UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"

12ª Conferência Luso-Brasileira de Ciência Aberta - Formato Híbrido

DADOS ABERTOS NO SETOR PÚBLICO: ANÁLISE DA ENGENHARIA DE CONSTRUÇÃO

Prof. ª Dr.ª Marise Miglioli Lorusso UNESP– Universidade Estadual Paulista "Júlio de Mesquita Filho"
Prof. ª Dr.ª Silvana Aparecida Borsetti Gregorio Vidotti - UNESP– Universidade Estadual Paulista "Júlio de Mesquita Filho"

INTRODUÇÃO

O tema desta pesquisa situa-se no campo da denominada "ciência aberta", contextualizada na Ciência da Informação, enfatizando os dados abertos na esfera pública, com foco na esfera pública municipal de São Paulo.
Apesar da existência de legislação concernente à obrigatoriedade da apresentação dos "dados abertos" na esfera pública, a disponibilização dos mesmos sugere falta de continuidade em sua atualização e contínuas modificações em sua arquitetura. Essa arquitetura aparenta ter variado de acordo com as lideranças, cujos propósitos são divergentes. Como fontes de dados pesquisáveis, os denominados "dados abertos" deveriam ser disponibilizados periodicamente, principalmente quando voltados à prestação de contas no Portal da Transparência. Um pouco "fosca", essa transparência tem refletido uma arquitetura de dinâmica e solidez contestáveis, num fluxo que parece dispersar-se em meio a eventos que o desestabilizam. As lideranças, por sua vez, sendo altamente rotativas, sem a necessária experiência em projetos de arquitetura da informação, raramente dispendem seu tempo para refletir sobre os mecanismos de produção, gerenciamento e disponibilização dos dados abertos de suas respectivas pastas. A partir daí torna-se necessário refletir sobre o atual estágio dos dados abertos do setor público e pensar acerca de sua arquitetura bem como sobre o fluxo informacional que lhe dá sustentação.
De acordo com a *Open Data Science Initiative*, a ideia acerca da ciência de dados aberta é a disponibilização das novas metodologias de análise "tão ampla e rapidamente quanto possível, com o mínimo possível de condições de utilização" (OPENDSI, 2019) além de prever a capacitação dos usuários e o compartilhamento ágil das informações disponibilizadas e propor um equilíbrio entre o compartilhamento de dados para benefício social.
No que se refere à construção de "dados abertos", parece ser útil compará-la com o padrão seguido pelos engenheiros (ao elaborarem projetos para construção de edificações). Quando se constrói uma edificação é de vital importância conhecer sua localização, o tipo de solo do terreno onde será erguida, o fluxo das pessoas que circularão por ela e o material a ser utilizado.
Da mesma forma, estudos prévios analisam o impacto no setor onde atuam e as consequências de sua inserção. Em terreno adequado, os dados podem fluir e serem mantidos com razoável segurança. Seu conteúdo, desde que estruturalmente estável, pode perdurar, mediante as devidas atualizações, por tempo indeterminado.
Edificações e estruturas de dados apresentam semelhanças mais amplas que as metafóricas. Edifícios podem estar condenados à ruína, ao desabamento, antes mesmos de serem totalmente concluídos; os dados abertos, idem. Neste caso, não basta a questão interativa e visual por meio de *URLs*. É imprescindível um fluxo coerente, com tecnologia cuja obsolescência não seja verificada em curto espaço de tempo, além de solidez nas fontes fornecedoras da informação: matéria-prima para uma engenharia de dados robusta e funcional.
Da mesma forma que engenheiros planejam suas edificações com saídas de emergência em caso de imprevistos ou sinistros, uma construção de dados abertos deve levar em conta as possibilidades de construção de meios de entrada e saída alternativos, sempre com vistas à disponibilização do conteúdo à sociedade como um todo. Trata-se de uma garantia que não pode ser truncada por mudanças políticas, assim como um edifício pode mudar de proprietário e/ou síndico, mas suas bases estruturais continuam sólidas. Esta pesquisa toma por bases os dados abertos das esferas públicas brasileiras, com ênfase à Prefeitura do Município de São Paulo, constituindo-se em estudo de caso descritivo, com base na análise documental *webgráfica*, complementado por questionário voltado a dois grupos: utilizadores dos dados abertos e gerenciadores dos portais.

