos no formato csv são arquivos texto, onde o conteúdo é tabular, separado por vírgulas e organizado sequencialmente por linhas. Entretanto, observou-se carência de metadados sobre os dados, sendo feita, apenas, a referência aos sítios cujos dados foram obtidos.*

202. *O Instituto Brasileiro de Geografia e Estatística (IBGE), por sua vez, disponibiliza no seu sítio informações provenientes de dados oficiais oriundos de pesquisas por ele realizadas, em sua maior parte, e de outras fontes governamentais (seriesestatisticas.ibge.gov.br/apresentacao.aspx). Dados do Censo também são fornecidos, embora não em formato bruto, mas agregados por Estado ou município (cidades.ibge.gov.br/xtras/temas.php?lang=&codmun=530010&idtema=94&search=distr+ito-federal|brasilia|censo-demografico-2010:-resultados-da-amostra-domicilios--). A pesquisa nos dados disponibilizados permite constatar que, embora a informação esteja acessível ao público em geral, os dados não atendem aos princípios de dados abertos (parágrafo 51), pois, por exemplo, muitos deles não são compreensíveis por máquina e não são primários (brutos).*

203. *Tendo em vista que os dados produzidos pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep) referem-se à área de educação, descrição das iniciativas de dados ali verificadas consta na seção 4.6.1.*

4.1.4 Iniciativas de dados abertos em Estados e Municípios

204. *Também na esfera estadual, o nível de comprometimento de cada estado com a abertura de dados mostra-se heterogêneo.*

205. *O Estado de São Paulo, por exemplo, lançou a iniciativa Governo Aberto SP, possui uma política de Gestão do Conhecimento e Inovação (peça 48, p. 34-36) e, recentemente, firmou parceria com o Reino Unido com o intuito de ampliar as bases de dados disponibilizadas, melhorar a acessibilidade dos portais oficiais do Estado e estimular a criação de novos aplicativos, serviços e negócios.*

206. *O sítio do Governo Aberto SP (www.governoaberto.sp.gov.br), consultado em 4/7/2014, contava com 390 conjuntos de dados (datasets) cadastrados e somente 22 disponibilizados para download. Dos 22 conjuntos de dados acessíveis por download, os dados estavam em um dos seguintes formatos: planilha Excel, arquivo texto, arquivo no formato csv, acesso a um web service – software que permite a uma aplicação acessar dados de outra por meio do envio e recebimento de dados em formato xml –, e acesso por uma Application Programming Interface (API). Uma API constitui um conjunto de rotinas e padrões estabelecidos por um software que age como intermediário entre os dados e os*

sistemas que vão reutilizá-los, permitindo que programas e aplicativos tenham acesso automático aos dados (pt.wikipedia.org/wiki/API).

207. Além disso, o Portal da Transparência Estadual (transparencia.sp.gov.br) oferece ao usuário consultas às informações patrimoniais e financeiras, receitas e despesas, apresentando os dados na página da internet ou em formato PDF, permitindo o download de algumas informações em planilhas Excel, além do acesso por web service aos dados orçamentários, o qual constitui o mesmo acesso disponibilizado no sítio de Governo Aberto SP.

208. Desde 2013, o Estado de Pernambuco possui um sítio de dados abertos e disponibiliza 33 conjuntos de dados nos formatos xml e csv (peça 49). O Distrito Federal, em seu sítio dados.abertos.df.gov.br, lançado em abril de 2014, disponibiliza dezesseis conjuntos de dados em formato csv.

209. O Rio Grande do Sul possui um portal de dados abertos (dados.rs.gov.br) que, no momento, disponibiliza setenta conjuntos de dados (datasets), muitos dos quais séries históricas, em sete áreas: educação, finanças, governo, segurança, trânsito, saúde, desenvolvimento, trabalho e renda. Os dados são disponibilizados principalmente no formato csv, mas também em outros formatos abertos.

210. Além do portal de dados abertos, o Estado mantém solução de Business Intelligence (BI) de acesso público e ferramenta de participação social, tais como o portal Transparência RS, o Mapa da Transparência e o Gabinete Digital. Criado em 2012, o Gabinete Digital do RS (gabinetedigital.rs.gov.br) é vinculado à Secretaria-Geral de Governo e busca estabelecer um canal de participação e diálogo entre governo e sociedade, por meio da incorporação de novas ferramentas que oferecem diferentes oportunidades ao cidadão de influenciar a gestão pública e exercer maior controle social sobre o Estado. Desde a sua criação, o governo já recebeu milhares de propostas e centenas de perguntas, que têm colaborado na priorização das ações do governo (gabinetedigital.rs.gov.br/resultados).

211. A título de exemplo da atuação de entidades estaduais provedoras de dados, pode-se citar a Fundação de Economia e Estatística (FEE) do Estado do Rio Grande do Sul, que possui um grande acervo de informações estaduais e as disponibiliza em vários formatos em sua página de dados abertos - dados.fee.tche.br.

212. Minas Gerais lançou em 2012 o Mapa de Resultados (mapaderesultados.mg.gov.br/index.html), que constitui uma ferramenta que fornece uma visão da evolução dos indicadores do Estado. Disponível no Portal Minas em Números (numeros.mg.gov.br), o mapa tem o objetivo de tornar públicos, em um mesmo ambiente, de forma didática e interativa, os principais indicadores do Governo de Minas, de seus municípios e regiões nos últimos dez anos, fornecendo subsídios para uma melhor tomada de decisão, provendo transparência e constante atualização das informações. Cumpre frisar que o portal Mapa de Resultados está indisponível no período de eleição.

213. Cabe observar que alguns dos estados supracitados apresentam suas informações financeiras em mapas construídos com ferramentas de Business Intelligence (BI), citando-se como exemplo os sítios mapa.rs.gov.br e mapaderesultados.mg.gov.br. Embora essas publicações promovam a transparência, permitindo o detalhamento do gasto até a especificação de um determinado favorecido, não há a disponibilização dos dados brutos em formatos abertos que possam ser reutilizados pelos cidadãos e outros entes de governo. Desse modo, as iniciativas não podem ser consideradas, estritamente, de dados abertos.

214. *Porto Alegre é outra cidade a ser citada pela adoção de uma política pública de dados abertos. Na busca por uma gestão qualificada, com incentivo à inovação e ao empreendedorismo resultando em novos negócios, a Prefeitura, dentre outras ações, criou o DataPoa, que constitui o portal de dados abertos de Porto Alegre e tem como objetivo convidar a comunidade a participar do desenvolvimento de soluções inteligentes para a cidade, criando laços colaborativos entre governo municipal, empresas, desenvolvedores e cidadãos.*

215. *Em julho de 2014, o portal declarava possuir 37 conjuntos de dados de onze diferentes grupos, tais como saúde, educação, água, clima, mobilidade, orçamento participativo e outros. Utilizando dados abertos sobre a saúde do município, foi criado o aplicativo Doctor Poa, que permite ao usuário localizar as unidades de atendimento e hospitais mais próximos a uma determinada localidade, além de uma busca por emergências em determinadas especialidades. O aplicativo é descrito no parágrafo 224.*

216. *Em relação a estudos nacionais sobre o nível de abertura de dados nos Estados, esse trabalho identificou o Índice de Transparência, que constitui em um indicador calculado a partir da análise dos portais de transparência dos governos estaduais realizada pela Associação Contas Abertas. Essa avaliação elenca sítios com menor ou maior grau de transparência, atribuindo notas de zero a dez com base em cem parâmetros, divididos em três grandes temas: Conteúdo (55% da nota final), Usabilidade (40%) e Série Histórica e Frequência de Atualização (5%). A terceira edição da avaliação, publicada em 30/5/2014, mostrou os estados do Espírito Santo (8,96), Pernambuco (8,14) e São Paulo (7,95) na liderança (indexdetransparencia.com/2014/05/30/es-pe-e-sp-sao-os-estados-mais-transparentes-do-brasil/).*

4.1.5 Iniciativas de dados abertos em Tribunais de Contas

217. *Quanto à atuação dos Tribunais de Contas em relação ao tema dados abertos, a título de exemplo, cita-se o Tribunal de Contas do Estado de Pernambuco (TCE-PE), que desenvolveu seu Portal do Cidadão com o intuito de democratizar o acesso às informações sobre a gestão governamental. No portal, há uma seção de dados abertos (cidadao.tce.pe.gov.br/portalcidadao/index.php/dados-abertos) em que são disponibilizados arquivos sobre execução orçamentária e financeira, e licitações e contratos, no formato aberto xml e no formato proprietário bak (arquivos de backup do banco de dados SQL Server). Por meio de sistema de Business Intelligence (BI), o Portal do Cidadão também apresenta os indicadores dos municípios do Estado em áreas como saúde e educação, com expectativa de abranger outras áreas, no futuro.*

218. *Nesse ambiente, para um determinado município selecionado, o usuário pode acessar graficamente análises comparativas dos municípios nas áreas de educação e saúde, e a evolução de indicadores como a taxa de mortalidade, despesa com saúde por habitante, quantidade de médicos por mil habitantes e indicadores ambientais, dentre outros. Também estão consolidadas no portal as informações sobre a execução orçamentária dos municípios do Estado de Pernambuco e sobre a evolução das principais obras do Estado.*

4.1.6 Soluções criadas pela sociedade que utilizam dados governamentais abertos

219. *Durante a fase de execução deste levantamento, a equipe de fiscalização realizou pesquisa e selecionou algumas soluções ou aplicativos criados pela sociedade com o intuito de ilustrar o que tem sido feito no país com o uso de dados abertos. Alguns exemplos são apresentados a seguir.*

Fi-Guardian – <http://www.fi-ware.org/2014/03/19/fi-guardian-smart-monitor-for-adverse-events-winners-of-smart-city-challenge-at-campus-party-brazil-2014>

220. *A solução criada por desenvolvedores de Nova Friburgo, região serrana do Rio de Janeiro, tem o objetivo de emitir alertas de eventos climáticos adversos com o intuito de evitar catástrofes. Um protótipo foi desenvolvido no âmbito do Campus Party 2014, tendo, inclusive, ganhado o maior prêmio na categoria Cidades Inteligentes.*

221. *O objetivo dos desenvolvedores é evoluir o Fi-Guardian para um aplicativo robusto, aperfeiçoando o monitoramento com a integração de dados de várias fontes oficiais utilizados pelo governo para gerenciar os desastres naturais, permitindo às cidades que incorporem sensores locais à solução e definam novas regras de detecção de eventos. Outros objetivos são ampliar e diversificar os canais de comunicação com o cidadão, além da promoção da interoperabilidade com outros sistemas, como exemplificado pela integração do Fi-Guardian com o sistema de alerta SMS da cidade de Nova Friburgo (alertaviacelular.com.br).*

Política Aberta – <http://www.politicaaberta.org/>

222. *O Política Aberta é um aplicativo que apresenta, de um lado, a relação dos maiores doadores de campanha do ano de 2012, do outro, a relação das empresas mais contratadas do governo federal no mesmo ano. Com o clique sobre as informações da empresa ou doador, o usuário obtém o detalhamento dos valores referentes às doações feitas e pagamentos recebidos.*

223. *O Política Aberta foi desenvolvido por Gustavo H. M. Oliveira, como parte de um projeto junto ao Edmond J. Safra Center for Ethics da Universidade de Harvard, e foi inspirado na solução Chicago Lobbyist (parágrafo 162). Todos os dados são provenientes do <http://www.portaltransparencia.gov.br/> e do Tribunal Superior Eleitoral.*

Doctor Poa – <http://www.absoluta.net/doctor-poa-pt>

224. *Criado a partir de dados do DataPoa, portal de dados abertos da cidade de Porto Alegre (parágrafos 214 e 215), o aplicativo Doctor Poa foi idealizado por um médico para atender cidadãos porto-alegrenses e turistas. Ele permite ao usuário localizar as unidades de atendimento e hospitais mais próximos, com base em sua posição geolocalizada, além de uma busca por emergências em determinadas especialidades, como traumatologia, cardiologia etc, e por unidades públicas ou particular.*

225. *A solução utiliza a base de instituições médicas da cidade, que está disponível para download, e fornece vários dos dados sobre elas: tipo de instituição; tipo de emergência; nome; endereço; telefone; latitude; longitude; endereço da página web da instituição; e especialidades. O exemplo prova que é viável reunir essas informações sobre a rede médica e, por óbvio, a criação de soluções como essa, se estendidas ao âmbito nacional, traria informação e agilidade a um número maior de usuários do serviço de saúde pública.*

Escola Que Queremos – <http://www.escolaquequeremos.org/>

226. *O principal indicador levado em conta no debate educacional para se falar de qualidade do ensino é o Índice de Desenvolvimento da Educação Básica (Ideb). Para muitos, apesar de ser uma referência importante, o Ideb ainda é insuficiente: abrange apenas o desempenho dos alunos em português e matemática e o rendimento escolar (taxa de aprovação), sem trazer dados de contexto que permitam entender as causas em eventuais falhas no aprendizado.*

227. *A versão piloto da solução Escola Que Queremos se propõe a apresentar a avaliação de escolas públicas considerando outras dimensões além do Ideb, tais como o Censo Escolar e a Prova Brasil, infraestrutura, valorização dos professores, insumos e*

gestão democrática. A solução apresenta vinte indicadores em cinco dimensões diferentes, para que cada um possa selecioná-los e compor seu próprio índice de qualidade, comparando-o com a média do município. A solução utiliza dados disponibilizados pelo Inep.

QEdu – <http://www.qedu.org.br>

228. A solução QEdu constitui um portal aberto e gratuito, que apresenta informações sobre a qualidade do aprendizado dos alunos do 5º e 9º anos em cada escola, município e estado do Brasil. O portal mostra informações sobre o aprendizado em matemática e português; o perfil dos alunos, professores e diretores; o número de matrículas; taxas de aprovação, abandono e reprovação; informações sobre infraestrutura escolar; e alguns indicadores sobre aprendizado dos alunos e qualidade de ensino. As fontes de dados do sítio são os dados da Prova Brasil, do Censo Escolar e do Ideb, com detalhamento das informações por estado, cidade e escola, todos disponíveis no portal do Inep. Os alunos participantes da Prova Brasil, seus professores e o diretor da escola respondem a questionários sobre o perfil, práticas de estudo, práticas de ensino, percepções e ocorrências na escola. Essas informações também são apresentadas no portal.

