

Governo Aberto

Felipe Mateos Castro de Souza - 11796909

Glaucia Pamponet Sobrinho - 11271000

Gustavo Tsuyoshi Ariga - 11857215

Henrique Tsuyoshi Yara - 11796083

Silas Bovolin Reis - 11796739

Sungwon Yoon - 9822261

19 de julho de 2023

Sumário

1	Introdução	2
2	Objetivo	3
2.1	Objetivo Específico	3
2.2	Justificativa da Escolha dos Municípios	4
3	Metodologia	6
3.1	Análise de Conteúdo	6
3.2	Comparação e Recomendações	9
4	Resultado	10
4.1	Avaliação W3C dos portais de transparência	10
5	Conclusão	16

Capítulo 1

Introdução

Diante do cenário de uma sociedade cada vez mais imersa na era digital e interativa, a gestão das escolas públicas não se esquivava dessa transformação. No meio de uma avalanche de informações e de expectativas em constante crescimento, a missão de prestar contas e democratizar os dados educacionais surge como uma jornada desafiadora, mas de vital importância, a ser percorrida. Apesar da relevância incontestável das escolas para o desenvolvimento integral dos alunos e para a consolidação das fundações socioeconômicas, estabelecer um sistema eficiente de prestação de contas continua sendo um desafio que temos pela frente.

Neste cenário, a democratização de dados educacionais surge como um protagonista crucial. A sociedade digital contemporânea, repleta de informações instantâneas e onipresentes, aguarda um acesso amplo e eficiente aos dados escolares. Esse passo, além de promover a transparência, é um impulso fundamental na direção da responsabilidade das escolas públicas. No entanto, o processo de democratização desses dados é como um caminho de duas vias, e ainda estamos navegando pelas curvas de sua compreensibilidade e acessibilidade.

O alicerce deste trabalho repousa sobre a teoria do Governo Aberto, que destaca a transparência, participação e colaboração como trampolim para uma governança eficaz. Essa teoria preconiza que, ao iluminar os cantos escuros dos dados governamentais, estamos incentivando a participação cidadã nas decisões públicas, proporcionando uma administração mais acertada. Levando esse preceito para o universo da educação, podemos projetar estratégias mais acuradas para aprimorar a prestação de contas e a democratização dos dados escolares.

Capítulo 2

Objetivo

O propósito principal deste projeto é avaliar a adesão às “Melhores Práticas para Dados na Web” da W3C [16] nas cidades de São Paulo com foco específico em conjuntos de dados relacionados à educação disponíveis em seu respectivo portal de transparência.

2.1 Objetivo Específico

Através deste estudo, pretendemos:

- **Analisar a qualidade e acessibilidade dos dados:** Avaliar se os conjuntos de dados educacionais nestas cidades estão em conformidade com as melhores práticas recomendadas pelo W3C [15], como a disponibilidade de descrições de alto nível dos conjuntos de dados, a granularidade dos dados, a existência de metadados apropriados e o uso de formatos abertos e amplamente utilizados.
- **Comparar as práticas entre diferentes cidades:** Comparar e contrastar a adesão a essas melhores práticas entre a cidade de São Paulo e outras capitais brasileiras, sendo estas: Campo Grande, Curitiba, Manaus e Salvador. O objetivo é metrificar, em termos de práticas de dados abertos, o desempenho de São Paulo através da comparação entre outros municípios.
- **Fornecer recomendações para aprimorar a abertura de dados:** Com base nas descobertas, oferecer sugestões para aprimorar a transparência e a abertura dos dados relacionados à educação em cada uma dessas cidades.

Para alcançar esses objetivos, este estudo utilizou alguns sites como recurso de apoio:

-
- Portal de Transparência da Educação do Estado de São Paulo [4]
 - QEdu [12]
 - Pátio Digital da Prefeitura Municipal de São Paulo [9]
 - Censo Escolar do INEP [3]
 - Portal da Transparência do Governo Federal – Seção Educação [14]
 - Inep consulta de Escolas [6]
 - Heuristic Evaluation. Usability Evaluation Materials. [7]
 - Document Analysis as a Qualitative Research Method [1]
 - Portal de Transparência - Curitiba [2]
 - Portal de Transparência - Salvador [11]
 - Portal da Transparência de Campo Grande [10]
 - Portal da Transparência de Manaus [8]

2.2 Justificativa da Escolha dos Municípios

A seleção dos municípios de São Paulo, Curitiba, Salvador, Manaus e Campo Grande para este estudo foi estrategicamente pensada considerando o seguinte:

São Paulo [13]: Como a maior cidade do Brasil, São Paulo possui o maior sistema educacional municipal do país. Além disso, a cidade é conhecida por seus esforços para promover a transparência e a inovação digital, tornando-a um caso interessante para análise.