OBJETIVO

Com base nesse escopo, o problema de pesquisa buscou analisar a base de sustentação dos dados abertos na esfera pública, pela observação da existência ou não de um fluxo coerente e dinâmico, sua compatibilidade com os princípios supramencionados e se, dentro desse contexto, a Prefeitura do Município de São Paulo está com uma engenharia solidamente aplicada, apresentando menor vulnerabilidade as constates mudanças de gestão, próprias das esferas públicas.

METODOLOGIA

A pesquisa utiliza a abordagem qualitativa, aplicando o método de estudos de caso único, triangulado com a análise documental *webgráfica* e, quando necessário e possível, bibliográfica, dos *sites* das esferas federal, estadual e municipal, com ênfase à cidade de São Paulo e ao Governo do Estado de São Paulo (ASSEMBLEIA LEGISLATIVA DO ESTADO DE SÃO PAULO, 2019).
Foi analisado o processo de engenharia de construção, a circulação dos dados e a propagação das informações, com foco na Prefeitura da Cidade de São Paulo, levando-se em conta os aspectos referentes à informação cinzenta e àquela considerada dispersa pelo labirinto dos dados abertos.
O objeto da pesquisa é a esfera da Prefeitura da Cidade de São Paulo no contexto do Governo do Estado de São Paulo, como estudo de casos baseados na realidade dos dados abertos construídos e alimentados pelo poder público municipal, contextualizado na macrorregião do Governo do Estado de São Paulo.
Pela análise do objeto de pesquisa foram considerados os atores envolvidos no processo: sociedade civil (munícipes e instituições diversas), fornecedores e parceiros além dos próprios servidores que necessitam consultar dados internos para construção de novos projetos.