229. Conforme a opção do usuário, são mostradas diferentes informações sobre as escolas, como, por exemplo, comparações de seus índices com os índices estaduais ou nacionais, evolução do desempenho da escola nos últimos anos, ou mesmo o detalhamento completo dos dados sobre a escola. A título de exemplo, o detalhamento da escola 01 do Planalto (código Inep 53000854) é reproduzido nas telas abaixo.

[Sobre](#)
[Destaques](#)
[Ajuda](#)

[Brasil](#) > [Distrito Federal](#) > [Brasília](#)
[Acompanhar](#)

01 Do Planalto > CEF estadual

[✓ Aprendizado](#)
[Compare](#)
[Evolução](#)
[Proficiência](#)
[Explore](#)
[Contexto](#)
[Pessoas](#)
[Censo](#)

Contexto da educação na escola - Rede Estadual

Confira alguns números e informações que ajudam a entender a qualidade da Educação e o contexto

Além da Prova Brasil o **Censo Escolar** é uma importante fonte de dados sobre a educação brasileira

Aprendizado dos alunos em 2011

Porcentagem de alunos com aprendizado adequado



Ideb em 2011 Novo!

Índice de Desenvolvimento da Educação Básica

Anos iniciais	4,3	=	0,84 (fluxo)	×	5,16 (aprendizado)
Anos finais	4,8	=	0,92 (fluxo)	×	5,20 (aprendizado)

Para ver o Ideb deste município em detalhes, visite o Portal Ideb



Matrículas em 2013



Anos iniciais (1ª a 4ª série ou 1º ao 5º ano)	281
Anos finais (5ª a 8ª série ou 6º ao 9º ano)	258

Informações Gerais

Código INEP	53000854
Localização	Urbana
Dependência	Estadual
Endereço	ACAMPAMENTO PACHECO FERNANDES - AE S/N Bairro: VILA PLANALTO CEP: 70802140
Telefone	(61) 3901-8207

Infraestrutura (dependências)



Existe sanitário dentro do prédio da escola?	Sim
Existe sanitário fora do prédio da escola?	Não
A escola possui biblioteca?	Sim
A escola possui cozinha?	Sim
A escola possui laboratório de informática?	Sim
A escola possui laboratório de ciências?	Não
A escola possui sala de leitura?	Não

Nível socioeconômico (NSE)

NSE não calculado

O nível socioeconômico sintetiza o perfil econômico e social dos estudantes de uma escola. Em geral deve-se comparar escolas cujo perfil dos alunos é semelhante.

[Saiba mais sobre o NSE ▶](#)

Figura 13 – Consulta à escola 01 Do Planalto (Fonte: QEdu.org.br)

Infraestrutura (dependências)



Existe sanitário dentro do prédio da escola?	Sim
Existe sanitário fora do prédio da escola?	Não
A escola possui biblioteca?	Sim
A escola possui cozinha?	Sim
A escola possui laboratório de informática?	Sim
A escola possui laboratório de ciências?	Não
A escola possui sala de leitura?	Não
A escola possui quadra de esportes?	Sim
A escola possui sala para a diretoria?	Sim
A escola possui sala para os professores?	Sim
A escola possui sala de atendimento especial?	Sim

Computadores e Internet



Internet	Sim
Banda larga	Sim
Computadores para uso dos alunos	730
Computadores para uso administrativo	5

Equipamentos da Escola



Aparelho de DVD	Sim
Impressora	Sim
Copiadora	Sim
Retroprojektor	Sim
Televisão	Sim

Telefone	(61) 3901-8207
----------	----------------

Nível socioeconômico (NSE)

NSE não calculado

O nível socioeconômico sintetiza o perfil econômico e social dos estudantes de uma escola. Em geral deve-se comparar escolas cujo perfil dos alunos é semelhante.

[Saiba mais sobre o NSE](#)

Perfil dos alunos de 5º ano em 2011

Foram respondidos 61 (81%) de um total de 75 questionários aplicados.



20% informam que a mãe ou a mulher responsável completou a 8ª série, mas não completou o ensino médio (antigo 2º grau)

80% entraram na escola na creche ou na pré-escola

Incentivo aos estudos

[mais](#)

Fonte: Prova Brasil/INEP 2011. QEdu.org.br

Perfil dos alunos de 9º ano em 2011

Foram respondidos 24 (86%) de um total de 28 questionários aplicados.



18% informam que a mãe ou a mulher responsável completou a faculdade

96% entraram na escola na creche ou na pré-escola

Incentivo aos estudos

[mais](#)

Fonte: Prova Brasil/INEP 2011. QEdu.org.br

Saneamento Básico



Abastecimento de água	Rede pública
Abastecimento de energia	Rede pública
Destino do esgoto	Rede pública
Destino do Lixo	Coleta periódica

Figura 14 - Consulta à escola 01 Do Planalto (Fonte: QEdu.org.br)

4.2 Aspectos gerais da legislação brasileira correlata a acesso à informação, dados abertos, governo aberto e participação social

230. A Constituição Federal, em seu Artigo 5º, inciso XXXIII, estabeleceu que: 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 (grifos nossos).

231. A Constituição Federal, no inciso II do § 3º de seu art. 37, ao abordar a necessidade de disciplinamento das formas de participação popular na administração

pública, enfatizou o acesso da população a registros administrativos e a informações sobre os atos de governo.

§ 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;

232. Quanto ao direito de acesso à informação, a Lei 12.527/2011, Lei de Acesso à Informação (LAI), foi publicada com o intuito de garantir o cumprimento dos incisos citados da Constituição Federal, devendo os procedimentos ali dispostos serem observados pela União, Estados, Distrito Federal e Municípios. Nesse sentido, a LAI dispõe, a uma, sobre a obrigação do Estado de prover transparência passiva, ou seja, a obrigação de responder às demandas por informação da sociedade, e, a duas, prevê novas obrigações de transparência ativa para os órgãos e entidades públicas, tais como as relativas aos chamados dados governamentais abertos.

233. Vale frisar que algumas obrigações de transparência ativa já haviam sido objeto de regulamentação inicial desde a vigência do Decreto 5.482/2005 e a Portaria Interministerial 140/2006, segundo os quais os órgãos e entidades da Administração Pública Federal, direta e indireta, devem manter em seus respectivos sítios eletrônicos página denominada “Transparência Pública”, exclusivamente para fins de controle social, com a divulgação de dados e informações relativas à sua execução orçamentária e financeira, compreendendo, entre outras, matérias relativas a licitações, contratos, convênios, despesas com passagens e diárias, além de outros conteúdos que vierem a ser estabelecidos. Tais normativos, contudo, ainda não abordavam requisitos específicos de divulgação de dados públicos contemplados na política de dados abertos.

234. Em relação aos dados abertos, o art. 8º da Lei 12.527/2011 e seus § 2º e 3º estabelecem que:

Art. 8º É dever dos órgãos e entidades públicas promover, independentemente de requerimentos, a divulgação em local de fácil acesso, no âmbito de suas competências, de informações de interesse coletivo ou geral por eles produzidas ou custodiadas. (...)

§ 2º Para cumprimento do disposto no caput, os órgãos e entidades públicas deverão utilizar todos os meios e instrumentos legítimos de que dispuserem, sendo obrigatória a divulgação em sítios oficiais da rede mundial de computadores (internet).

§ 3º Os sítios de que trata o § 2º deverão, na forma de regulamento, atender, entre outros, aos seguintes requisitos:

I - conter ferramenta de pesquisa de conteúdo que permita o acesso à informação de forma objetiva, transparente, clara e em linguagem de fácil compreensão;

II - possibilitar a gravação de relatórios em diversos formatos eletrônicos, inclusive abertos e não proprietários, tais como planilhas e texto, de modo a facilitar a análise das informações;

III - possibilitar o acesso automatizado por sistemas externos em formatos abertos, estruturados e legíveis por máquina;

IV - divulgar em detalhes os formatos utilizados para estruturação da informação;

V - garantir a autenticidade e a integridade das informações disponíveis para acesso;

VI - manter atualizadas as informações disponíveis para acesso;

VII - indicar local e instruções que permitam ao interessado comunicar-se, por via eletrônica ou telefônica, com o órgão ou entidade detentora do sítio; e

VIII - adotar as medidas necessárias para garantir a acessibilidade de conteúdo para pessoas com deficiência, nos termos do art. 17 da Lei 10.098, de 19 de dezembro de 2000, e do art. 9º da Convenção sobre os Direitos das Pessoas com Deficiência, aprovada pelo Decreto Legislativo 186, de 9 de julho de 2008. (grifos nossos)

235. *Sobre o acesso a informações pessoais, a LAI, em seu art. 31, aborda a questão, consignando que:*

Art. 31. O tratamento das informações pessoais deve ser feito de forma transparente e com respeito à intimidade, vida privada, honra e imagem das pessoas, bem como às liberdades e garantias individuais.

236. *Contudo, nem a LAI nem o Decreto 7.724/2012 define o rol de informações consideradas pessoais. Nesse sentido, tramita na Câmara dos Deputados projeto de lei para regulamentar o conceito e uso dos dados pessoais. Trata-se do Projeto de Lei 4.060/2012, de autoria do Deputado Milton Monti (PR-SP), que “dispõe sobre o tratamento de dados pessoais, e dá outras providências”. O projeto atualmente aguarda Parecer do Relator na Comissão de Ciência e Tecnologia, Comunicação e Informática (CCTCI). No Projeto de Lei 4.060/2012, considera-se dado pessoal “qualquer informação que permita a identificação exata e precisa de uma pessoa determinada”.*

237. *Como normativo infralegal mais detalhado e específico em relação a dados abertos, tem-se a Instrução Normativa SLTI/MP 4, de 18 de abril de 2012, que instituiu a Infraestrutura Nacional de Dados Abertos (Inda). Consoante, o art. 1º da IN – SLTI/MP 4/2012, a Inda é “uma política para garantir e facilitar 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”. A IN – SLTI/MP 4/2012 estabelece como membros da Inda, obrigatoriamente, todos os órgãos do Sisp e facultativamente, mediante a assinatura do termo de adesão, os demais órgãos e entidades dos Poderes Executivo, Legislativo e Judiciário, das esferas Federal, Estadual, Distrital e Municipal.*

238. *Segundo o disposto no art. 1º da IN – SLTI/MP 4/2012, os objetivos da Inda são:*

I – definir, estruturar e coordenar a política de dados abertos, bem como estabelecer o seu modelo de funcionamento;

II – promover o ordenamento na geração, armazenamento, acesso, e compartilhamento de dados para uso do Poder Executivo federal e da sociedade;

III- 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;

IV – promover 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;

V – apoiar, capacitar e fornecer suporte para a publicação de dados abertos aos órgãos e entidades do Poder Executivo federal ou que aderirem à INDA que não possuem prática, cultura e atribuições finalísticas de disseminação de dados;

VI – buscar a melhoria contínua da publicação de dados abertos, baseando-se nas melhores práticas concebidas nos cenários nacional e internacional;

VII – 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;

VIII – promover e apoiar o desenvolvimento da cultura da publicidade de dados e informações na gestão pública;

IX – disponibilizar tecnologias e apoiar as ações dos órgãos e entidades do Poder Executivo federal ou que aderirem à INDA na implementação da transparência ativa por meios digitais; e

X – promover a participação social na construção de um ecossistema de reuso e de agregação de valor dos dados públicos (grifos nossos).

239. *Consoante o art. 5º da IN – SLTI/MP 4/2012, a gestão da Inda cabe a um comitê gestor, presidido pela SLTI/MP e composto por representantes da sociedade civil, do setor acadêmico e de nove entes federais, quais sejam, Ministério do Planejamento, Orçamento e Gestão (MP), Casa Civil da Presidência da República, CGU, Ministério da Ciência,*

Tecnologia e Inovação (MCTI), Ministério do Desenvolvimento Social (MDS), Ministério da Educação (MEC), Ministério da Saúde (MS), Secretaria-Geral da Presidência (SGPR) e Instituto Brasileiro de Geografia e Estatística (IBGE).

240. *Ao comitê gestor da Inda, cabe (art. 6): i) estabelecer diretrizes para o desenvolvimento, implantação, manutenção e gestão da evolução do Portal Brasileiro de Dados Abertos; ii) definir o modelo de licença para os dados abertos; iii) priorizar e recomendar aos órgãos e entidades quanto à abertura dos dados e informações; e iv) elaborar e monitorar o plano de ação para a implantação e gestão da Inda.*

241. *Já em relação à legislação relativa ao governo aberto, o Decreto s/n de 15 de setembro de 2011 instituiu o plano de ação nacional, por meio do qual o Brasil, como um dos países que celebraram a Parceria para Governo Aberto (OGP), tendo sido co-líder nos anos de 2011 e 2012, assumiu compromissos perante essa parceria (peça 6, p. 2).*

242. *De acordo com o art. 1º, o plano de ação nacional sobre governo aberto destina-se a promover ações e medidas que visem ao incremento da transparência e do acesso à informação pública, à melhoria na prestação de serviços públicos e ao fortalecimento da integridade pública.*

243. *Também, por meio do Decreto s/n de 15 de setembro de 2011, em seu art. 3º, o Governo Federal criou o Comitê Interministerial de Governo Aberto (Ciga). O Ciga é a instância decisória responsável por orientar a implementação e a elaboração dos planos de ação de governo aberto do Brasil que deverão ser, segundo as regras da OGP, internacionais e bianuais. O comitê é composto por dezoito ministérios, representados por seus respectivos ministros, e coordenado pela Casa Civil da Presidência da República.*

244. *Internamente ao Ciga existe um Grupo Executivo (GE Ciga), coordenado pela Controladoria-Geral da União, que funciona como braço executor das políticas, diretrizes e ações definidas pelo Ciga. Além da CGU, integram o GE Ciga: Secretaria-Geral da Presidência da República, Casa Civil da Presidência da República, Ministério do Planejamento, Ministério da Justiça, Ministério da Fazenda e Ministério das Relações Exteriores (peça 51, p. 5).*

245. *O Ciga possui as seguintes atribuições, conforme o art. 4º do Decreto s/n de 15 de setembro de 2011:*