Curitiba [2]: Curitiba é frequentemente referida como uma cidade inovadora no que se refere a soluções urbanas e também possui um histórico forte no que diz respeito à transparência e uso de dados abertos. Além disso, sua posição como uma das principais cidades da região Sul a torna uma adição valiosa ao estudo.

Salvador [11]: Como uma das maiores cidades do Nordeste e com uma complexa rede de ensino, Salvador oferece uma visão valiosa das práticas de dados abertos nesta região. Além disso, a cidade tem feito esforços significativos para melhorar a educação e a transparência, tornando-a uma escolha relevante para este projeto.

Manaus [8]: Manaus é a cidade mais populosa da região Norte e tem um contexto educacional único, devido a seus desafios geográficos e socioeconômicos. Analisar a abertura e transparência dos dados educacionais nesta cidade proporcionará uma perspectiva importante das práticas de dados abertos nesta região.

Campo Grande [10]: Representando a região Centro-Oeste, Campo Grande tem um sistema educacional significativo e é conhecida por seus esforços em relação à governança digital. A inclusão desta cidade fornecerá uma visão mais completa das práticas de dados abertos em todo o Brasil.

Cada cidade foi selecionada não apenas pela sua representatividade regional, mas também pelas suas características únicas em termos de tamanho, desafios educacionais e compromisso com a transparência e os dados abertos. Acreditamos que essa diversidade fortalecerá a análise e as conclusões do nosso estudo.

Capítulo 3

Metodologia

Neste projeto, empregaremos uma metodologia que combina análise de conteúdo para avaliar a adesão das cidades selecionadas às Melhores Práticas para Dados na Web da W3C [16].

3.1 Análise de Conteúdo

Na análise de conteúdo, nos concentraremos nos portais de transparência de cada cidade para coletar e analisar os conjuntos de dados relacionados à educação. Esta etapa inclui o seguinte:

- **Coleta de dados:** Visitaremos cada portal [13] [2] [11] [8] [10] e coletaremos os conjuntos de dados relevantes à educação disponíveis. Isso inclui dados sobre escolas, professores, alunos, desempenho dos alunos, financiamento e outros aspectos relevantes da educação.
- **Avaliação da qualidade dos dados:** Avaliaremos a qualidade dos dados coletados de acordo com os critérios estabelecidos pelas Melhores Práticas para Dados na Web da W3C [15]. Isso inclui a análise da completude, precisão, relevância e consistência dos dados.
- **Avaliação da acessibilidade dos dados:** Avaliaremos a facilidade de acesso aos dados nos portais, a clareza das informações fornecidas e a presença de metadados que facilitam a compreensão e o uso dos dados.

No quesito da qualidade dos dados, é necessário definir qual será a medida específica adotada para cada ponto do W3C:

1. **Fornecer metadados:** Verificado se o arquivo disponibilizado contém cabeçalho descritivo para cada coluna da planilha;
2. **Fornecer metadados descritivos:** Verificado se é disponibilizado no arquivo ou em algum link da página visitada se há uma descrição que

define especificamente o que as informações de cada coluna da planilha representa;