RESULTADOS E DISCUSSÃO

A pesquisa teve como ponto de partida a questão do conhecimento prévio (ou não) dos dados abertos, seu conceito, sua acessibilidade e sua interoperabilidade, tomando em conta o ponto de vista do usuário dos dados.
Para essa prévia, formulou-se um questionário simplificado com seis questões tipo sim/não, cuja temática girou em torno da experiência de cada usuário com os dados abertos da macrorregião de São Paulo. Os usuários foram escolhidos aleatoriamente, a partir do efeito 'bola de neve', onde cada respondente tem a possibilidade de repassar à própria rede de contatos com o perfil solicitado, a fim de colaborar com a pesquisa. Todos os respondentes deveriam estar cursando ou serem egressos de instituições de nível universitário, públicas ou privadas.
Do total de respondentes, 62% afirmaram ter conhecimento (ou noções) do conceito de "dados abertos", enquanto 38% demonstraram não saber de que se tratava; no que se refere à utilização dos conteúdos dos dados abertos, 73% dos respondentes afirmaram não utilizar os conteúdos disponíveis nas bases que consultaram, enquanto 27% afirmaram fazer usos deles; quanto a necessidade ou não dos conteúdos dos dados abertos, 49% afirmaram ter necessidade de dados abertos contra 51% que não sentem essa necessidade; dos 125 respondentes que afirmaram utilizar os dados abertos, 34% consideraram os dados como relevante e/ou pertinentes as suas pesquisas, enquanto 66% afirmaram o oposto; os mesmos 125 respondentes foram questionados sobre a similaridade ou padronização nas bases de sua pesquisa e 34% dos mesmos afirmaram ter encontrado semelhança no modo de operar as bases, enquanto 66% julgaram ter que 'aprender' a operar a base a cada pesquisa feita; finalmente, quando demandados sobre a atualização dos dados, 11% declararam que os dados são atualizados, 30% declararam que os dados não são atualizados e 59% dos respondentes que utilizaram as bases não conseguiram identificar.
Esses dados já apontavam para algumas contradições como, por exemplo, o fato de parte dos respondentes afirmar seu desconhecimento sobre o conceito de dados abertos ao mesmo tempo em que consideram necessária sua utilização.
Do resultado das entrevistas foi possível ter uma percepção parcial dos integrantes das equipes que deliberam, trabalham ou simplesmente alimentam as bases de dados abertos a serem disponibilizadas no Portal da Prefeitura do Município de São Paulo.
Em primeiro lugar, destaca-se a questão do tempo de permanência do funcionário dentro do órgão público. Essa informação é primordial quando se fala em 'alicerçamento sólido' dos dados abertos em instituições governamentais. No resultado apresentado é possível perceber que a totalidade dos servidores envolvidos (entrevistados) pertence à categoria de 'comissionados' (100%) e possui indicação de instância superior (100%). Isso equivale a dizer que não existem garantias de permanência ou continuidade de seus trabalhos em múltiplas gestões governamentais, tendo em visto serem 'cargos de confiança', sujeitos a alterações, extinções e a contínuas mudanças de unidades. Embora cargos efetivos também possam passar por tais mudanças, a celeridade do processo é menor, tendo em vista a legislação vigente.
A capacitação para o pleno desempenho das atividades laborativas junto às bases de dados é parcialmente realizada (76%), voltada para a alimentação das bases (88%) e não para a compreensão da totalidade do fluxo (59%). O desconhecimento demonstrado por alguns pontos focais no que se refere à suas atividades com a possibilidade de relacionamento ao fluxo de dados abertos sugere uma 'fenda' na solidez do processo (29%). A maioria dos entrevistados (59%) admite desconhecer o fluxo das informações para o sistema de dados abertos. Mesmo aqueles que obtiveram capacitação apresentaram dúvidas entre teoria e prática.
A reutilização de dados abertos não é visível (71%) para os entrevistados e não possui registro no Portal da Transparência da Prefeitura do Município de São Paulo. Esse fato pode ser comprovado também nos Portais do Governo do Estado de São Paulo e Portal de Dados do Governo Federal.