- a) promover a difusão do plano de ação nacional sobre governo aberto junto à sociedade e a articulação intragovernamental para a sua implementação e execução;*
- b) identificar os meios necessários à elaboração, implementação e monitoramento do plano de ação nacional sobre governo aberto;*
- c) apreciar as propostas que lhe forem submetidas por seu Grupo Executivo;*
- d) avaliar os resultados e propor alterações ou revisões ao plano de ação nacional sobre governo aberto; e*
- e) aprovar parâmetros, métodos e práticas para implementação, coordenação, execução e avaliação do plano de ação nacional sobre governo aberto, quando necessário.*

246. *O Decreto s/n, de 15 de setembro de 2011, art. 1º, incisos II e III, também destaca o dever da gestão pública na promoção da participação social, quando estabelece como diretrizes do plano de ação nacional sobre governo aberto, entre outras: a) o fomento à participação social nos processos decisórios e; b) o estímulo ao uso de novas tecnologias na gestão e prestação de serviços públicos, que devem fomentar a inovação, fortalecer a governança pública e aumentar a transparência e a participação social.*

247. *Igualmente, a IN – SLTI/MP 4/2012 apresenta, como um de seus objetivos (art. 1º, inciso X) a promoção da participação social na construção de um ecossistema de reuso e de agregação de valor dos dados abertos. Inserida nesse espírito, a própria norma estabelece que os cidadãos e entidades da sociedade civil interessados nas atividades da Inda poderão compor a sua estrutura e participar de sua implementação (art. 3º, § 1º e 2º).*

248. *Recentemente, em maio de 2014, a Casa Civil instituiu a Política Nacional de Participação Social – PNPS (Decreto 8.243/2014), com o objetivo de fortalecer e articular os mecanismos e as instâncias democráticas de diálogo e a atuação conjunta entre a Administração Pública Federal e a sociedade civil. No contexto deste trabalho, cabe destaque ao inciso VII do art. 18 do Decreto 8.243/2014, o qual estabelece que, na criação de ambientes virtuais de participação social, deve ser priorizada a exportação de dados em formatos abertos e legíveis por máquinas.*

249. *Por fim, entende-se necessário destacar que o tratamento das informações sigilosas e pessoais é vital para a definição de quais são os dados públicos de interesse coletivo ou geral.*

250. *Tem-se, portanto, que a LAI os demais normativos aqui mencionados constituem instrumentos importantes no sentido de permitir e impor a realização de ações na direção da transparência e da abertura das informações governamentais. Contudo, e considerando o caráter inovador do tema, há assuntos ainda não normatizados, como o tratamento das informações sigilosas e pessoais, aspecto fundamental para a definição de quais são os dados públicos de interesse coletivo ou geral.*

4.3 Atuação dos órgãos nacionais gestores de dados abertos, governo aberto e implementação e gestão da LAI

251. *As principais instituições públicas responsáveis pelas iniciativas de dados abertos, governo aberto e implementação da Lei de Acesso à Informação são: o Ministério do Planejamento, Orçamento e Gestão (MP), a Controladoria-Geral da União (CGU), a Casa Civil da Presidência da República e a Secretaria-Geral da Presidência da República. Além desses órgãos, foram instituídos órgãos colegiados destinados à gestão desses temas: o Comitê Gestor da Infraestrutura Nacional de Dados Abertos e o Comitê Interministerial de Governo Aberto (Ciga). A seguir, serão enumeradas as competências de cada uma dessas instituições nos temas levantados.*

Casa Civil da Presidência da República

252. *A Casa Civil da Presidência da República é um ator estratégico nos temas de governo aberto e de implementação e gestão da LAI.*

253. *No tocante a governo aberto, conforme o disposto no art. 5º do Decreto s/n de 15 de setembro de 2011, a Casa Civil da Presidência da República é o órgão coordenador do Ciga.*

254. *Também compete à Casa Civil, por força do disposto no Decreto 7.724/2012, art. 68, inciso VII, definir, em conjunto com a Controladoria-Geral da União, diretrizes e procedimentos complementares necessários à implementação da LAI. Ainda sobre esse tema, a Casa Civil exerce a função de Secretaria-Executiva da Comissão Mista de Reavaliação de Informações, consoante o Decreto 7.724/2012, art. 46, inciso I c/c art. 53, parágrafo único.*

Ministério do Planejamento, Orçamento e Gestão

255. *A unidade do MP responsável pelos temas citados nesta seção é o Departamento de Governo Eletrônico (DGE) da Secretaria de Logística e Tecnologia da Informação (peça 6, p. 4), conforme disposto no Decreto 8.189/2014, normativo que aprova a estrutura regimental do ministério.*

256. *À SLTI compete propor políticas, planejar, coordenar, supervisionar e orientar normativamente as atividades de governo eletrônico relacionadas a padronização e à disponibilização de serviços eletrônicos interoperáveis, acessibilidade digital e abertura de dados (art. 34, inciso I, alínea “d”).*

257. *Mais especificamente, consoante o art. 38, ao DGE foram designadas as seguintes atribuições:*

I - coordenar, disciplinar e articular a implantação de ações integradas de governo eletrônico na administração pública federal;

II - promover e coordenar atividades relacionadas à prestação de serviços públicos por meios eletrônicos na administração pública federal;

III - definir, publicar e disseminar padrões e normas em governo eletrônico referentes a acessibilidade, interoperabilidade, serviços e conteúdos públicos digitais e coordenar a sua implementação;

IV - promover a transparência e a participação da sociedade através dos meios digitais, quanto a consultas públicas e abertura de dados; e

V - sistematizar e disseminar informações relacionadas às ações de governo eletrônico da administração pública federal.

258. *Em cumprimento ao art. 4º da IN – SLTI/MP 4/2012, a SLTI/MP mantém o Portal Brasileiro de Dados Abertos (dados.gov.br), concebido para ser um sítio eletrônico de referência para a busca e acesso aos dados públicos, seus metadados, informações, aplicativos e serviços relacionados. Maiores detalhes sobre a plataforma e gestão do portal constam na seção 4.4 deste relatório.*

259. *Além disso, a SLTI/MP preside o Comitê Gestor da Inda. Nesse sentido, a SLTI/MP vem adotando iniciativas, das quais destacamos a construção do manual para elaboração de planos de dados abertos (peça 38). O manual para elaboração de planos de dados abertos tem a finalidade de apoiar os órgãos da APF no planejamento e padronização das ações de publicação de dados abertos do governo brasileiro. Baseado neste manual, o MP publicou seu próprio plano de dados abertos (peça 39).*

260. *Merece consideração iniciativa constante no plano de ação da Inda atinente à publicação de um modelo de maturidade de dados abertos do governo brasileiro que, à época da conclusão deste trabalho, estava previsto para outubro de 2014. Segundo o plano de ação, esse modelo pretende orientar as instituições detentoras de dados públicos no processo de referenciar seus dados publicados e pretende classificar essas instituições de acordo com o nível de qualidade de publicação dos seus dados, visando aumentar o engajamento e interação com a sociedade civil (peça 50, p. 31-32).*

261. *Outro ponto de atuação da SLTI/MP no tocante a dados abertos refere-se à definição quanto à necessidade e à viabilidade de adoção ou criação de um modelo de licenças para dados abertos no Brasil. Trata-se de um modelo que defina as regras de como a informação pode ser usada, reusada e distribuída. A definição em relação à adoção ou não de um modelo de licenças constitui meta incluída no plano de ação da Inda para os anos de 2013 e 2014. Entretanto, o resultado das ações da SLTI se resumem, até então, ao estabelecimento de acordo de cooperação com a Universidade de Brasília para realização de estudo sobre o modelo de licenças (peça 6, p. 8-9).*

262. *No que concerne ao governo aberto, o MP, representado pelo DGE/SLTI, participa no Grupo Executivo do Comitê Interministerial de Governo Aberto (GE Ciga).*

263. *Com relação ao monitoramento da LAI, o art. 69 do Decreto 7.724/2012 atribuiu ao MP, em conjunto com a CGU, responsabilidade pelo estabelecimento de procedimentos, regras e padrões de divulgação de informações ao público e pelo detalhamento dos procedimentos necessários à busca, estruturação e prestação de informações no âmbito do Sistema de Informação ao Cidadão (SIC). Quanto a isso, o MP informou que, embora não tenha sido estabelecido ato formal conjunto entre o ministério e a CGU, tem adotado iniciativas que podem ser consideradas como parte de um processo de amadurecimento relacionado à implementação da LAI e decorrentes das necessidades de disseminação de dados e informações, tais como: i) a publicação do manual para elaboração de planos de dados abertos; ii) a construção do guia de publicação do rol de informações classificadas e desclassificadas; e iii) a construção do guia para criação da seção “Acesso à Informação” nos sítios dos órgãos e entidades federais, dentre outros (peça 6, p. 2).*

Controladoria-Geral da União (CGU)

264. *A CGU atua representando o Brasil na Parceria de Governo Aberto (OGP). Isso porque a CGU coordena o GE Ciga, tendo, dentre outras funções, a de coordenar a implementação e a execução do plano de ação nacional sobre governo aberto (peça 51, p. 5).*

265. *Na CGU, o envolvimento com o tema governo aberto se dá em diferentes níveis: Ministro, Secretaria-Executiva, Assessoria Internacional, Secretaria de Transparência e Prevenção da Corrupção (STPC/CGU), Diretoria de Transparência e Controle Social e Coordenação-Geral de Transparência e Governo Aberto.*

266. *Com relação à transparência e à implementação da LAI, a referida lei, art. 41, incisos I, II e IV, c/c Decreto 7.724/2012, art. 68, incisos II, III e V, estabelece que compete à CGU:*

Decreto 7.724/2012, art. 68

(...)

II - promover campanha de abrangência nacional de fomento à cultura da transparência na administração pública e conscientização sobre o direito fundamental de acesso à informação;

III - promover o treinamento dos agentes públicos e, no que couber, a capacitação das entidades privadas sem fins lucrativos, no que se refere ao desenvolvimento de práticas relacionadas à transparência na administração pública;

(...)

V - preparar relatório anual com informações referentes à implementação da Lei 12.527, de 2011, a ser encaminhado ao Congresso Nacional;

267. *Finalizando, vale comentar que com relação à sua atuação nos termos do art. 69 do Decreto 7.724/2012, a CGU informou que vem tomando as seguintes providências (peça 51):*

a) disponibilização, manutenção e atualização do Portal da Lei de Acesso à Informação (www.acessoainformacao.gov.br);

b) em conjunto com MP, Casa Civil, Secretaria de Comunicação e Gabinete de Segurança Institucional da Presidência da República, elaboração de guias para orientação aos órgãos em relação à seção de acesso à informação, e divulgação de informações classificadas e desclassificadas;

c) desenvolvimento e gestão do Sistema Eletrônico do Serviço de Informações ao Cidadão (e-SIC), com o intuito de uniformizar procedimentos para acesso à informação;

- d) elaboração do manual do e-SIC;
- e) criação da rede de SICs (em fase de planejamento).

268. Além disso, vale comentar sobre o Portal da Transparência do Governo Federal (portaldatransparencia.gov.br), uma iniciativa da CGU lançada em novembro de 2004 para aumentar a transparência da gestão pública, com vistas a permitir que o cidadão acompanhe como o dinheiro público está sendo utilizado e exerça o controle social. Regulamentado pelo Decreto 5.482/2005, o portal apresenta informações sobre os gastos diretos do Governo Federal, as transferências de recursos federais, os beneficiários de programas sociais, as receitas previstas, os convênios celebrados, o rol de empresas sancionadas, os servidores públicos, o controle social, a relação de órgãos e entidades que disponibilizam página de transparência pública, dentre outras.

269. No menu “Download de Dados” (<http://portaldatransparencia.gov.br/downloads/>), são disponibilizados vários conjuntos de dados em formato aberto, permitindo ao usuário obter planilhas com os dados que são apresentados no portal.

Secretaria-Geral da Presidência da República

270. A Secretaria-Geral da Presidência da República (SG/PR) é também um ator importante na temática de dados abertos e principalmente na questão da participação social.

271. No que tange a dados abertos, a SG/PR tem representação no comitê gestor da Inda e uma de suas secretarias, a Secretaria Nacional de Articulação Social (SNAS/SG/PR), tem o poder de convidar um representante da sociedade civil para compor o comitê.

272. Além disso, o Departamento de Participação Social da SNAS/SG/PR é um ator importante na promoção da participação social, uma vez que, conforme o disposto no Decreto 7.688/2012, art. 16, a ele compete estruturar formas de consulta e participação social na gestão pública:

Decreto 7.688/2012

Art. 16. Ao Departamento de Participação Social compete:

I - propor a criação e a articulação de institutos de consulta e participação social na gestão pública;

273. Nessa esteira, o Departamento de Participação Social da SNAS/SG/PR mantém o Portal da Participação Social (participa.br), que oferece uma série de ferramentas para participação dos cidadãos, redes, movimentos sociais e organizações que querem dialogar com o Governo Federal na construção de políticas públicas. Frisa-se que o Participa.br encontra-se indisponível durante o período eleitoral de 2014.

4.4 Portal Brasileiro de Dados Abertos

274. O portal brasileiro de dados abertos (dados.gov.br), além de servir como um catálogo que contém os dados abertos produzidos ou custodiados pelos entes da administração pública, disponibiliza links para orientações técnicas sobre o tema, tais como a cartilha técnica para publicação de dados abertos no Brasil (peça 36), a cartilha técnica para desenvolvedores (peça 37) e o guia de dados abertos da OKF (opendatahandbook.org/pt_BR), destinados tanto ao público interno quanto à sociedade, especialmente os desenvolvedores de aplicações.

275. O próprio MP abriu parte dos dados públicos de um dos sistemas estruturantes da Administração Pública Federal, pelos quais é responsável, o Sistema de Convênios e

Contratos de Repasse da Administração Pública Federal (Siconv), e os catalogou no Portal Brasileiro de Dados Abertos, no endereço dados.gov.br/dataset/siconv.