3. **Fornecer metadados estruturais:** Verificado se é disponibilizado no arquivo ou em algum link da página visitada a estrutura que compõem os metadados da planilha;
4. **Fornecer informações sobre a licença de dados:** Verificado se é disponibilizado um arquivo ou um link de acesso que descre a licença de dados;
5. **Fornecer informações sobre a procedência dos dados:** Verificado se é informado como foi obtido os dados informados nos arquivos avaliados;
6. **Fornecer informações de qualidade de dados:** Verificado se é informado se há dados faltantes, quantos se houver, o formato dos dados, etc;
7. **Fornecer indicador de versão:** Exibição na página de download dos arquivos, ou no próprio arquivo, a versão dos dados disponibilizados;
8. **Fornecer o histórico de versão:** Exibição de uma tabela contendo o histórico de versão. se houver indicação de versão;
9. **Usar URIs persistentes como identificadores de conjuntos de dados:** Nos dados agrupados, fornecimento de URIs que levam aos dados específicos;
10. **Usar URIs persistentes como identificadores dentro de conjunto de dados:** Nos dados específicos, exibição de URIs que levam a outros dados vinculados a este;
11. **Atribuir URIs para as versões dos conjuntos de dados e séries:** Nos indicadores de versões, atribuir URIs que levam à planilha;
12. **Usar formatos de dados padronizados legíveis por máquinas:** Os dados disponibilizados devem estar em arquivos csv, xlsx, xml ou outro formato que uma máquina consiga manipular os dados contidos;
13. **Usar representações de dados que sejam independentes de localidade:** Dados como data, moedas e palavras são entendíveis independente da localidade do usuário;
14. **Fornecer dados em formatos múltiplos:** Deve disponibilizar os dados em, ao menos, dois formatos diferentes;
15. **Reutilizar vocabulários, dando preferência aos padronizados:** Os dados que são disponibilizados, se forem repetidos, devem estar de maneira igual, não com palavras semelhantes ou formatos diferentes;
16. **Escolher o nível de formalização adequado:** Ser retirado ou colocado formato de formalização dos dados que melhor define o que está sendo representado pelas informações exibidas;

-
17. **Fornecer download em massa:** Permitir, através de diferentes de usuários, ou do mesmo, a extração dos dados sem bloqueio por quantidade de dados;
 18. **Fornecer subconjuntos para conjuntos de dados extensos:** Poder desagrupar dados que são construídos a partir de outros dados;
 19. **Usar negociação de conteúdo para disponibilizar dados em formatos múltiplos:** Em uma API, receber os tipos de formatos aceitos pelo cliente e disponibilizar os dados em um dos formatos recebidos;
 20. **Fornecer acesso em tempo real:** Ao tentar realizar o download dos dados, deve ser executado o procedimento no mesmo momento, não deve ser enviado por e-mail ou por notificação após x minutos;
 21. **Fornecer dados atualizados:** Os dados são atualizados diariamente, de segunda a sexta, desconsiderando feriados;
 22. **Fornecer uma explicação para os dados que não estão disponíveis:** Se houver dados faltantes, conter uma explicação dentro de uma aba de observações ou na página de download;
 23. **Disponibilizar dados por meio de uma API:** Os dados podem ser obtidos através de uma requisição em um endpoint de uma API;
 24. **Usar padrões Web como base para construção de APIs:** Utilizar um estilo arquitetônico baseado nas tecnologias da própria Web;
 25. **Fornecer documentação completa para as APIs:** Fornecer informações completas na Web sobre a API. Atualizar a documentação conforme características adicionadas ou modificações realizadas;
 26. **Evitar alterações que afetem o funcionamento de sua API:** Ao realizar uma requisição para a API recentemente atualizada, os dados devem ser idênticos aos obtidos antes da atualização;
 27. **Preservar identificadores:** Os identificadores dos dados disponibilizados não devem ser faltantes no arquivo;
 28. **Avaliar a cobertura do conjunto de dados:** Os dados devem estar disponíveis independente da localidade do usuário;
 29. **Coletar feedback de consumidores de dados:** Deve haver uma aba para avaliação ou comentário na página de download do arquivo;
 30. **Compartilhar o feedback disponível:** Deve haver uma aba que disponibilize as avaliações de usuários terceiros;
 31. **Enriquecer dados por meio da geração de novos dados:** Dados secundários devem ser obtidos por meio de dados primários, ambos disponibilizados no arquivo;

-
32. **Fornecer visualizações complementares:** Fornecer visualizações complementares, como gráficos, grids, mapas, heat maps, entre outras formas de exibição;
 33. **Fornecer feedback para o publicador original:** Fornecer feedback para o responsável dos dados disponibilizados;
 34. **Obedecer os termos de licença:** Se houver disponibilizado os termos de licença, não infringi-los;
 35. **Citar a publicação original do conjunto de dados:** Se for um dado republicado, citar a publicação original por meio de referências.

3.2 Comparação e Recomendações

Após a análise de conteúdo, compararemos nossas descobertas entre as diferentes cidades. Isto nos permitirá identificar as cidades que estão liderando em termos de práticas de dados abertos e aquelas que podem precisar de melhorias.