CONSIDERAÇÕES FINAIS

A abertura de bases de dados governamentais pode trazer uma série de benefícios para diferentes setores da sociedade além do próprio órgão público que decide abrir suas bases. Estruturando-os de forma adequada e publicando-os sistematicamente, com procedimentos permanentes e atualizáveis, o setor público pode auxiliar no processo de transparência e garantir a ampliação da disponibilidade das informações de caráter até então visível somente em sua estrutura interna.
"Não é possível administrar, ou conhecer, aquilo que não possui indicadores para medição." A abertura dos dados públicos, em suas diversas esferas, possibilita conhecer e gerar parâmetros para medição de eficiência/eficácia de atividades. Também possibilita que pesquisadores, independentes ou vinculados a uma instituição, tenham acesso a informações sobre a formulação e execução de políticas públicas, bem como sobre a aplicação dos recursos públicos.
Muitas bases paralelas, com dados divergentes, categorizados de forma aleatória apontam critérios de cientificidade duvidosa, amparados, muitas vezes, pela dimensão particularizada das políticas públicas vigentes.
Com relação à análise do fluxo informacional é possível perceber que, embora a trajetória procure cumprir a legislação vigente, não fica evidenciada a clareza do modelo de construção aqueles que participam e/ou contribuem para o processo de geração dos dados abertos.
Nos aspectos de disseminação, existem 195 conjuntos de dados distribuídos pelos diversos órgãos das administrações direta e indireta da Prefeitura de São Paulo, com nenhum seguidor para todos. Na observação de cada uma das setenta e cinco instituições registradas no Portal foi percebida ênfase à 'folha de pagamento' em detrimento de outras informações relevantes.
A eficácia da disseminação dos dados abertos para a consolidação de informações na esfera pública carece de divulgação nas redes sociais e em mídias externas, bem como não parece atingir os servidores (envolvidos ou não), como mostrou a pesquisa aleatoriamente realizada como prévia.
No contexto das demais esferas, a PMSP aparenta cumprir os protocolos federais e estaduais no que se refere à escolha de plataformas, formatos e demais recursos técnicos, mas seu entendimento da necessidade de disseminar seus dados por meio de estruturas sólidas, com registros permanentes ainda parece longe de se tornar uma realidade.
Essa realidade conduz à percepção de que existem 'fissuras' estruturais na engenharia de dados abertos da PMSP, tanto no que tange aos critérios de escolha dos pontos focais como no que diz respeito aos critérios e planejamento do processo para escolha e abertura dos dados.
A legislação faz referência sobre o caráter de exceção dos dados que envolvem "preservação da pessoa", respeitando seus "dados bancários, sigilosos e pessoais", mas a disponibilização, em larga escala, da folha de pagamento como primeira opção de dado aberto, em detrimento de dados relevantes, reutilizáveis, parece atender aspirações diversas daquelas preconizadas pela Lei.
Como em toda a fundação da engenharia, é importante evitar as 'infiltrações', causadoras de graves danos no médio e longo prazo, conforme sua dimensão. Essas 'infiltrações' ocorrem na esfera pública quando os dados são gerados em função de modelos de gestão pública ao invés de serem construídos em processo sólido, sem intervenções de cunho pessoal.
Tais dados também deveriam analisar a manutenção das plataformas e sistemas vigentes, sem o risco de paralisação por motivos orçamentários, políticos ou temporais. Assim, o resguardo dos dados em mecanismos disponíveis com garantias será sempre mais oportuno do que as 'soluções caseiras' que possuem tempo de validade e cujos responsáveis não são permanentes nem deixam equipes montadas.
Assim, uma engenharia baseada apenas nos parâmetros e na tecnologia gerada por viés político tenderá a sucumbir tão logo esse viés rume por outra direção. Esse fato torna imprescindível uma engenharia de alicerces fundamentados na perenidade, solidamente edificados para servirem eficazmente ao fluxo de informações que, pelo menos reduzam a efemeridade sugerida pela contínua mudança de políticas públicas para o setor.

