

Recuperação de dados abertos para inovação cívica e ciência cidadã: uma análise do novo portal brasileiro

Open data recovery for civic innovation and citizen science: an analysis of the new brazilian portal

Patrícia Nascimento Silva

¹ Universidade Federal de Minas Gerais, Belo Horizonte-MG, Brasil. ORCID: <https://orcid.org/0000-0002-2405-8536>

Autor para correspondência/Mail to: Patrícia Nascimento Silva, patricians.prof@gmail.com

Recebido/Submitted: 09 de janeiro de 2023; **Aceito/Approved:** 04 de março de 2024



Copyright © 2024 Nascimento Silva. Todo o conteúdo da Revista (incluindo-se instruções, política editorial e modelos) está sob uma licença Creative Commons Atribuição 4.0 Internacional. Ao serem publicados por esta Revista, os artigos são de livre uso para compartilhar e adaptar e é preciso dar o crédito apropriado, prover um link para a licença e indicar se mudanças foram feitas. Mais informações em <http://revistas.ufpr.br/atoz/about/submissions#copyrightNotice>.

Resumo

Introdução: Este short paper provoca uma discussão urgente sobre o acesso, a recuperação e o reuso de dados abertos no país, impactados diretamente pela reestruturação do Portal Brasileiro de Dados Abertos, ocorrida no final de novembro de 2022. O Portal foi modificado sob a justificativa de facilitar o consumo de dados pelos usuários. Diante da importância desse instrumento para a Política de Dados Abertos do país, sobretudo para a inovação cívica, esse estudo fez uma análise do novo Portal sob a óptica da recuperação de Informação. **Método:** Para tanto foi realizada uma pesquisa descritiva e exploratória no novo Portal mapeando novos recursos de recuperação de informação e uma comparação com as funcionalidades existentes no Portal antigo. **Resultados:** Na análise realizada em dezembro de 2022 foi possível observar que grande parte dos recursos do Portal antigo foram mantidos, porém apresentam falhas. Novos recursos foram criados, mas não estão em conformidade com os princípios dos dados abertos e boas práticas existentes. **Conclusão:** Apesar de uma arquitetura minimalista e agradável ao usuário, o Portal perdeu funcionalidades essenciais de recuperação de informação, como a ordenação dos resultados, e apresenta inconsistências em recursos de filtragem. Espera-se que este trabalho contribua para direcionar correções e melhorias basilares, neste instrumento tão importante para o efetivo acesso à informação pela sociedade.

Palavras-chave: Dados abertos; Recuperação de informação; Inovação cívica; Ciência cidadã; Portal Brasileiro de Dados Abertos.

Abstract

Introduction: This short paper provokes an urgent discussion about access, recovery and reuse of open data in the country, directly impacted by the restructuring of the Brazilian Open Data Portal, which took place at the end of November 2022. The Portal was modified under the justification of facilitating data consumption by users. Given the importance of this instrument for the country's Open Data Policy, especially for civic innovation, this study analyzed the new Portal from the perspective of information retrieval. **Method:** For this purpose, a descriptive and exploratory research was conducted in the new Portal, mapping information retrieval resources and a comparison with the existing functionalities in the old Portal. **Results:** In the analysis conducted in December 2022, it was possible to observe that most of the features of the old Portal were maintained, but they have flaws. New features have been created but are not in line with existing open data principles and best practices. **Conclusion:** Despite a minimalist and user-friendly architecture, the Portal has lost essential information retrieval features, such as sorting results, and presents inconsistencies in filtering resources. It is hoped that this work will contribute towards directing basic corrections and improvements in this important instrument for effective access to information by society.

Keywords: Open data; Information retrieval; Civic innovation; Citizen science; Brazilian Open Data Portal.

INTRODUÇÃO

O Portal Brasileiro de Dados Abertos é um dos instrumentos mais importantes da Política de Dados Abertos do Brasil, instituída em 2016, por meio do Decreto Nº 8.777. Nesta fonte são catalogados os conjuntos de dados abertos de todos os órgãos do Governo Federal, ou seja, dados governamentais abertos (DGA). Os dados abertos são dados com formatos e licenças abertas e disponíveis na *internet*, livre para utilização pela sociedade [Agune, Gregório Filho, e Bolliger \(2010\)](#). Os DGA, ou dados abertos governamentais, são dados abertos, sobre o governo e suas atividades. O Portal é visto como uma fonte de qualidade, fácil acesso e fomento a participação mais articulada e ativa da sociedade no processo político, elementos marcantes na inovação cívica [Open Knowledge Brasil \(2022\)](#).