276. *O portal brasileiro de dados abertos é baseado na plataforma de código aberto para portais de dados denominada CKAN (ckan.org/) da Open Knowledge Foundation (okfn.org). O CKAN é um padrão para portal de acesso a dados abertos, que possibilita acesso às informações por meio de uma API, com vistas a simplificar a automação e o reuso das informações por outras soluções tecnológicas.*

277. *Durante a fase de elaboração deste relatório, buscando avaliar a situação desse portal em momento em que a LAI e o próprio portal já contavam com mais de dois anos de vida, foi efetuada, em 21/7/2014, uma análise quantitativa e qualitativa do seu conteúdo. A análise foi feita a partir do catálogo disponível em dados.gov.br/dataset e do acesso a algumas das bases disponibilizadas no portal.*

278. *O catálogo do portal é dividido em conjuntos de dados, cada um contendo uma ou mais fontes de dados de temas inter-relacionados que podem ser individualmente acessadas. O portal denomina essas fontes de dados como recursos. Ou seja, o portal trata conjunto de dados como o agrupamento de recursos inter-relacionados. Cada conjunto de dados é identificado por um título, apresenta uma breve descrição do seu conteúdo e é indexado segundo algumas de suas propriedades, a saber: formato de seus arquivos, grupos temáticos e palavras chave (tags ou etiquetas). Os recursos podem ser disponibilizados em mais de um formato e identificados por mais de uma palavra chave. Um conjunto de dados recebe um selo “open data” se ele satisfaz a definição de dado aberto da organização The Open Definition (opendefinition.org) ou, em caso contrário, recebe o selo “Não abertamente licenciado”.*

279. *Para exemplificar, tomou-se, à época, o conjunto de dados Lista de Eleitores Filiados aos Partidos Políticos (dados.gov.br/dataset/filiados-partidos-politicos), pertencente ao grupo temático Governo e Política e com as seguintes etiquetas: TSE, eleitores, eleitores filiados, filiação partidária, partidos e partidos políticos. Esse conjunto continha 810 recursos, em formato csv, correspondentes aos eleitores filiados a cada um dos partidos políticos do Brasil em cada Unidade da Federação, como, por exemplo: Filiados ao PSC no Distrito Federal e Filiados ao PSDB em São Paulo.*

280. *No momento da análise, o catálogo apresentado pelo portal disponibilizava 55 conjuntos de dados e declarava que eles continham 2.135 recursos, distribuídos segundo suas propriedades nas várias etiquetas e grupos temáticos. Receberam o selo “open data” somente onze conjuntos de dados, o que representa 20% do total de conjuntos de dados.*

281. *Após alguns testes de busca no catálogo e algumas pesquisas feitas pelo sítio, foi possível identificar diversos conjuntos de dados que estavam de fato no portal, encontrados por caminhos de busca alternativos, mas que não puderam ser encontrados diretamente a partir da busca no catálogo. Citam-se como exemplos de recursos que constavam do sítio, mas não do catálogo, os seguintes: Cadastro Nacional de Entidades Sociais, que consta na página principal do portal, no elemento “Dados em destaque”, e que pode ser acessado no link dados.gov.br/dataset/cadastro-nacional-de-entidades-sociais; Unidades Básicas de Saúde (UBS), acessado pelo link dados.gov.br/dataset/unidades-basicas-de-saude-ubs; e Orçamento Federal, acessado pelo link dados.gov.br/dataset/orcamento-federal.*

282. *Essas discrepâncias levantam dúvidas com relação à integridade do catálogo disponibilizado e às estatísticas quantitativas por ele oferecidas, tal como o número de conjuntos de dados existentes no portal. Notou-se, por exemplo, que a totalização de*

conjuntos de dados por etiqueta na página principal mostrava valores diferentes daqueles obtidos ao acessar diretamente o catálogo. Diante disso, foi feito contato com a equipe gestora do portal na SLTI/MP, que reportou a existência de um problema de indexação que estaria sendo corrigido pelo Serpro, o hospedeiro do portal. Assim, de forma alternativa à realização dessa análise, a equipe gestora do Ministério sugeriu a utilização de um catálogo disponível no ambiente de desenvolvimento, cuja construção se baseou numa versão mais recente que aguardava homologação. Essa nova versão foi disponibilizada na internet no dia 1º/8/2014, contendo 239 conjuntos de dados e 2.135 recursos.

283. *No entanto, estima-se que uma parte considerável dos 239 conjuntos de dados catalogados na nova versão do portal - 96 conjuntos ou 40% do total - é simples referência a recursos disponíveis em outros sítios na internet. É o que ocorre com os conjuntos de dados sob a etiqueta Inde. Do total de 135 conjuntos de dados, 96 apenas redirecionam o usuário ao visualizador de mapas da Infraestrutura Nacional de Dados Espaciais (Inde; visualizador.inde.gov.br), e os demais 39 conjuntos de dados, além de fazer esse redirecionamento, oferecem recursos no formato csv.*

284. *Nessa nova versão do portal, constam com o selo “open data” somente 47 conjuntos de dados, considerando-se as licenças descritas como “Licença aberta para bases de dados do Open Data Commons”, “Domínio Público”, “Atribuição” e “Aberta”, o que representa 20% do total.*

285. *Quanto à usabilidade, a página inicial do portal, no endereço dados.gov.br, permite uma visualização rápida dos conjuntos mais acessados e os atualizados mais recentemente. Nesse painel, lê-se que o conjunto atualizado mais recentemente, o foi no dia dois do último junho, ou seja, sessenta dias antes dessa consulta. Essa página inicial parece ter como objetivo centralizar as buscas do interessado pelos dados abertos. Nela são apresentados, além das cinco publicações mais recentes, notícias e uma caixa de pesquisa textual.*

286. *O portal não oferecia, na versão anterior, meios fáceis de identificar a data da última atualização de um recurso disponibilizado ou o período de tempo ao qual ele corresponde. Na nova versão do portal, os recursos são melhor descritos e é possível uma pré-visualização do seu conteúdo, além da possibilidade de cópia. Boa parte dos recursos são arquivos que podem ser importados e, em alguns deles, é possível obter alguns detalhes do período considerado no próprio nome do arquivo. Em outros recursos, o nome da fonte não fornece qualquer informação útil em relação ao período ou a versão dos dados, exigindo a importação dos arquivos e análise do seu conteúdo para obter maiores esclarecimentos.*

287. *O que se observa é que o portal disponibiliza dados de relevância, tais como as bases do Inep relativas a desempenho escolar, avaliação da educação, etc.; dados do portal da transparência, o cadastro nacional de empresas inidôneas e suspensas, a relação dos servidores do poder executivo federal; diversas bases do IBGE; dados do TSE tais como a lista de eleitores filiados aos partidos políticos; dados do TCU, tais como a relação de licitantes inidôneas e dos inabilitados para função pública entre outros.*

288. *Em todos os casos citados acima, o portal republica dados existentes em portais dos próprios órgãos fornecedores das respectivas bases. Nesses casos, o portal de dados abertos serve de centralizador, disponibilizando dados de diversas fontes em um único local, o que facilita a busca realizada pelo cidadão.*

289. Quanto à atualização das informações, na maioria dos conjuntos de dados pesquisados, os dados disponibilizados no portal estão bastante defasados com relação aos fornecidos diretamente pelo órgão/entidade de origem. Essa defasagem tende a confundir o usuário e a obrigá-lo a recorrer, de qualquer forma, ao sítio internet do fornecedor para obter a informação mais atual. Como exemplo, o conjunto de dados intitulado *Despesas Diretas do Poder Executivo Federal* publicadas no Portal Transparência contém oito recursos, sendo o arquivo mais atual despesas referentes ao ano de 2011, enquanto no Portal da Transparência os dados são atualizados mensalmente e, no momento dessa análise, os mais atuais se referiam ao mês de maio de 2014.

290. Outro exemplo: no caso do conjunto de dados *Convênios do poder executivo federal*, publicados no Portal da Transparência, que deveria ter atualização semanal, não é possível verificar o período compreendido pelos dados a partir do nome do arquivo disponibilizado nem pelo detalhamento ou pré-visualização desse recurso, uma vez que o campo *Última Atualização* consta como desconhecido. Somente após a cópia do arquivo, verifica-se que se trata de uma versão que compreende convênios com data de publicação entre abril de 1989 e março de 2013. Porém, no portal da transparência, é possível obter a mesma base atualizada até 22/7/2014 (uma semana anterior à data da pesquisa).

291. Da mesma forma, nota-se que os recursos disponibilizados representam, muitas vezes, uma pequena fração do conteúdo informacional existente no órgão fornecedor. É o caso, por exemplo, do IBGE, que disponibiliza, no portal brasileiro de dados abertos somente treze conjuntos de dados de um conjunto de centenas de bases disponíveis no próprio sítio do IBGE na internet. E, considerando-se a lista das dez séries estatísticas mais acessadas no sítio do IBGE, (downloads.ibge.gov.br/downloads_top.php), somente duas estão disponibilizadas no portal brasileiro de dados abertos: uma relativa ao Índice de Preços ao Consumidor (IPCA), sendo que no sítio do IBGE é possível obter a série histórica até junho de 2014 e no portal de dados abertos a última série disponível vai de 2006 a 2011; e a outra relativa à malha geométrica dos municípios, sendo que no sítio do IBGE é possível obter a série de vários anos (2001, 2005, 2007, 2010 e 2013; downloads.ibge.gov.br/downloads_geociencias.htm) e no portal de dados abertos somente os dados referentes a 2010.

292. Outra questão relevante diz respeito à documentação dos recursos. Na maioria dos conjuntos de dados pesquisados nas duas versões do portal brasileiro de dados abertos, havia informação sumária descrevendo o conteúdo dos recursos e muito pouca descrição dos dados em si (metadados). Para compreender as informações disponibilizadas, é muitas vezes necessário acessar o sítio internet da instituição fornecedora dos dados e buscar as explicações necessárias à sua compreensão.

293. Ao que parece, não é papel da equipe gestora do portal brasileiro de dados abertos gerenciar a frequência de atualização dos dados catalogados pelo fornecedor, sua documentação, nem os critérios de seleção dos dados tornados abertos dentre as bases ou recursos que a instituição fornecedora detém. Os próprios fornecedores, ou seja, órgãos e entidades que aderem ao portal brasileiro de dados abertos, é que devem gerir os recursos publicados. Essa gestão tem sido heterogênea, com alguns conjuntos de dados atualizados e bem documentados no portal brasileiro de dados abertos, enquanto outros conjuntos de dados encontram-se bastante defasados e com documentação precária.

294. Dessa forma, conclui-se que o Portal Brasileiro de Dados Abertos, como ambiente centralizador do tema, disponibiliza dados da APF de relevância. Da breve análise do portal verificou-se as seguintes dificuldades (parágrafos 274 a 293): a maioria dos conjuntos de dados disponibilizados não recebem o selo open data por não

preencherem os requisitos inerentes ao conceito de dados abertos; a maioria dos conjuntos de dados estão desatualizados; os dados disponibilizados por organizações públicas nesse portal são, em sua maioria, uma fração dos dados disponibilizados em seus próprios sítios e sua gestão tem sido heterogênea, com alguns conjuntos de dados atualizados e bem documentados, outros bastante defasados e com documentação precária; e não se evidenciou a disponibilização de informações estatísticas acerca do uso e reuso dos dados abertos, tais como número de acesso ou de downloads realizados.

4.5 Atuação das empresas públicas prestadoras de serviços de TI em Dados Abertos e Big Data

295. *Nesta seção, descreve-se o cenário de atuação do Serviço de Processamento de Dados do Governo Federal (Serpro) e da Empresa de Tecnologia e Informações da Previdência Social (Dataprev) em relação ao tema de dados abertos e de tecnologias de Big Data. Essa descrição se justifica em virtude do fato de essas entidades serem prestadores de serviços de TI para instituições e, além disso, custodiantes de um grande volume de dados de interesse coletivo e geral produzidos pela Administração Pública Federal.*

4.5.1 Dataprev

296. *A Dataprev é uma empresa pública vinculada ao Ministério da Previdência Social (MPS), que tem por objetivo estudar e viabilizar tecnologias de informática, na área da previdência e assistência social, compreendendo prestação de serviços de desenvolvimento, processamento e tratamento de informações, atividades de teleprocessamento e comunicação de dados, voz e imagem, assessoramento e assistência técnica no campo de sua especialidade.*

297. *Por tratar-se de uma entidade destinada a custodiar e divulgar informações de alto valor para a gestão e o acompanhamento da previdência, assistência social, seguro desemprego e fiscalização do trabalho no Brasil, a Dataprev é um ator estratégico no que diz respeito a governo aberto e dados de interesse social.*

298. *São descritas, a seguir, as informações obtidas por meio da resposta ao Ofício de Requisição 3-137/2014 (peça 32), assim como as colhidas na reunião realizada em 24/4/2014 com representantes da Dataprev, dentre os quais o Sr. Raphael Pacheco, assessor de diretoria da presidência da empresa e responsável pelas iniciativas de dados abertos.*

Oferta de serviços de abertura de dados

299. *O serviço de publicação de dados abertos constitui um novo negócio para a Dataprev, que tem investido recursos com pesquisa e desenvolvimento na tecnologia e na infraestrutura necessária para a futura prestação desse serviço aos seus clientes (peça 32, p. 2-3).*

300. *A Dataprev elaborou minutas referentes ao modelo de negócio e ao modelo de comercialização de serviço de publicação de dados abertos (peças 33 e 34) que se encontram em fase de análise e cuja aprovação pela direção da empresa, de acordo com o Plano de Ação 2014, está prevista para o ano de 2014 (peça 32, p. 2).*

301. *Dessa forma, o serviço de publicação de dados abertos não se encontra no catálogo de serviços da empresa e, portanto, ainda não pode ser contratado pelos atuais e potenciais clientes da Dataprev (peça 32, p. 2).*

302. *No entanto, apesar da Dataprev ainda não ter sido demandada ou contratada por seus clientes para prestação do serviço de disponibilização de dados públicos em*

formato aberto, a empresa acredita ser de interesse do Governo Federal que esses dados, nos diversos órgãos, sejam publicados em formato aberto e, portanto, prepara-se para prover esse serviço (peça 32, p. 4).