REFERÊNCIAS

ASSEMBLEIA LEGISLATIVA DO ESTADO DE SÃO PAULO. **Portal dos dados abertos**. Disponível em: <https://www.al.sp.gov.br/dados-abertos/> Acesso em 25.05.2019
BRASIL. CONTROLADORIA GERAL DA UNIÃO. Guia de implantação do Portal da Transparência. Disponível em: < http://www.cgu.gov.br/Publicacoes/transparencia-publica/brasil-transparente/arquivos/guia_portaltransparencia.pdf>_Acesso em 22.05.2019
BRASIL. CONTROLADORIA GERAL DA UNIÃO. Guia sobre informações qualificadas. Disponível em: < http://www.acessoainformacao.gov.br/lai-para-sic/sic-apoio-orientacoes/guias-e_orientacoes/guia_informacoesclassificadas.pdf/@@download/file/Guia_InformacoesClassificadas.pdf>_Acesso em 22.05.2019
BRASIL. CONTROLADORIA GERAL DA UNIÃO. **Manual de lei de acesso à informação para Estados e municípios**. Disponível em: < http://www.cgu.gov.br/Publicacoes/transparencia-publica/brasil-transparente/arquivos/manual_lai_estadosmunicipios.pdf>_Acesso em 21.05.2019
BRASIL. **Kit de dados abertos**: infraestrutura nacional de dados aberto. Disponível em: <http://kit.dados.gov.br/> Acesso em 31.05.2019
BRASIL. Leis e decretos. **Portal da transparência**. Disponível em: <http://www.portaltransparencia.gov.br/> Acesso em 21.05.2019
BRASIL. LEIS E DECRETOS. **Lei de Acesso à Informação**: Lei nº 12.527, de 18 de novembro de 2011. Disponível em: < http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/l12527.htm>_Acesso em 21.04.2019
COMMAseparated values. Disponível em: <CSV: https://en.wikipedia.org/wiki/Comma-separated_values> Acesso em 22.05.2019
EAVES, David. **Creating Municipalities that work like the web, in municipal interface**. 2009. Disponível em: < https://eaves.ca/publications/> Acesso em 21.07.2019
FUNDAÇÃO GETÚLIO VARGAS. **Lançamento de dados abertos para São Paulo**. Disponível em: <http://dapp.fgv.br/dapp-e-open-knowledge-lancam-indice-de-dados-abertos-para-sao-paulo/> Acesso em 31.05.2019
GELLMAN, Robert. **Privacy and biometric ID systems**: an approach using fair information practices for developing countries. Disponível em: <https://okfn.org/> Acesso em 31.05.2019
GEOJSON. The Geojson specification (RFC 7946). Disponível em: <geoJSON: http://geojson.org/>_Acesso em 31.05.2019
GOOGLE DEVELOPERS. **Introdução a documentação KML**. Disponível em: <https://developers.google.com/kml/documentation/?hl=pt-br>_Acesso em 24.04.2019
KRANTZ, Peter. **Publishing open data: do you really need an API?** Disponível em: <https://www.peterkrantz.com/2012/publishing-open-data-api-design/>_Acesso em 24.05.2019
MERRIAM, S. B. **Qualitative research and case study applications in education**. San Francisco: Allyn and Bacon, 1998.
MERRIAM, Sharam B. Case studies as qualitative research. In: **Qualitative research and case study applications in education**. 2. ed. rev and exp. California: Jossey-Bass, 1998. p. 26-43
NEVINSKI, Felipe. **Dados "meio" abertos**: sobre o uso e reuso dos dados governamentais brasileiros. Open Knowledge Foundation, 2013. Disponível em: <http://br.okfn.org/2013/08/28/dados-meio-abertos-sobre-o-uso-e-reuso-dos-dados-governamentais-brasileiros/>_Acesso em 25.04.2019
OPEN DATA SCIENCE INICIATIVE. Disponível em < http://opendsi.cc/about/> Acesso em 21.06.2019
OPEN DEFINITION. **Open definition 2.1**: open knowledge. Disponível em: <http://opendefinition.org/od/index.html>_Acesso em 21.05.2019
OPEN KNOWLEDGE FOUNDATION. **Guidelines on open government data for citizen engagement.** Disponível em: <https://blog.okfn.org/2012/09/28/open-data-and-access-to-information-advocates-unite/> Acesso em 31.05.2019
OPEN KNOWLEDGE FOUNDATION. **Open data handbook**. Disponível em: <http://opendatahandbook.org/> Acesso em 31.05.2019
SÃO PAULO, SP. **Dados abertos**. Disponível em: <http://dados.prefeitura.sp.gov.br/pt_PT/> Acesso em 21.05.2019