A inovação cívica é entendida como uma forma de inovação para as cidades baseada na participação cidadã e em novas formas de governança urbana aliadas à tecnologia. Quando associada ao modelo de conhecimento e inovação, “tríplice hélice” (academia, mercado e governo), pesquisadores como [Carayannis e Campbell \(2011\)](#) indicam uma evolução para um modelo baseado na “quádrupla hélice” que considera a sociedade do conhecimento e a inclusão da sociedade civil. Desta forma, os cidadãos são convidados a participar, contribuindo com seu conhecimento e criatividade para geração de inovação [Depiné \(2018\)](#).

As iniciativas de ciência cidadã têm se consolidado como uma nova forma de interação entre os cientistas profissionais e os cidadãos. É uma das vertentes do movimento de ciência aberta e permite os cidadãos participem de atividades formais de pesquisa desenvolvidas nos principais centros de investigação do mundo, propiciando engajamento de não cientistas com a ciência, com o método científico e com novas possibilidades educacionais que

podem ser colocadas à disposição da sociedade. O indivíduo participante pode ter uma atividade de aprendizado ou lazer ao mesmo tempo em que realiza atividades de produção e análise de conhecimentos, coleta de dados e informações relevantes à ciência e às questões de importância social, ambiental e territorial (Instituto Brasileiro de Informação em Ciência e Tecnologia [Ibict], (2024); Universidade de São Paulo [USP], 2024).

O Portal de Dados Abertos foi lançado oficialmente em 2012, para ser o ponto referencial de acesso e recuperação de dados públicos no Brasil. Seu conteúdo foi dividido em conjuntos de dados, contendo um ou mais recursos sobre temas relacionados. Neste primeiro momento, o instrumento disponibilizava poucos conjuntos de dados e cumpria em partes com os princípios e recomendações de melhores práticas para publicação de dados abertos.

Com a instituição da Política de dados Abertos, em 2016, a publicação dos conjuntos de dados no Portal tornou-se compulsória e crescente, demandando uma maior estruturação do instrumento, principalmente em relação à organização e recuperação dos dados publicados. Em congruência com outras experiências internacionais, o Portal não tinha a função de armazenar arquivos de dados e sim de ser um catálogo, classificado segundo metadados que facilitam a sua localização e recuperação a partir do local em que estão hospedados pelos órgãos Possamai (2016).

A recuperação de informação ocupa-se da representação, do armazenamento, da organização e do acesso a itens de informação, como documentos, páginas Web, catálogos online, registros estruturados e semiestruturados, objetos multimídias, etc. Baeza-Yates e Ribeiro-Neto (2013). Com isso, alterações na organização e arquitetura dos dados e informações impactam diretamente nas formas de recuperação de informação.

Em 25 de novembro de 2022, foi publicado o novo Portal Brasileiro de Dados Abertos, uma surpresa para os cidadãos, que não vivenciaram um período de transição ou migração. Nesta nova versão, a arquitetura do Portal foi alterada, modificando a organização e impactando na recuperação de informação pelos usuários. Diante dessa mudança e sua repercussão na sociedade, esta pesquisa buscou mapear as alterações relacionadas à recuperação de informação.

Para tanto, foi feito um breve estudo do novo Portal, identificando e mapeando as novas funcionalidades de recuperação de informação, e um estudo comparativo com a versão anterior do Portal. O questionamento que deu início ao estudo foi: Quais são as alterações no novo Portal de Dados Abertos? O estudo foi motivado pela grande quantidade de reclamações de usuários, manifestado nos comentários de *lives* e eventos de divulgação do novo Portal e justifica-se por ser investigar um dos instrumentos mais relevantes da Política de Dados Abertos do país que permite o acesso e o reuso dos dados abertos pela sociedade.

PROCEDIMENTOS METODOLÓGICOS E RESULTADOS

A pesquisa é caracterizada, em relação ao objetivo, como descritiva e exploratória, com abordagem qualitativa. O instrumento utilizado foi o novo Portal Brasileiro de Dados Abertos, disponível no endereço: <https://dados.gov.br/Brasil> (2022a), em dezembro de 2022, em uma análise das funcionalidades e interfaces sob a óptica dos recursos de recuperação de informação.