Abertura de dados custodiados pela Dataprev

303. *A Dataprev realizou um evento com seus desenvolvedores internos, denominando-o de Hackathon interno. O evento utilizou dados custodiados pela Dataprev, mas sem participação direta de seus clientes. Segundo a empresa, os principais benefícios advindos do Hackathon interno foi, entre outros, a constatação da importância da qualidade e da relevância dos dados a serem publicados (peça 32, p. 2).*

304. *Além desse evento, a Dataprev publica, em caráter experimental, algumas bases de dados, das quais destacam-se o Anuário Estatístico da Previdência Social (Aeps), publicado em formatos abertos (json, xml e csv). Esse anuário corresponde às informações divulgadas pelo Ministério da Previdência Social (previdencia.gov.br/estatisticas), originalmente disponibilizados em arquivos pdf e tabelas xls (peça 32, p. 3). Os dados abertos do anuário apresentam-se publicados sob a forma de um conjunto de mais de cem conjuntos de dados (datasets) no portal de dados abertos da Dataprev (dadosabertos.dataprev.gov.br).*

305. *Ressalta-se que o Ministério da Previdência Social não participou da decisão de publicar dados do Anuário Estatístico da Previdência Social, no portal de dados abertos da Dataprev. Entretanto, houve ciência da intenção e autorização prévia para tal por parte da Coordenação Geral do Regime Geral de Previdência (peça 32, p. 4).*

306. *Na entrevista realizada em 24/4/2014, os representantes da Dataprev apontaram as seguintes dificuldades para a abertura de dados:*

306.1. *carência de definições únicas para termos correlatos empregados em sistemas diferentes, o que limita as possibilidades de integração desses dados (seria necessário reduzir as ambiguidades de vocabulário com a construção de uma ontologia);*

306.2. *problemas relativos à integridade e à confiabilidade dos dados; e*

306.3. *escassez de profissionais qualificados na área de ciência de dados (data science).*

Uso de tecnologias de Big Data pela Dataprev

307. *A Dataprev associou-se à Fundação Osvaldo Cruz/MS (FioCruz) para tornar disponíveis as bases de pesquisa na área de saúde realizadas com recursos públicos.*

308. *Foi firmado um convênio com a UFMG, em dezembro de 2013, para uma consultoria para realização de dois projetos pilotos em um projeto de dois anos, a título de investimento da empresa.*

309. *No primeiro projeto piloto com a Universidade, que se encontra em andamento, foi implantada uma infraestrutura para instalação de uma plataforma de processamento massivo de dados para execução de algoritmos de mineração de dados aplicados a uma questão específica de negócio, qual seja, os problemas que o INSS vem enfrentando para o agendamento das perícias médicas. Nesse projeto, a Dataprev realizou a extração dos dados e a UFMG está realizando a etapa de análise e caracterização dos dados.*

310. *Os principais benefícios esperados do convênio são materializados no conhecimento técnico a ser adquirido pelas equipes da empresa e na diminuição do risco tecnológico na escolha de uma nova solução relacionada ao tema (peça 32, p. 3).*

Conclusão

311. *Em face do exposto no levantamento realizado junto à Dataprev, é possível concluir que:*

a) a empresa custodia e difunde informações de alto valor social;

b) trata-se de uma instituição sensível à importância da política de dados abertos, que aderiu à Inda e ao portal brasileiro de dados abertos e que, antecipando-se a futuras demandas de seus clientes, desenvolveu um serviço para a abertura de dados de outros órgãos e entidades governamentais, além de ter concluído com sucesso a abertura dos dados do Anuário Estatístico da Previdência Social;

c) o modelo de negócio em relação ao novo serviço permanece incerto, visto que ainda não houve contratações desse serviço de abertura de dados pelos clientes da Dataprev;

d) como dificuldades para a abertura de dados, os representantes da empresa apontaram: carência de definições únicas para termos correlatos empregados em sistemas diferentes, o que limita as possibilidades de integração desses dados; problemas relativos à integridade e à confiabilidade dos dados; e escassez de profissionais qualificados na área de ciência de dados;

e) em relação a projetos de Big Data, a empresa ainda está iniciando projeto piloto de pesquisa, mediante convênio com a UFMG, com o intuito de adquirir conhecimento técnico e diminuir o risco tecnológico na escolha de uma eventual solução relacionada ao tema.

4.5.2 Serpro

312. *O Serpro é uma empresa pública vinculada ao Ministério da Fazenda (MF), que tem por negócio prestar serviços em tecnologia da informação e comunicações para o setor público.*

313. *Trata-se de uma entidade que se destina a desenvolver programas e serviços para permitir um maior controle e transparência sobre a receita e os gastos públicos, o que representa seu principal mercado de atuação. Além disso, desenvolve programas para ações estruturadoras da Administração Pública Federal, sob a gestão do Ministério do Planejamento, Orçamento e Gestão (MP), e programas que contemplam questões sociais de acessibilidade e inclusão digital. Desse modo, o Serpro também é um ator estratégico no que diz respeito a governo aberto e dados de interesse social.*

314. *São descritas e analisadas, a seguir, as informações obtidas por meio da resposta aos itens constantes no Ofício de Requisição 1-137/2014 (peça 12), assim como as colhidas na reunião realizada em 9/4/2014 na sede da empresa com seus representantes acerca dos temas dados abertos e Big Data.*

Serviços de manutenção do Portal Brasileiro de Dados Abertos (dados.gov.br)

315. *O Serpro hospeda o ambiente de hardware e software do Portal Brasileiro de Dados Abertos (dados.gov.br), tendo sua infraestrutura localizada no centro de dados do Serpro no Rio de Janeiro. O portal também é baseado na plataforma de código aberto para portais de dados CKAN, da OKFN (peça 24, p. 1).*

316. *À época dos trabalhos de campo, o Serpro informou que o Portal Brasileiro de Dados Abertos contava com 231 conjuntos de dados disponibilizados.*

Ausência de oferta de serviços de abertura de dados a seus clientes

317. *Entretanto, o Serpro manifestou que não há serviço específico e exclusivo para dispor dados públicos, de propriedade dos seus clientes, para publicação no Portal Brasileiro de Dados Abertos (peça 24, p. 2). Ademais, destacou que não firmou até então*

qualquer contrato que preveja disponibilização de dados dos sistemas sob sua guarda para o referido portal (peça 24, p. 4).

318. Além disso, o Serpro informou que não publica informações dos sistemas estruturantes do MP (Siape, Siasg, Comprasnet, SCDP e Siconv) nem da Secretaria do Tesouro Nacional (Siafi) diretamente no Portal Brasileiro de Dados Abertos (peça 24, p. 2).

319. O Serpro ressaltou que, nos contratos com esses clientes, há previsão de entrega de dados específicos e outras formas de extrações ao cliente demandante, as quais são faturadas sob o serviço de “apuração especial”. A entidade salientou que o destino do dado entregue ao cliente demandante não é de responsabilidade do Serpro (peça 24, p. 3).

320. O contrato celebrado entre o Serpro e o MP, por exemplo, define o serviço “apuração especial” como funcionalidades executadas apenas uma vez (grifo nosso) para: corrigir problemas ou atualizar dados nas bases de dados das aplicações; gerar um relatório específico ou arquivo para o usuário por meio de recuperação de informações nas bases da aplicação. Esses serviços, quando demandados, são cobrados seguindo a metodologia de pontos por função estabelecida pelo Sisp. De forma análoga, o contrato entre o Serpro e a STN define “apuração especial” como o desenvolvimento de rotinas para extração eventual de dados de um sistema ou para acerto do conteúdo da base de dados, sendo faturado também por ponto de função (peça 24, p. 3).

321. Assim, o Serpro envia dados dos sistemas estruturantes do MP para o portal da transparência da Controladoria Geral da União, conforme formatos e leiautes definidos pelo MP. Os dados do SCDP e do Siconv, semanalmente, e os dados do Siape, Siasg e do Comprasnet, mensalmente.

322. No caso do contrato com a STN, por exemplo, o Serpro disponibiliza dados do Siafi no formato de arquivos txt para download por usuários habilitados. Ou seja, nesse caso há restrição de acesso aos dados (peça 24, p. 2).

323. Convém lembrar que, em geral, os dados são disponibilizados em arquivos txt, formato não estruturado, portanto. Além disso, não são fornecidos metadados para permitir o entendimento do dado que está sendo entregue, o que seria necessário para serem considerados dados abertos (peça 24, p. 4).

324. No modelo de apurações especiais, os próprios gestores do Serpro apontaram as seguintes dificuldades para a abertura de dados (peça 24, p. 4):

- a) baixos ganho de escala e reuso;
- b) muito esforço na realização de cada demanda, visto que exige formatação e estruturação dos dados nos padrões definidos para dados abertos (organização para possibilitar o processamento por máquinas e inclusão de metadados);
- c) exigência de conhecimento especializado dos modelos internos das bases de dados;
- d) concorrência por recursos humanos com o desenvolvimento e evolução dos sistemas.

Uso de tecnologias de Big Data pelo Serpro

325. Com relação ao assunto, o Serpro informou que, nos últimos dois anos, vem prospectando tecnologias que permitam acessar grandes volumes de dados com baixa latência. Dentre essas tecnologias, citam Appliances de bancos de dados, ou seja, dispositivos de hardware com software integrado especificamente projetado para um propósito particular, no caso, consultar grandes bancos de dados de ambientes de Business Intelligence.

326. Outra linha de prospecção, segundo os gestores do Serpro, é a disponibilização dos dados dos sistemas via sua replicação, possibilitando seu acesso, em tempo real e sem concorrer com o ambiente produtivo. Esse tipo de abordagem tem sido chamada de Dados como Serviço (Data as a Service). Mas nessa prospecção, o Serpro ainda utiliza bancos de dados tradicionais na replicação dos dados. Uma evolução desta abordagem seria o uso de banco de dados baseados nas técnicas de Map Reduce (research.google.com/archive/mapreduce.html), tais como Hive/Hadoop (en.wikipedia.org/wiki/Apache_Hive), em substituição aos bancos de dados tradicionais atualmente utilizados. Segundo os interlocutores do Serpro, essa evolução no conjunto das arquiteturas utilizadas seria o que o mercado vem chamando de “Big Data”, o que ainda não é praticado pela empresa.

Conclusão

327. Em face do exposto no levantamento realizado junto ao Serpro, conclui-se que:

- a) a empresa custodia e difunde informações de alto valor social;
- b) trata-se de uma instituição que ainda não se mostra sensível à importância dos dados abertos, uma vez que, embora mantenha o Portal Brasileiro de Dados Abertos, não demonstrou estar se preparando para oferecer serviço exclusivo e específico para abertura de dados de seus clientes;
- c) o modelo de apurações especiais adotado pelo Serpro para disponibilizar dados públicos, de propriedade dos seus clientes, seja para publicação no Portal Brasileiro de Dados Abertos ou para outra destinação, mostra-se ineficiente e oneroso para a Administração Pública Federal, em face das dificuldades reconhecidas pela própria empresa pública;
- d) a empresa está prospectando tecnologias que permitam acessar grandes volumes de dados em alta velocidade, mas ainda não utiliza tecnologias que o mercado vem chamando de Big Data.

4.6 Oportunidades de utilização de dados públicos de funções essenciais de governo

328. Considerando a relevância e impacto das ações de governo nas funções de saúde e educação, nessa seção serão relatados aspectos da abertura de dados observados em dois entes públicos responsáveis por produzir e/ou custodiar dados públicos nas áreas de educação e saúde, a saber, o Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep) e o Departamento de Informática do SUS (Datasus).

329. Serão descritas as iniciativas de abertura de dados, aspectos relacionados à segurança e privacidade, existência de inventário de dados, dificuldades enfrentadas na disponibilização dos dados, e as iniciativas em tecnologias de Big Data, entre outros aspectos. Ambos os entes foram visitados pela equipe de fiscalização e a eles foram solicitadas informações por meio de ofícios de requisição.

4.6.1 Educação – Inep

330. O Inep é uma autarquia federal vinculada ao Ministério da Educação (MEC), cuja missão é promover censos, exames, pesquisas e avaliações sobre o Sistema Educacional Brasileiro, com o objetivo de subsidiar políticas públicas na área da educação, bem como produzir informações claras e confiáveis aos gestores, pesquisadores, educadores e público em geral.

331. Por tratar-se de uma entidade destinada a gerar e divulgar informações de alto valor para a gestão e o acompanhamento da educação no Brasil, o Inep é um ator estratégico no que diz respeito a governo aberto e dados de interesse social.

332. *As informações do Inep possibilitam conhecer o desempenho escolar e as atividades de ensino por meio de indicadores cujo nível e detalhe abrange as informações sobre a escola agregada até o nível do país como um todo. Esses indicadores permitem comparar a realidade do ensino entre localidades, mas também entre o Brasil e os demais países. Além de informações estatísticas e indicadores, o Inep possui ainda uma rica base de microdados, captados em censos e exames periódicos de abrangência nacional. Nesses microdados encontra-se retratada a realidade das escolas, professores e estudantes, assim como a ação do Estado, com grande riqueza de detalhes.*

333. *No contexto desse diagnóstico em torno do tema dos dados abertos, embora as informações estatísticas e indicadores tenham reconhecido valor, é nos microdados que reside e o maior potencial de análise e aproveitamento pela sociedade em geral.*

334. *O Inep faz parte do Comitê de Estatísticas Sociais, coordenado pelo IBGE. O comitê tem por objetivo instituir a cooperação dos órgãos produtores de informações sociais e buscar o fortalecimento e aprimoramento das informações estatísticas sociais oficiais, visando ampliar o uso, a variedade e qualidade das informações (ces.ibge.gov.br).*

Abertura de dados

335. *O Inep declara possuir ao todo trinta diferentes bases de dados, sendo que as informações de dez dessas bases são disponibilizadas com o nome de microdados em seu portal. Os microdados constituem conjuntos de dados em formato aberto destinados ao uso e reuso pela sociedade (peça 31, p. 10-12). Embora disponibilize as informações de apenas dez das suas trinta bases de dados, o Inep afirma que os dados abertos representam mais de 80% de todas as informações de sua propriedade.*

336. *Segundo a autarquia, um inventário contendo documentação sobre suas bases de dados é publicado internamente com acesso restrito aos seus servidores. Os artefatos que compõem o inventário das bases de dados são: modelo de dados, dicionário de dados, scripts e layout para importação e exportação de dados. Essa documentação é atualizada sempre que existe alteração na sua estrutura. Layouts são criados quando existe a necessidade de importar e exportar dados para outros órgãos ou sistemas (peça 31, p. 10).*

337. *A publicação dos dados educacionais é feita no portal do Inep, portal.inep.gov.br, onde se provê o acesso a dados brutos e também a dados agregados. Dados brutos são dados desagregados ao maior nível de detalhamento possível, livres de qualquer processamento ou consolidação. Dados agregados são dados que de alguma forma foram consolidados ou integrados.*

338. *O portal do Inep permite o download de arquivos contendo os microdados gerados por avaliações, pesquisas e exames, em formatos csv e ASCII, sendo também possível a sua leitura utilizando os softwares Statistical Analysis System (SAS) e Statistical Package for the Social Sciences (SPSS). O download traz informações adicionais para o entendimento dos dados (metadados), tais como o dicionário de dados e o manual de utilização.*

339. *O Inep disponibiliza, ainda, dados agregados sob a forma de sinopses estatísticas, indicadores educacionais e índices educacionais. Esses dados são fornecidos no portal, em sua maioria em formato csv. Planilhas proprietárias e PDF também são formatos utilizados.*

340. *Parte dos microdados publicados no portal do Inep também estão disponibilizados no Portal Brasileiro de Dados Abertos do Governo Federal*

(dados.gov.br), que contém, além desses, outros dados sobre o Sistema Brasileiro de Educação. Em 25/6/2014, esse portal contava com vinte conjuntos de dados referentes à função de governo educação, sendo quinze deles produzidos pelo Inep. Além de dados brutos, tais como o cadastro das escolas da educação básica (dados.gov.br/dataset/instituicoes-de-ensino-basico) ou o cadastro das instituições de ensino superior (dados.gov.br/dataset/instituicoes-de-ensino-superior), oferece, também, dados agregados, com visões específicas das informações educacionais, como, por exemplo, a média de alunos por turma e a média de horas-aula diária da educação básica.