SÃO PAULO, SP. **Governo aberto**. Disponível em: <http://www.governoaberto.sp.gov.br/> Acesso em 21.05.2019
SÃO PAULO, SP. **Transparência na PMSP**. Disponível em: <http://transparencia.prefeitura.sp.gov.br/Paginas/home.aspx> Acesso em 21.05.2019
SÃO PAULO. **Base de dados do Governo do Estado de São Paulo**. Disponível em: <http://dados.gov.br/dataset?tags=SP> Acesso em 21.05.2019
SÃO PAULO. **Base de dados da Secretaria Estadual de Educação**. Disponível em: <https://dados.educacao.sp.gov.br/> Acesso em 21.05.2019
SÃO PAULO. **Leis e decretos**. Disponível em: <http://www.saopaulo.sp.leg.br/transparencia/dados-abertos/> Acesso em 21.05.2019
SÃO PAULO. **Transparência de dados**. Disponível em: <http://www.transparencia.sp.gov.br/> Acesso em 21.05.2019
STUDER, R.; BENJAMINS, R. R.; FENSEL, D. Knowledge engineering: principles and methods. **Data & Knowledge Engineering**, v. 25, n. 1-2, p. 161-197, 1998. Disponível em: https://www.sciencedirect.com/science/article/pii/S0169023X97000565. Acesso em: 23.01.2019.
STAKE, R. E. **The art of case study research**. Thousand Oaks, CA: Sage, 1995.
STAKE, R. E. The case study method in social inquiry. In R. Gomm, M. Hammersley, and P. Foster (Eds.), **Case study method:** Key issues, key texts. Thousand Oaks, CA: Sage, 2000. p. 20-26.
STAKE, R. E. Qualitative case studies. In N. K. Denzin, & Y.S. Lincoln (Eds.), **The handbook of qualitative research**. 3rd ed. Thousand Oaks, CA: Sage, 2005. p. 443-466
STAKE, R. E. (2008). Qualitative case studies. In N. K. Denzin, & Lincoln, Y. S. (Eds.), **Strategies of qualitative inquiry**. Los Angeles: Sage, 2008. P. 119-149
SUNLSUNLIGHT FOUNDATION. Making government and politics more accountable and transparent. Disponível em: <https://sunlightfoundation.com/>_Acesso em: 25.06.2019
SYMPLA. **Open data day – por que dados abertos?** Disponível em: <https://www.sympla.com.br/por-que-dados-abertos---open-data-day__475545> Acesso em 20.05.2019
TAUBERER, Joshua. **Open government data**: the principles, practices, and a history of the open government data movement. E-Book Disponível em: <https://opengovdata.io/>_Acesso em 23.04.2019
THE WORLD BANK. **Guide pratique des données ouvertes.** Disponível em: <http://opendatatoolkit.worldbank.org/fr/>_Acesso em 27.04.2019
UNITED KINGDON. **Datasets**. Disponível em: <http://data.gov.uk/code-conducl> Acesso em 23.05.2019
UNITED KINGDON. HAMPSHIRE COUNTY CONUNCIL. **Hampshire county open licence**. Disponível em: <http://www3.hants.gov.uk/opendata/licence.htm>_Acesso em 23.04.2019
UNITED KINGDON. THE NATIONAL ARCHIVES. **Open government licence for public sector information. Disponível em:** < http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/>_Acesso em 24.05.2019
UNITED STATES OF AMERICA. **Open data project**. Disponível em: < https://project-open-data.cio.gov/license-examples/>_Acesso em 23.04.2019
VICTORINO *et al*. Uma proposta de ecossistema de Big data para a análise de dados abertos governamentais conectados. **Inf. & Soc.:Est**., João Pessoa, v.27, n.1, p. 213-230, jan./abr. 2017
W3C. **Extensive markup language**: XML. Disponível em: <http://www.w3.org/XML/>_Acesso em 21.04.2019
YIN, R. K. **Estudo de caso**: planejamento e métodos. 4. ed. Porto Alegre: Bookman, 2010.
Yin, R. K. (1999). Enhancing the quality of case studies in health services research. **Health Services Research**, v. 34, n. 5, p. 1209-1224.
Yin, R. K. (2003a). **Applications of case study research**. 2nd ed.. Thousand Oaks, CA: Sage.
Yin, R. K. (2005). Introduction. In R. K. Yin (Ed.), **Introducing the world of education**: a case study reader Thousand Oaks, CA: Sage. P. 13-22

E-mail para contato com os autores: Marise Miglioli Lorusso marise.m.lorusso@uneap.br
Silvana Vidotti silvana.vidotti@unesp.br