Inicialmente foram mapeadas as funcionalidades relacionadas à recuperação da informação existentes no Portal antigo, que foi disponibilizado em um endereço temporário após a migração para o novo Portal: Brasil (2022b). É importante destacar que no novo Portal não há nenhuma informação sobre este endereço disponibilizado para o Portal antigo. A autora obteve o acesso durante uma *live* realizada pela Controladoria-Geral da União, para divulgar o novo Portal, em dezembro de 2022. As funcionalidades existentes em relação à recuperação da informação, conforme apresentado Quadro 1, estavam relacionadas a: (1) Busca, (2) Filtros e (3) Ordenação dos resultados, critérios que refletiam classificações temáticas dos metadados dos conjuntos de dados catalogados no Portal e técnicas de recuperação de informação como sistemas de filtragem. Após a definição dos critérios foi iniciada a análise do novo Portal, identificando e analisando cada um dos recursos mapeados.

Funcionalidade	Nome/Tipo	Descrição do recurso
(1) Busca	Busque no Portal	Pesquisa aberta por texto.
(2) Filtros	Organizações	Filtro contendo todos os órgãos que publicam dados no Portal.
	Temas	Representação temática em 21 categorias.
	Etiquetas	Etiquetas diversas com palavras-chave aleatórias.
	Formatos	Todos os tipos de formatos existentes para os conjuntos de dados disponibilizados.
	Licenças	Todas as licenças existentes para os conjuntos de dados disponibilizados.
(3) Ordenação	Ordenar por	Opções de ordenação por: Relevância, Nome Crescente, Nome decrescente, modificada pela última vez.

Quadro 1. Critérios de análise e resultados no novo Portal
Fonte: Elaborado pela autora (2023).

No primeiro critério analisado, (1) Busca: observou-se uma inovação com relação a localização e ao destaque da funcionalidade, que é apresentada na página principal como primeiro recurso revelado ao usuário. Por ser uma busca aberta, não há nenhum tipo de informação, dica ou direcionamento para o usuário. Somente informações sobre o total de conjuntos de dados, recursos, reuso, organizações e usuários, sem nenhum detalhe ou opção de filtro. Destaca-se que não é possível realizar uma busca sem informar um texto, desta forma, o usuário que não tem familiaridade com o nome dos conjuntos de dados poderá encontrar dificuldades para realizar a pesquisa. No Portal antigo era possível visualizar todos os conjuntos de dados, categorizados pelos filtros (indicados no Quadro 1), na opção “Dados”, que era apresentada na página inicial do Portal.

Ao realizar uma busca textual, o Portal recuperou os totais de Conjuntos De Dados, Organizações, Temas, Reuso, Notícias, Discussões e apresentou os conjuntos de dados relacionados. Ao selecionar um conjunto de dados específico foram exibidas informações sobre a descrição, recursos, ações, conjunto de dados similares e discussões. O novo *layout* representa até 15 conjuntos de dados por página, por meio de uma área quadrada e clicável para acessar os detalhes. Com isso, para buscas com muitos resultados o usuário terá que navegar por várias páginas para acessar o recurso desejado.

É importante destacar que a busca geral do novo Portal recupera conjuntos de dados que tenham qualquer similaridade com o texto informado e não somente com o título do conjunto de dados. Uma consequência dessa grande amplitude da busca (alta revocação) é a recuperação de conjuntos de dados que não possuem nenhuma relação direta com o texto informado (baixa precisão). Em um exemplo realizado com a palavra “moeda” foram recuperados conjuntos de dados relacionados como: “Taxas de câmbio - todos os boletins diários” e conjuntos de dados distintos como: “lista de Participantes do STR”, todos em uma mesma página sem indicação de ranking, relevância ou ordenação.

O segundo critério (2) filtros: foi identificado no novo Portal, no entanto parece estar “encoberto” para os usuários. Após procurar em várias páginas do Portal, os filtros foram localizados (Figura 1) ao clicar na opção “Conjunto de dados”, visualizada na página de detalhe de um conjunto de dados pesquisado ou no rodapé do Portal, e depois na opção “Mais opções de filtros”.