341. *A publicação dos microdados no Portal Brasileiro de Dados Abertos demonstra que o Inep aderiu à Inda, embora as bases lá disponibilizadas não preencham a definição de dado aberto utilizada pelo portal e, por isso, não recebam o selo “open data” (parágrafo 278), além de muitas delas estarem desatualizadas. Em consulta realizada no dia 25/6/2014, verificou-se que a última versão lá disponibilizada do censo escolar refere-se a 2011, enquanto que, no portal do Inep, a última versão é do censo de 2013. Da mesma forma, os microdados da Prova Brasil, no portal de dados abertos, tem como última versão a edição de 2009, enquanto, no portal do Inep, a última versão é a de 2011.*

Classificação das informações e tratamento de dados pessoais

342. *Com relação à proteção da privacidade, tem-se que é responsabilidade do Inep assegurar o sigilo e a proteção dos dados pessoais, conforme determina o Decreto 6.425/2008 (peça 31, p. 6). O Inep conta com Comitê de Segurança da Informação e Comitê de Classificação das Informações. A autarquia informa que os dados pessoais de alunos, docentes e funcionários coletados pelos diversos sistemas só podem ser visualizados por quem os informou, pela escola ou por instituição de ensino superior (peça 31, p. 12). Para o controle de acesso, o Inep possui uma hierarquia de tipos de usuários com perfis definidos. Esse controle cabe quando do acesso direto aos sistemas do Inep ou para requisições de informações com fulcro na LAI (transparência passiva).*

343. *Dessa forma, as informações disponibilizadas à sociedade não contemplam nomes de pessoas nem documentos pessoais que possibilitem a identificação do indivíduo. A distinção dos indivíduos é feita por meio dos códigos e números. Os microdados do Exame Nacional do Ensino Médio, por exemplo, individualizam o aluno por meio de seu número de inscrição, inviabilizando sua identificação para aquele que não dispor desse número.*

Dificuldades à abertura de dados no Inep

344. *Ao ser questionado sobre as dificuldades à abertura dos dados, a autarquia citou, basicamente, a restrição de recursos materiais adequados e de pessoal para a coleta de informações (peça 31, p. 6). Nota-se que o requerido sigilo das informações pessoais de alunos e docentes também se mostra um empecilho para a disseminação dos dados, na medida em que ainda constitui, hoje, uma barreira para que seja público o acesso ao Sistema de consulta Inepdata (peça 31, p. 15).*

345. *Ademais, nas reuniões realizadas por esta equipe de fiscalização, o Inep citou a preocupação com a qualidade dos dados. Um dos aspectos ressaltados pelo Inep se refere a informações incorretas ou desatualizadas que dificultam ou impedem o acompanhamento da evolução da educação do aluno.*

346. *Com relação a dificuldades quanto à correção dos dados, vale comentar sobre um tipo específico de dados passíveis de abertura, os dados georreferenciados ou de geolocalização. Dá-se o nome de georreferenciamento à determinação precisa de um ponto na superfície terrestre. A utilização de dados georreferenciados tem sido muito*

difundida nos dias atuais, pois, além de fornecer a localização exata e única de um objeto, possibilita o desenvolvimento de soluções e aplicações que utilizem informações espaciais. Considerando-se a área da educação, dados georreferenciados podem contribuir não somente ao informar a localização de uma escola, mas também ao possibilitar a oferta de soluções, como, por exemplo, de um aplicativo que informe o rol de escolas próximas a uma determinada localidade ou as escolas próximas a um determinado percurso, que pode ser o percurso desde a casa de um cidadão até seu trabalho.

347. *Soluções internacionais que utilizam dados educacionais fazem o uso contínuo de georeferenciamento. Como exemplo, pode-se citar: illustreets.co.uk, aplicativo do Reino Unido que fornece indicadores sobre segurança pública, educação, transporte e padrão de vida de uma determinada localidade selecionada pelo usuário (parágrafo 151).*

348. *No caso do Inep, em 2012, a autarquia incluiu dentre as questões do Censo Escolar da Educação Básica os dados referentes à latitude e a longitude da localização das escolas. Entretanto, como se trata de um dado de coleta recente e não tão disponível nas escolas como os outros dados administrativos, há dificuldades na obtenção dessa informação, o que, muitas vezes, compromete seus requisitos de confiabilidade, consistência e fidedignidade. Segundo a autarquia, a principal dificuldade para a coleta adequada é a falta de definição sobre os padrões, equipamentos e processos de levantamento da latitude e longitude com vistas à localização de estabelecimentos por ponto, estabelecendo a precisão requerida para levantamentos desse tipo de dado (peça 31, p. 7-8).*

349. *Em relação a essa questão, a autarquia informou que vem tomando providências para que os dados alcancem os requisitos necessários à divulgação, tais como: consulta à CONCAR/IBGE e à Infraestrutura Nacional de Dados Espaciais (Inde) sobre os padrões que devem ser adotados; treinamento da equipe do Inep para a disseminação dessas orientações; pesquisa sobre o equipamento e a solução a ser adotada para obtenção da localização da escola; e adoção de críticas de consistência para garantia da qualidade da informação (peça 31, p. 8).*

350. *Embora de baixa qualidade, dados referentes a georeferenciamento já são divulgados, conforme informado pelo próprio Inep (peça 31, p. 7). A equipe de fiscalização realizou o download do conjunto de dados denominado Instituições de Ensino Básico, com os dados referentes ao cadastro das escolas da educação básica da região centro-oeste em 2012, obtido do portal dados.gov.br, em 22/5/2014. Constatou-se a existência de atributos referentes à latitude e à longitude de cada registro de escola. Todavia, dos 11.364 registros de escolas listados no arquivo, 3.624 deles não continham valores dos atributos de georeferenciamento, ou seja, estavam vazios os valores de latitude e longitude, o que confirmou a baixa qualidade dessa informação. Para exemplificar, foram apensados ao processo alguns registros desse conjunto de dados (peça 52).*

351. *A dificuldade de captação de dados primários também é um problema vivenciado também por outros países. No Seminário Internacional Sobre Dados Educacionais, ocorrido em Brasília em 9/5/2014, do qual participou um membro desta equipe de auditoria, essa dificuldade foi colocada e apresentadas sugestões para superá-la, tais como: i) implementação de campanha para conscientizar as escolas da utilidade do fornecimento de dados; ii) auxílio à gestão das escolas; iii) mais recursos de captação de dados e demais incentivos.*

Iniciativas de dados abertos do Inep com participação social

352. *Com o intuito de incentivar a sociedade a se apropriar das bases de dados educacionais, em 2013, o Inep, a partir de um Termo de Cooperação Técnica com a Fundação Lemann, realizou a primeira edição do Hackathon de Dados Educacionais. O Inep entende que o evento segue as diretrizes da Lei de Acesso a Informação, que incentiva a participação popular, a divulgação de informações de interesse público e o fomento ao desenvolvimento da cultura de transparência na administração pública (peça 31, p. 2).*

353. *O evento contou com trinta participantes, em sua maioria programadores, divididos em oito equipes, que foram desafiadas a elaborar soluções que utilizassem os microdados da Prova Brasil 2011. Além dos dados, o Inep proveu informações de auxílio à compreensão dos dados, como escalas de desempenho e matrizes de referência (peça 31, p. 2).*

354. *Como resultado, foram apresentadas uma diversidade de soluções criativas e abordagens para diferentes problemas, tendo sido premiados os três melhores projetos. “A Escola que Queremos” (escolaquequeremos.org) foi o projeto vencedor e constituiu a construção de um indicador da escola que levasse em consideração mais do que o fluxo escolar e a proficiência do estudante, utilizando catorze variáveis originárias do Censo da Educação Básica e do questionário socioeconômico da Prova Brasil (parágrafo 226). Foram citados, ainda, os seguintes benefícios advindos do evento:*

- a) evidenciação da necessidade de debater a questão dos dados abertos na área educacional, dando maior transparência às informações oriundas dos sistemas de avaliação e do censo;*
- b) promoção do encontro da comunidade que pensa dados abertos e desenvolvimento de softwares em torno do tema educacional, atraindo novos talentos para propor soluções em TI que aprimorem as formas de divulgação dos dados educacionais;*
- c) reunião de diferentes propostas de grupos de diversas regiões do Brasil, estimulando a troca de conhecimento e interlocução entre diferentes instituições e comunidades;*
- d) aproximação dos pesquisadores do Inep ao grupo de estudantes, jornalistas, desenvolvedores e demais interessados em explorar os dados educacionais, oportunizando a troca de conhecimento e a possibilidade de o Inep captar novas ideias sobre o uso dos dados educacionais;*
- e) inserção da educação no campo do debate sobre dados abertos de forma prática e colaborativa.*

355. *Os projetos, contudo, não foram continuados. Após a entrega da premiação, não houve mais contatos do Inep com as equipes vencedoras. No entendimento do Inep, a questão da continuidade e sustentabilidade dos projetos deve ser amadurecida nos próximos eventos e depende de um maior envolvimento entre as diretorias do Inep, sobretudo, das Diretorias de Tecnologia da Informação (DTDIE) e de Disseminação de Dados Educacionais (Deed).*

356. *O segundo Hackathon sobre dados educacionais ocorreu em maio de 2014, cujo foco foi a sustentação dos projetos e não mais a entrega de protótipos. Para o evento, houve a preocupação em garantir a diversidade dos perfis dos participantes que, além de programadores, eram designers e educadores, e a liberdade de uso de qualquer microdado disponibilizado pelo Inep, sem a restrição de uma informação específica, como tinha sido estabelecido no evento anterior. Houve 89 inscritos, tendo sido selecionados quarenta participantes, que foram distribuídos em dez grupos.*

357. *De forma a dar mais sustentabilidade aos projetos oriundos do evento, o edital previu duas semanas de imersão das equipes na autarquia para que os projetos fossem desenvolvidos pelas equipes vencedoras em conjunto com o Inep.*

358. *Por fim, o projeto vencedor foi Dados Abertos API, que tem o objetivo de facilitar o acesso a bases públicas de dados do Inep por meio de APIs. A definição de API foi comentada no parágrafo 205. O Inep coloca que a escolha desse projeto indica a importância da pauta de dados abertos e de se facilitar a apropriação dos dados públicos. Acrescentou que os outros dois projetos vencedores apontam para pautas de relevância social como a falta de creches e o problema da formação dos professores, mostrando que dados do Inep e de outros órgãos podem servir para evidenciar problemas sociais e auxiliar na resolução desses problemas (peça 54, p. 2).*

Conclusão

359. *Em face do exposto no levantamento realizado junto ao Inep, conclui-se que:*

- a) o instituto custodia, gera e difunde informações de alto valor social (parágrafos 330-332);*
- b) trata-se de uma instituição sensível à importância dos dados abertos e que, paralelamente ao seu trabalho de coleta, análise e produção de informações estatísticas, procura disponibilizar microdados em formatos abertos e estimular seu uso (parágrafos 336-341 e 352 a 356);*
- c) sobre as dificuldades à abertura dos dados, a autarquia destaca, basicamente, a restrição de recursos materiais adequados e de pessoal para a coleta de informações, bem como problemas na qualidade dos dados (parágrafos 344 a 351);*
- d) o instituto aderiu à Inda e ao Portal Brasileiro de Dados Abertos do Governo Federal, tendo disponibilizado cerca de quinze conjuntos de dados. Esses dados não preenchem todos os requisitos da definição de dado aberto utilizada pelo portal e muitos deles estão desatualizados (parágrafos 341);*
- e) em termos de melhorias, o instituto poderia facilitar o acesso e a compreensão de seus microdados se publicasse um inventário centralizado descrevendo suas bases e a semântica dos dados ali contidos. Um controle mais efetivo da qualidade dos dados e a consolidação dos dados de posicionamento geográfico das escolas tornaria os microdados ainda mais úteis e férteis. A atualização dos microdados no Portal Brasileiro de Dados Abertos é importante para que sua utilização seja difundida.*
- f) a disponibilidade de dados confiáveis permite evoluir na análise e posterior debate de políticas educacionais, distanciando-as de visões ideológicas ou preferências pessoais para decisões mais embasadas, testáveis e refutáveis, além de fornecer mais insumos para o controle da sociedade.*

4.6.2 Saúde – Datasus

360. *O Datasus é o Departamento de Informática do Sistema Único de Saúde (SUS), órgão da Secretaria de Gestão Estratégica e Participativa (Sgep) do Ministério da Saúde (MS) com a responsabilidade de coletar, processar e disseminar informações sobre saúde.*

361. *Ao Datasus compete (Decreto 8.065/2013, art. 38):*

- I - fomentar, regulamentar e avaliar as ações de informatização do SUS, direcionadas à manutenção e ao desenvolvimento do sistema de informações em saúde e dos sistemas internos de gestão do Ministério da Saúde;*
- II – desenvolver, pesquisar e incorporar produtos e serviços de tecnologia da informação que possibilitem a implementação de sistemas e a disseminação de informações para ações de saúde, em consonância com as diretrizes da Política Nacional de Saúde; (...)*

IV - manter o acervo das bases de dados necessários ao sistema de informações em saúde e aos sistemas internos de gestão institucional; (...)

362. *Por tratar-se de uma unidade destinada a custodiar e divulgar informações de alto valor para a gestão e o acompanhamento da saúde no Brasil, o Datasus é um ator estratégico no que diz respeito a governo aberto e dados de interesse social.*

363. *Repisa-se que o Datasus não produz os dados da saúde, sendo apenas o custodiante dos dados fornecidos pelas unidades de saúde de estados e municípios, a saber, as secretarias de saúde estaduais e municipais, em uma estrutura de competências concorrentes entre União, estados e municípios.*

364. *Isto pode ser exemplificado na operação do sistema Siasus, gerido pelo Datasus e processado nas unidades ambulatoriais credenciadas pelo SUS com a finalidade registrar os quantitativos e valores a serem pagos aos prestadores de serviços (públicos e privados), produzindo informações locais que são consolidadas a nível nacional no Banco de Dados Nacional de Informações Ambulatoriais do SUS (www2.datasus.gov.br/DATASUS/index.php?area=0601).*

Abertura de dados

365. *Como diretriz para abertura dos dados custodiados, o Datasus informou que no PDTI 2014/2015 do Ministério da Saúde há o seguinte objetivo: “promover e fortalecer as ações de TI que contribuam para o efetivo acesso à atenção, à promoção, à gestão e à vigilância em saúde, para o governo aberto e para o engajamento da sociedade”. Vinculados a esse objetivo, constam os seguintes projetos/ações (peça 30, p. 7):*

- a) projeto/ação 5.1.1 – promover ações de divulgação do Portal do Cidadão;*
- b) projeto/ação 5.1.2 – elaborar plano de disseminação da informação nas mídias e ferramentas disponíveis na comunidade;*
- c) projeto/ação 5.1.3 – definir e implementar ações que favoreçam o acesso a informações.*

366. *Em que pese o objetivo citar a promoção e fortalecimento de ações para governo aberto, o Ministério não apresentou no PDTI supramencionado metas vinculadas a governo aberto nem à abertura de dados.*

367. *Na reunião de 12/5/2014, os gestores do Datasus informaram que existem 171 sistemas de informação mantidos pelo departamento, conforme se observa no PDTI do Ministério da Saúde (peça 62, p. 44), nos quais se encontram fragmentadas as informações referentes à saúde pública no Brasil, sem padronização nem garantia interoperabilidade entre os sistemas.*

368. *Em resposta ao Ofício de requisição 5-137/2014, o Datasus informou que possui inventário das bases de dados. Compõem este inventário a descrição de cada sistema, sigla, identificação do gestor, dentre outras informações. O inventário é de acesso restrito ao Ministério da Saúde (peça 30, p. 3).*

369. *Segundo os gestores, o Datasus ainda não publica os microdados de seus sistemas de informação na Infraestrutura Nacional de Dados Abertos (parágrafo) devido à ausência de decisão das áreas de negócio do Ministério da Saúde, gestoras dessas informações, que aguardam a normatização do processo de solicitação e autorização da liberação dos dados (peça 30, p. 6-7). Todavia, publica dados de seus sistemas no padrão da Infraestrutura Nacional de Dados Espaciais (Inde) (peça 30, p. 3).*

370. *Por exemplo, o portal Sala de Apoio à Gestão Estratégica (Sage), criado pela Portaria – MS 406/2012, disponível no endereço www.saude.gov.br/sage, permite a visualização da geolocalização de farmácias populares pelos usuários.*

371. *Contudo, o Datasus confirmou que nenhuma das bases que armazenam dados de geolocalização estão integradas às informações de outros sistemas, tais como as informações do Cadastro Nacional de Estabelecimentos de Saúde (CNES), que mantém informações dos estabelecimentos públicos e privados, mas sem visualização de dados de geolocalização (peça 30, p. 7-8).*

372. *Segundo informações do chefe de gabinete da Secretaria de Gestão Estratégia do Ministério da Saúde (SGEP), durante a reunião realizada em 12/5/2014, essa funcionalidade de integração de dados de vários sistemas sob uma interface baseada em mapas encontrava-se em desenvolvimento e seria oferecida no futuro sítio mapadasaude.saude.gov.br, que se encontrava em fase de testes, com acesso limitado.*

373. *Também na entrevista, foi informado que embora o CNES preveja o cadastramento de informações de geolocalização (latitude e longitude) dos estabelecimentos de saúde, frequentemente as unidades municipais, estaduais e federais, e principalmente os estabelecimentos privados, não fornecem essa informação. A equipe de fiscalização não identificou incentivos previstos para que esse tipo de informação seja informado de maneira recorrente e confiável via sistema.*

374. *Assim, a qualidade e confiabilidade dos dados fornecidos varia bastante entre os diversos sistemas de informação. A falta de qualidade e confiabilidade dos dados é ainda maior naqueles sistemas cujos dados são incluídos pelos estabelecimentos privados que não têm incentivos em cadastrar adequadamente seus dados nos sistemas de informação do Datasus.*

375. *Outra ferramenta de disponibilização de dados públicos oferecida pelo Datasus refere-se ao portal TabNet, que possibilita a geração de relatórios sobre informações da saúde no formato de arquivos .dbc, um formato proprietário. Esses relatórios não foram concebidos na atual perspectiva de dados abertos, e sim para fornecer dados brutos para gestores da saúde e pesquisadores. Para isso, o Datasus fornece gratuitamente há 25 anos o software de tabulação de dados Tabwin para tornar possível o uso dos dados armazenados nos arquivos .dbc. Desse modo, para que sejam utilizados os dados desses arquivos, é necessário que o Tabwin seja instalado no computador de quem pretende extrair e analisar dados contidos nos relatórios do TabNet.*

376. *Em geral, as informações acessíveis pelo portal do TabNet encontram-se agregadas em nível municipal, mas há a opção de obter arquivos com microdados anonimizados (arquivos cujo conteúdo busca impedir a identificação das pessoas ali referenciadas ao retirar ou mascarar, total ou parcialmente, atributos chave tais como nome, número de telefone ou de documentos de identificação) para uso com o TabWin (cnes.datasus.gov.br, menus “serviços”, “arquivos (outros)”).*

377. *Para avaliar as formas de acesso aos dados da saúde custodiados pelo Datasus, foi testada a capacidade e dificuldade em obter, extrair e reutilizar microdados do Sistema de informações de Mortalidade, cujo procedimento está detalhadamente descrito nos subitens.*

377.1. *O portal da saúde (datasus.saude.gov.br) reagrupa informações de diversos sistemas gerenciados pelo Datasus. Procurando em seus diversos menus, é possível encontrar endereços para baixar dados em arquivos eletrônicos (datasus.saude.gov.br/arquivos-de-dados). Seleciona-se a opção de “Arquivos*

dissemináveis para tabulação do Sistema de informações de Mortalidade”; Em seguida, no menu de consulta, seleciona-se, por exemplo, a modalidade de dados, declarações de óbito, ano 2012 (mais recente disponível) e todas as unidades da federação; confirma-se a seleção de todos esses arquivos, um por unidade federativa apenas para o ano de 2012, compactados em um único arquivo .zip. Recebido e descompactado o arquivo, tem-se, então, 27 arquivos no formato “DOXX2012.DBC” onde XX corresponde à sigla do estado.

377.2. Para poder utilizar as informações desses arquivos, é necessário obter e instalar o programa TabWin, de tabulação de dados, disponibilizado gratuitamente pelo Datasus (www2.datasus.gov.br/DATASUS/index.php?area=060805&item=3). Dentre os programas instalados junto com o TabWin encontra-se o dbf2dbc.exe, que permite descompactar os arquivos .dbc baixados para o formato .dbf utilizado pelo TabWin. Uma vez esses dados carregados no TabWin, é possível exportá-los em um formato padrão aberto, como .csv, para reuso em novas aplicações.

377.3. Assim, foi gerado um arquivo .csv de todos os óbitos do estado de Roraima em 2012 e, utilizando-se os mesmos procedimentos, seria possível fazer o mesmo com todas as demais unidades da federação e anos disponíveis. Para montar todos esses dados em uma base única e íntegra de microdados de óbitos no Brasil, nos últimos cinco anos, por exemplo, seria necessário utilizar um sistema gerenciador de banco de dados (SGBD) para agregar os diversos arquivos individuais.

377.4. Após essa etapa de consolidação dos dados anuais por Estados, seria possível realizar novas análises e eventualmente descoberta de novos conhecimentos que possam ter relevância para a saúde pública.

378. Assim, como descrito acima, e assumindo que esse seria o procedimento para obter outras informações custodiadas pelo Datasus, pode-se avaliar que estes se encontram no padrão de dados abertos correspondente a duas estrelas na escala de classificação, 5 Stars Open Data (5stardata.info). Isso se dá em virtude de que as informações encontram-se acessíveis pela internet, em documento eletrônico estruturado, mas em formato proprietário que requer o programa TabWin, disponível gratuitamente para os usuários de informações do Datasus, para acesso e reuso dos dados.

379. Para atender algumas características de dados abertos consideradas na LAI (possibilitar o acesso automatizado por sistemas externos em formatos abertos, estruturados e legíveis por máquina; e serem atualizados) e maximizar as chances de que esses dados sejam reaproveitados em projetos de pesquisa e novos serviços à população, seria necessário, pelo menos, que esses dados fossem automaticamente atualizados com as suas versões disponíveis mais recentes e que estivessem acessíveis em formatos estruturados e não proprietários como .csv, o que corresponderia a três estrelas na escala de classificação 5 Stars Open Data.

Classificação das informações e tratamento de dados pessoais

380. O processo de classificação da informação custodiada pelo Datasus é disciplinado pela Portaria 1.583, de 19 de julho de 2012. No portal da saúde, pode-se verificar a classificação realizada no ano de 2013.

381. Ademais, o Datasus informou que estão sendo implementadas condições para tramitação e segurança da informação classificada na esfera do ministério, conforme a Lei 12.527/2011 e Decreto 7.845/2012. Esse projeto consta do PDTI 2014-2015 do órgão, vinculado ao objetivo “estabelecer padrões de segurança da informação”.

382. Quanto à questão da proteção ao sigilo de informações médicas pessoais, o Datasus informou que os procedimentos adotados para o tratamento de dados que envolvam a privacidade das pessoas são consubstanciados na possibilidade de se separar os dados de identificação do indivíduo daqueles relacionados à identificação do evento sem si (peça 30, p. 6), o que se costuma de chamar de anonimização. Inclusive, o PDTI 2014-2015 contempla a ação “implantação do processo de vinculação e disponibilização de dados anonimizados para pesquisadores”, vinculado ao objetivo de “estabelecer padrões de segurança da informação”. Nesse sentido, o Datasus adquiriu e está implementando o uso de ferramentas de criptografia de dados e de gerenciamento de identidades e trilhas de auditoria. Em paralelo, o Ministério informou que estão em andamento os seguintes projetos (peça 30, p. 6):

382.1. elaboração de portaria de acesso e cessão de dados com informações pessoais, cuja minuta aguardando aprovação do Ministério. Esse normativo pode influenciar a forma como o Datasus disponibiliza microdados em que as informações pessoais, até o momento de conclusão deste relatório, eram anonimizadas;

382.2. instituição da gestão de riscos para os ativos da Divisão de Análise e Administração Estratégica de Dados do MS (bases de dados, servidores, conexões, etc);

382.3. sistema de gestão de acesso e cessão às bases de dados do MS.

383. Quanto ao processo decisório utilizado em face de pedidos de acesso externo a dados pessoais, o Datasus declarou que está condicionado às hipóteses contidas no art. 31 da Lei 12.527/2011, bem como no disposto no Decreto 7.724/2012 e na Portaria GM 1.583/2012. Além disso, tal liberação de acesso está condicionada à aprovação dos gestores dos sistemas.

Uso de tecnologias de Big Data pelo Datasus

384. Finalmente, quanto ao uso de tecnologias relacionadas a Big Data, o Datasus declarou que tem sido representado em eventos, workshops e seminários sobre o tema, prospecta tecnologias e está buscando parcerias com instituições como Fiocruz, COPPE/UFRJ e FGV, entre outras, para desenvolvimento de técnicas de Big Data na análise de dados da saúde, já tendo realizado discussões com as instituições relacionadas. Porém, afirmou o gestor, ainda não há projeto específico definido (peça 30, p. 8).

385. Por outro lado, o Datasus manifestou que é preciso estabelecer processos e procedimentos de acesso às bases de dados, para que se possa extrair conhecimento sem colocar em risco a privacidade do cidadão e evitar o risco do uso inescrupuloso do acesso (peça 30, p. 8).

Conclusão

386. Em face do exposto no levantamento realizado junto ao Datasus, conclui-se que:

a) o departamento custodia e difunde informações de alto valor social (parágrafos 362 a 363);

b) trata-se de uma instituição que aderiu à Inda e que já possui processos de trabalho para tornar disponível tanto dados estatísticos quanto microdados por meio das ferramentas TabNet/TabWin (parágrafos 369 a 379);

c) o Datasus poderia aperfeiçoar o acesso aos dados, fornecendo-os diretamente em formatos abertos que não requeiram o uso de ferramentas específicas, aderindo à Inda e integrando seus sistemas;

d) o Datasus entende a necessidade de se estabelecer processos e procedimentos para a abertura de dados, de forma a evitar o risco do uso inescrupuloso do acesso e garantir a privacidade do cidadão (parágrafos 380 a 383);

e) o Datasus está buscando parcerias com instituições visando ao uso de tecnologias relacionadas a Big Data na análise de dados da saúde, porém ainda não há projeto definido (parágrafo 384).

4.7 Soluções brasileiras atuais comparadas às soluções internacionais pesquisadas

387. Em relação às funções de governo de saúde e educação, observou-se que muitas das soluções internacionais citadas neste trabalho constituem aplicações simples para a identificação ou comparação de instituições em termos de qualidade, apresentando, basicamente, os indicadores das instituições selecionadas.

388. Na educação, este trabalho cita as soluções Find The Best Primary Schools (parágrafos 171 a 175) e Sistema de Medición de la Calidad de la Educación (parágrafo 159), que apresentam indicadores das escolas, com o intuito de auxiliar os pais na escolha da escola a matricular seus filhos e no acompanhamento do desempenho da instituição escolhida. Ao se pesquisar soluções nacionais similares na área educacional, a equipe de fiscalização identificou as seguintes soluções, descritas anteriormente: QEdu (parágrafos 228 a 229) e Escola Que Queremos (parágrafos 226 a 227).