Figura 1. Funcionalidade Conjunto de Dados

Fonte: Recuperado do Portal Brasileiro de Dados Abertos [Brasil], 2022

Na página principal do Portal não é possível acessar a opção de filtros, de forma direta, para auxiliar as buscas do usuário. Na funcionalidade de conjuntos de dados os seguintes filtros foram identificados:

- Tema: apresenta as 21 categorias conforme o filtro Temas do Portal antigo;
- Formato: inclui uma lista extensa de formatos, mas não recupera dados corretos. Em uma busca pelo formato “api” os conjuntos de dados do Banco Central, órgão que disponibiliza diversos conjuntos neste formato, não foram recuperados;
- Organização: apresenta as organizações cadastradas, similar ao filtro Organizações do Portal antigo. Contudo, uma vez selecionado não há opção para limpar o filtro;
- Palavra-chave: apresenta inúmeras palavras, sem uma descrição ou documentação para o usuário, é similar ao filtro de etiqueta do Portal antigo;

- Reuso: apresenta as opções de possuir ou não reuso. No momento da pesquisa não foi identificado nenhum reuso cadastrado.
- Licença: apresenta as licenças, similar ao filtro Licenças do Portal antigo.

O terceiro critério (3) Ordenação: não foi identificado em nenhuma página do novo Portal. Os resultados das buscas são apresentados por página e sem o recurso de ordenação, assim a manipulação dos resultados com vários conjuntos de dados torna-se impraticável para o usuário. Também não foi identificada nenhuma função de exportação dos resultados, opção que permitiria o tratamento em outra ferramenta, externa ao Portal, definida pelo usuário.

Foram identificados novos recursos e funcionalidades na apresentação dos conjuntos de dados para os usuários, a saber: Seguir, Avaliar, Contato (organização relacionada ao conjunto de dados), Ações (Sugerir *tag*, Reportar um problema, Sugerir melhoria), Discussões (novo tópico). Todos estes recursos, a princípio, permitem a interação do usuário, mas ao acessá-los é exigida a autenticação com uma conta GOV BR. Conforme apontado por Silva (2022), este tipo de exigência já foi observado em outros instrumentos da Política de Dados Abertos do Brasil, e está em oposição ao sexto princípio dos dados governamentais abertos. O acesso aos dados deve ser não discriminatório, ou seja, os dados devem estar disponíveis para todos, sem exigência de requerimento ou cadastro. Uma possível solução para manter a interação, sem identificação dos usuários, seria o gerenciamento das postagens e mensagens anônimas enviadas, com a avaliação prévia dos administradores antes da publicação.

O Conjuntos de dados similares também foi identificado como um novo recurso, mas não pode ser avaliado, já que não possuía informações cadastradas. Também não foi identificada nenhuma “Discussão” publicada, dessa forma ainda não se sabe se além da autenticação, o usuário terá sua identidade divulgada ao participar de uma discussão no novo Portal. Métricas como quantidade de seguidores e *downloads*, escala de satisfação e localização são indicados nos detalhes dos conjuntos de dados, no entanto por serem recentes também estavam sem informações. As métricas são relevantes e podem contribuir para o monitoramento da Política de Dados Abertos no país, principalmente em relação ao controle social e uso da sociedade.

CONCLUSÕES

Este *short paper* objetivou fazer uma breve análise do novo Portal Brasileiro de Dados Abertos sob a óptica da recuperação de informação. Foi identificada a criação de recursos, assim como a manutenção de funcionalidades, no entanto inconformidades foram identificadas em relação à recuperação de dados e informação.

Com relação ao critério (1) busca: na página inicial do portal é disponibilizada somente uma busca simples, por texto, sem outros recursos para recuperação de informação. As buscas por texto figuram considerar muitos metadados dos conjuntos de dados, inclusive descrições e conceitos, o que implica diretamente na diminuição da precisão e no aumento da revocação das consultas. Com relação critério (2) filtros: os mesmos não são apresentados de forma direta para o usuário e não recuperam todos os dados catalogados no Portal, apresentando inconsistências. Por fim, no critério (3) ordenação: os resultados das buscas não podem ser ordenados e sua visualização, paginada com poucos elementos, dificulta a manipulação de resultados numerosos. Foram identificados novos recursos, que podem fomentar o diálogo com a sociedade, mas a exigência a identificação dos usuários poderá comprometer o uso.

Atividades envolvendo a inovação cívica estão diretamente relacionadas à recuperação de informação em instrumentos como o Portal de Dados Abertos e até o final de dezembro de 2022 o instrumento ainda apresentava muitas inconsistências, impactando diretamente no acesso à informação por cidadãos e pesquisadores de todo o país. Em março de 2023 a Controladora Geral da União (CGU) divulgou o ajuste de mais de 3000 conjuntos de dados, mas sem alterações nas funcionalidades aqui analisadas. Uma *Application Programming Interface* (API) do novo Portal também foi lançada em 2023, documentada na ferramenta Swagger: <https://dados.gov.br/swagger-ui/index.html>. A disponibilização desse serviço contribui diretamente com a recuperação no Portal e fortalece ações de interoperabilidade no contexto do governo digital do Brasil. Espera-se que este estudo contribua para direcionar outras correções e melhorias basilares na recuperação de informação deste instrumento tão importante para o efetivo acesso à informação pela sociedade.

Agradecimento A autora agradece a Universidade Federal de Minas Gerais pelo apoio à pesquisa no Projeto: 30201*42.

REFERÊNCIAS

- Agune, R. M., Gregório Filho, A. S., & Bolliger, S. P. (2010). *Governo aberto sp: disponibilização de bases de dados e informações em formato aberto*. Recuperado de https://www.prefeitura.sp.gov.br/cidade/secretarias/upload/controladoria_geral/arquivos/C3_TP_GOVERNO%20ABERTO%20SP%20DISPONIBILIZACAO%20DE%20BASES%20DE%20DADOS.pdf
- Baeza-Yates, R., & Ribeiro-Neto, B. (2013). *Recuperação de informação: conceitos e tecnologia das máquinas de busca* (2a. ed.). Porto Alegre: Bookman.
- Brasil. (2022a). *Portal brasileiro de dados abertos*. Recuperado de <https://dados.gov.br/home>
- Brasil. (2022b). *Portal brasileiro de dados abertos*. Recuperado de <https://dev.dados.gov.br/>
- Carayannis, E. G., & Campbell, D. F. J. (2011). Open innovation diplomacy and a 21st century fractal research, education and innovation (freie) ecosystem: building on the quadruple and quintuple helix innovation concepts and the “mode 3” knowledge production system. *Journal of the Knowledge Economy*, 2(3), 327–372.
- Depiné, A. (2018). *Inovação cívica: a tecnologia como aliada da cidadania*. Recuperado de <https://via.ufsc.br/wp-content/uploads/2018/12/revistaVIA-5ed.pdf>
- Open Knowledge Brasil. (2022). *Inovação cívica*. Recuperado de <https://ok.org.br/eixos-de-atuacao/inovacao-civica/>
- Possamai, A. J. (2016). Portal brasileiro de dados abertos: novas práticas para o fortalecimento da democracia e da gestão pública na era digital. In S. P. d. Silva, R. C. Bragato, & R. C. Sampaio (Eds.), *Democracia digital, comunicação política e redes: teoria e prática* (p. 365-392). Porto Alegre: Folio Digital; Letra e Imagem. Recuperado de <https://portolivre.fiocruz.br/democracia-digital-comunicacao-politica-e-redes-teoria-e-pratica>
- Silva, P. N. (2022). Acesso à informação no Brasil: política de dados governamentais abertos. In E. e. a. Geraldês (Ed.), *Dez anos da lei de acesso à informação: limites, perspectivas e desafios* (p. 689-703). São Paulo: Intercom. Recuperado de https://www.researchgate.net/publication/374506630_Acesso_a_informacao_no_Brasil_politica_de_dados_governamentais_abertos

Como citar este shortpaper (APA):

Nascimento Silva, Patrícia (2024). Recuperação de dados abertos para inovação cívica e ciência cidadã: uma análise do novo portal brasileiro. *AtoZ: novas práticas em informação e conhecimento*, 13, 1 – 7. Recuperado de: <http://dx.doi.org/10.5380/atoz.v13.89352>

NOTAS DA OBRA E CONFORMIDADE COM A CIÊNCIA ABERTA

CONTRIBUIÇÃO DE AUTORIA

Papéis e contribuições	Patrícia Nascimento Silva
Concepção do manuscrito	X
Escrita do manuscrito	X
Metodologia	X
Curadoria dos dados	X
Discussão dos resultados	X
Análise dos dados	X

EQUIPE EDITORIAL

Editora/Editor Chefe

Paula Carina de Araújo (<https://orcid.org/0000-0003-4608-752X>)

Editora/Editor Associada/Associado Júnior

Karolayne Costa Rodrigues de Lima (<https://orcid.org/0000-0002-6311-8482>)

Editora/Editor de Texto Responsável

Cristiane Sinimbu Sanchez (<https://orcid.org/0000-0002-0247-3579>)

Seção de Apoio às Publicações Científicas Periódicas - Sistema de Bibliotecas (SiBi) da Universidade Federal do Paraná - UFPR

Editora/Editor de Layout

Tiago Batista Pedra (<https://orcid.org/0009-0000-7385-7273>)