389. Comparando-se as soluções internacionais com as nacionais, pode-se identificar funcionalidades, naquelas, que faltam às soluções desenvolvidas no Brasil e que seriam úteis aos cidadãos em suas pesquisas por informações educacionais. Merece destaque a diversidade de caminhos que a solução Find The Best Primary Schools dispõe ao usuário para a seleção da escola. Ele pode fazer a busca pelo nome ou parte do nome da escola pré-selecionada; pode obter uma lista das escolas próximas a uma localidade por definida pelo usuário, dentro de um perímetro também definido por ele; ou, se preferir, pode visualizar o mapa do país, ou de uma localidade, com a marcação das escolas existentes para selecionar aquelas a comparar.

390. A informação do georreferenciamento das escolas proporciona a criação de soluções e também a construção de novas funcionalidades nas soluções existentes. O Inep vem tentando obter o georreferenciamento das escolas, mas, conforme relatado (parágrafos 346 a 350), tem enfrentado dificuldades na coleta dessas informações. Do exposto, entende-se ser essa questão uma oportunidade de melhoria para o Inep em relação à publicação dos dados escolares.

391. Um outro aspecto observado em Find The Best Primary Schools foi a apresentação de informações financeiras, além dos indicadores escolares. Para as escolas selecionadas, a solução apresenta as receitas e os gastos totais e por categorias (separa as despesas com pessoal envolvido no ensino, no apoio educacional e no suporte, despesas com recursos usados na aprendizagem tecnológicos e não, instalações, alimentação, energia etc), além de apresentar taxas de gastos por aluno. Também apresenta, numa mesma tela, o detalhamento dessas informações para cada escola selecionada, dispostas lado a lado, o que auxilia o usuário na comparação das escolas.

392. Na área de saúde, o levantamento identificou a solução Find The Best Hospitals (parágrafo 182). Essa solução apresenta os dados dos hospitais e instituições de saúde do Sistema Nacional de Saúde da Inglaterra (National Health Service – NHS) que foram selecionados pelo usuário, proporcionando uma análise comparativa dos índices relativos ao corpo clínico, às estatísticas das cirurgias, à taxa de mortalidade e ao tempo de atendimento, à avaliação dos pacientes, dentre outros.

393. *Ao pesquisar por soluções nacionais similares, a equipe de auditoria identificou o aplicativo Doctor Poa (parágrafos 224 a 225) como uma solução que se aproxima dessa solução internacional. O aplicativo, embora seja municipal e não traga indicadores de desempenho dos hospitais, permite ao usuário localizar as unidades de atendimento e hospitais mais próximos, com base em sua posição geolocalizada, além de uma busca por emergências em determinadas especialidades.*

394. *Ambas as soluções respondem a perguntas simples dos cidadãos: quais são as especialidades de cada hospital? Qual fica mais perto da sua casa ou trabalho?*

395. *Fazendo o mesmo exercício realizado na área da educação, ao se comparar as duas soluções, observa-se que o aplicativo Doctor Poa traz informações básicas dos hospitais selecionados, quais sejam: se é público ou privado, endereço, telefone, especialidades oferecidas e link. Por outro lado, a solução Find The Best Hospitals apresenta, além dessas informações, informações referentes aos serviços oferecidos e aos indicadores de desempenho dos hospitais, tais como:*

a) avaliações e taxas: tem-se, como exemplo, a avaliação média do paciente, avaliação do corpo médico, taxa de mortalidade, tempo médio de tratamento emergencial;

b) departamentos médicos e tratamentos oferecidos;

c) corpo médico;

d) indicadores NHS em relação a procedimentos usuais; são apresentados, por exemplo, o número de cesarianas realizadas, avaliação das cirurgias de hérnia, índice de satisfação das cirurgias de prótese de quadril e avaliações da qualidade da endoscopia.

e) tratamento de câncer: são apresentados indicadores, como, por exemplo, a avaliação do paciente quanto ao tratamento, a taxa de mortalidade em relação à taxa nacional, a taxa de readmissão etc.;

f) indicadores do tratamento emergencial: são apresentados, por exemplo, o tempo médio de atendimento no hospital, a taxa média para o atendimento da ambulância;

g) infraestrutura: são elencadas as facilidades e serviços não clínicos oferecidos pelo hospital;

h) hospitais similares próximos ao hospital selecionado.

396. *Todas essas são informações úteis aos cidadãos quando necessita de atendimento médico, auxiliando na identificação da disponibilidade do tratamento de que necessita, na escolha da unidade de melhor qualidade e, também, no acompanhamento do desempenho do hospital em que se é paciente. O aplicativo também responde a outras perguntas, também importantes para os cidadãos, como: qual o melhor hospital próximo da minha casa ou trabalho? Qual é a qualidade do exame de endoscopia realizado pelo hospital? Como está a qualidade do hospital que tenho frequentado? Qual a taxa de mortalidade referente a um determinado tipo de cirurgia?*

397. *Nota-se que, nas duas soluções analisadas, não há publicação de informações financeiras.*

398. *Da comparação, é importante ressaltar que a solução Doctor Poa constitui uma solução de âmbito municipal e que informações adicionais quanto ao desempenho dos hospitais agregariam valor a esse serviço prestado ao cidadão. Ademais, entende-se tratar-se de um exemplo de serviço a ser oferecido por outros municípios, estados, e mesmo em nível nacional.*

399. *Ainda na área da saúde, observa-se que o Datasus torna públicos alguns conjuntos de dados de geolocalização referentes à saúde, tais como os dados de*

geolocalização dos estabelecimentos de saúde ou farmácias populares. Porém, esses dados não são integrados às informações de outros sistemas, tais como o CNES, que mantém informações detalhadas dos estabelecimentos públicos e privados de saúde no país, mas não apresenta dados de geolocalização (parágrafos 371 a 372).

400. Além disso, há problemas relacionados à qualidade e confiabilidade das informações de geolocalização coletadas. É o caso do próprio CNES, que prevê o cadastramento dessas informações, mas, muitas vezes, as unidades municipais, estaduais e federais, e principalmente os estabelecimentos privados, não fornecem essa informação, não havendo incentivos previstos para que esse dado seja preenchido na atual estrutura de competências concorrentes entre União, estados e municípios, em que decisões são tomadas no âmbito de comissão tripartite (parágrafos 373 a 374).

401. Esses problemas de falta de integração entre sistemas e dados disponibilizados, bem como oscilações na qualidade e confiabilidade dos dados influenciam negativamente a criação de soluções a partir deles. A propósito, a equipe não conseguiu comparar alguma solução cívica nacional que utilize dados disponibilizados pelo Datasus (como fez em relação ao tema educação) com soluções internacionais.

Conclusão

402. Do exposto, entende-se que a disponibilização das informações em formato aberto sobre as instituições de saúde do país, incluindo a sua geolocalização, constitui insumo para o surgimento de novos serviços ao cidadão, disponibilizando informações que hoje não estão disponíveis ao cidadão comum.

5. Conclusão

403. Este trabalho buscou conhecer iniciativas de publicação de dados abertos na Administração Pública Federal. O conceito sobre dados abertos se insere na concepção de governo aberto, e se refere a projetos e ações que visam à promoção da transparência, à luta contra a corrupção, ao incremento da participação social e ao desenvolvimento de novas tecnologias, de modo a tornar os governos mais abertos, responsáveis por suas ações e aptos a atender às necessidades dos cidadãos (parágrafo 24).

404. Entende-se que a transparência e a abertura de dados figuram como aspectos correlatos e essenciais de um governo moderno. As práticas de governo aberto, quando implementadas, podem gerar benefícios tanto para a sociedade como para o próprio governo. No primeiro caso, destacam-se o aperfeiçoamento da participação social, da compreensão e da fiscalização das ações governamentais, propiciando o empoderamento do cidadão. No segundo, a melhoria do processo de tomada de decisão, do gerenciamento dos dados governamentais, da eficiência da gestão, de gerenciamento de riscos e dos serviços oferecidos aos cidadãos. Ao disponibilizar seus dados, o governo também favorece a criação de negócios e serviços inovadores, seja com valor comercial ou social (parágrafos 74 a 79).

405. Com a abertura dos dados governamentais, em âmbito internacional, organizações, cidadãos, acadêmicos, empresas e instituições públicas têm construído soluções para a produção e o compartilhamento da informação pública, levando conhecimento ao cidadão e estabelecendo um novo canal de comunicação entre governo e sociedade. Também chamadas de aplicações cívicas, as soluções existentes variam quanto ao nível de complexidade e podem ser voltadas a agilizar os serviços públicos oferecidos aos cidadãos, promover a cooperação na execução dos serviços públicos ou guiar o governo na elaboração e condução das políticas públicas (parágrafo 150).

406. Ainda no cenário internacional, observou-se que há anos organizações de vários países se empenham para compartilhar princípios e práticas de governo aberto e disseminá-las entre os governos. Apesar do interesse crescente em relação ao tema e do número de iniciativas cada vez mais volumoso, observou-se que o entendimento é de que ainda há muito a evoluir, de modo que os esforços para a consolidação de uma cultura de dados abertos permanecem nas agendas dos governos internacionais (parágrafo 118).

407. No cenário nacional, com a criação do Portal da Transparência, com o advento da Lei de Acesso a Informação (LAI), com a integração do Brasil à OGP, e com a criação da Infraestrutura Nacional de Dados Abertos (Inda) e do Portal Brasileiro de Dados Abertos (dados.gov.br), o país tem intensificado ações para a promoção da transparência, da prestação de contas, da participação popular na administração pública e, principalmente, da disponibilização de informações governamentais por meio dos dados abertos.

408. Pelo contexto desse trabalho e sua importância, foi realizada uma análise perfunctória do Portal Brasileiro de Dados Abertos, que constitui página central do tema no âmbito do Governo Federal. Concluiu-se que, embora o portal disponibilize dados de relevância, as dificuldades observadas demonstram a necessidade de evolução, tanto em termos de quantidade de conjuntos de dados disponibilizados, como em relação a formato e atualidade dos dados (parágrafos 274 a 293).

409. Em que pese o trabalho já realizado pelos atores dessa política, em destaque o esforço realizado pela SLTI/MPOG, ainda há muito a se fazer na direção de se estabelecer uma cultura de abertura de dados no Brasil. As ações realizadas ainda não se mostraram suficientes para a implementação e popularização do governo aberto no âmbito federal. Em princípio, para que os resultados esperados dessa política se intensifiquem, é necessário estabelecer um conjunto maior de incentivos para as instituições, que envolvam ações de conscientização sobre a importância do tema de modo a estimular o aumento da participação coordenada e ativa das instituições públicas em iniciativas de abertura de dados, abrangendo os três poderes, Executivo, Legislativo e Judiciário.

410. Em um cenário em que a abertura de dados públicos se populariza, aumentam as possibilidades de criação de soluções pelo governo ou pela própria sociedade, o que favorece o aparecimento de novas oportunidades de mercado para aqueles que desejem elaborar alternativas inovadoras para melhoria da prestação de serviços públicos. Portanto, a disponibilização de dados públicos de interesse coletivo ou geral tem grande potencial para agregar valor ao negócio da administração pública e à própria sociedade, com incremento da participação social e empoderamento do cidadão, que se coadunam com os objetivos estratégicos constantes do plano estratégico do TCU, quais sejam “contribuir para a melhoria do desempenho da Administração Pública” e “contribuir para a transparência da Administração Pública”.

411. Assim, conclui-se que as informações presentes nesse levantamento se constituíram em importante fonte para subsidiar futuros trabalhos da Sefti e do Tribunal em torno da política de dados abertos, das iniciativas de governo aberto, da prestação de serviços de governo eletrônico e da melhoria da tomada de decisão na Administração, bem como da aferição do cumprimento dos termos da Lei de Acesso à Informação.

412. Diante disso, propor-se-á que a Sefti considere, no seu planejamento, a conveniência de inserir ações de controle para avaliação da política de dados abertos em órgãos e entidades da APF.

6. Propostas de Encaminhamento

413. Diante do exposto, submetem-se os autos à consideração superior com a seguinte proposta:

413.1. recomendar à Secretaria de Fiscalização de Tecnologia da Informação (Sefti) que, em atenção às possibilidades de atuação identificadas neste levantamento, examine a viabilidade da inclusão, em futuros planos de fiscalização, da realização de ações de controle voltadas à verificação da implementação da política de abertura de dados nas unidades da Administração Pública Federal, bem como da realização de ações de estímulo à conscientização dessas unidades acerca de conceitos, objetivos, legislação, indicadores, riscos e estruturas de governança relacionados ao tema dados abertos;

413.2. encaminhar cópia da deliberação que vier a ser adotada, bem como do relatório, do voto, do relatório da unidade técnica e dos respectivos anexos que a fundamentarem:

413.2.1. à Controladoria-Geral da União;

413.2.2. à Secretaria de Logística e Tecnologia da Informação (SLTI/MP);

413.2.3. à Casa Civil da Presidência da República;

413.2.4. à Secretaria-Geral da Presidência da República;

413.2.5. ao Serviço de Processamento de Dados do Governo Federal (Serpro);

413.2.6. à Empresa de Tecnologia e Informações da Previdência Social (Dataprev);

413.2.7. ao Departamento de Informática do Sistema Único de Saúde (Datasus);

413.2.8. ao Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep);

413.2.9. ao Conselho Nacional de Justiça;

413.2.10. ao Conselho Nacional do Ministério Público;

413.2.11. à Diretoria-Geral da Câmara dos Deputados;

413.2.12. à Diretoria-Geral do Senado Federal;

413.2.13. à Secretaria Geral da Presidência do Tribunal de Contas da União – Segepres/TCU;

413.3. levantar o sigilo deste processo, por conter informações relevantes às instituições públicas quanto ao tema dados abertos, à exceção das peças 33 e 34, uma vez que contêm documentos que não foram tornados públicos pela instituição proprietária;

413.4. arquivar o presente processo, com fulcro no art. 169, inciso V, do RITCU.”

2. O diretor e o titular da Sefti ratificaram a proposta consignada no relatório acima transcrito (peças 64 e 65).

É o relatório.