Com base em nossas descobertas, formularíamos recomendações para cada cidade, sugerindo maneiras de melhorar a transparência, a qualidade e a acessibilidade dos dados de educação.

Os resultados desta análise permitirão não apenas uma avaliação aprofundada das práticas de dados abertos nos municípios selecionados, mas também fornecerão uma base para recomendações que possam promover uma maior transparência e acessibilidade dos dados de educação no futuro.

Capítulo 4

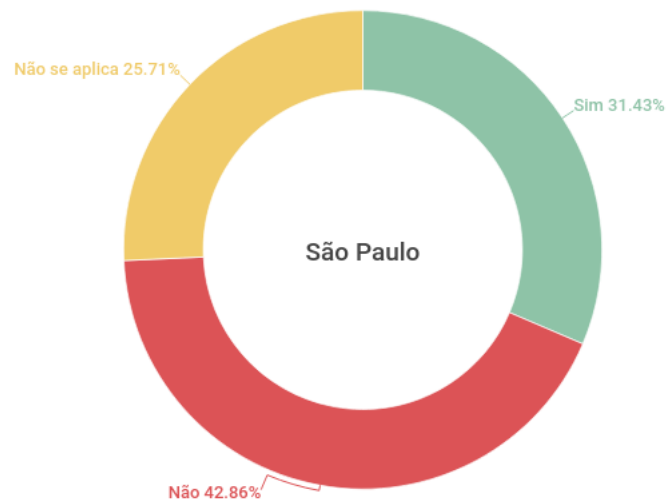
Resultado

4.1 Avaliação W3C dos portais de transparência

Para avaliar e quantificar a transparência das escolas municipais, foram utilizados os critérios de melhores práticas propostos pelo W3C [16]. O W3C pontua 35 melhores práticas [15] que podem ser seguidas para se alcançar uma boa padronização dos dados disponibilizados. A avaliação inicial se deu com os dados fornecidos pelo Programa de Transferência de Recursos Financeiros (PTRF) [13], da Secretaria de Educação do Município de São Paulo, disponibilizado através dos dados abertos do Portal de Transparência. O resultado pode ser visualizado na tabela a seguir. No gráfico subsequente à tabela, é possível avaliar os valores de forma proporcional.

Boas Práticas W3C	Atende o Requisito
Fornecer Metadados	Sim
Fornecer metadados descritivos	Não
Fornecer metadados estruturais	Não
Fornecer Informações sobre a licença de dados	Sim
Fornecer informações sobre a procedência dos dados	Não
Fornecer informações de qualidade de dados	Não
Fornecer indicador de versão	Não
Fornecer o histórico de versão	Não se aplica
Usar URIs persistentes como identificadores de conjuntos de dados	Não
Usar URIs persistentes como identificadores dentro de conjunto de dados	Não
Atribuir URIs para as versões dos conjuntos de dados e séries	Não
Usar formatos de dados padronizados legíveis por máquinas	Sim
Usar representações de dados que sejam independentes de localidade	Sim
Fornecer dados em formatos múltiplos	Não
Reutilizar vocabulários, dando preferência aos padronizados	Sim
Escolher o nível de formalização adequado	Não se aplica
Fornecer download em massa	Sim
Fornecer subconjuntos para conjuntos de dados extensos	Não
Usar negociação de conteúdo para disponibilizar dados em formatos múltiplos	Não se aplica
Fornecer acesso em tempo real	Sim
Fornecer dados atualizados	Sim
Fornecer uma explicação para os dados que não estão disponíveis	Não se aplica

Disponibilizar dados por meio de uma API	Não
Usar padrões Web como base para construção de APIs	Não se aplica
Fornecer documentação completa para as APIs	Não se aplica
Evitar alterações que afetem o funcionamento de sua API	Não se aplica
Preservar identificadores	Sim
Avaliar a cobertura do conjunto de dados	Sim
Coletar feedback de consumidores de dados	Não
Compartilhar o feedback disponível	Não se aplica
Enriquecer dados por meio da geração de novos dados	Não
Fornecer visualizações complementares	Não
Fornecer feedback para o publicador original	Não se aplica
Obedecer os termos de licença	Sim
Citar a publicação original do conjunto de dados	Não

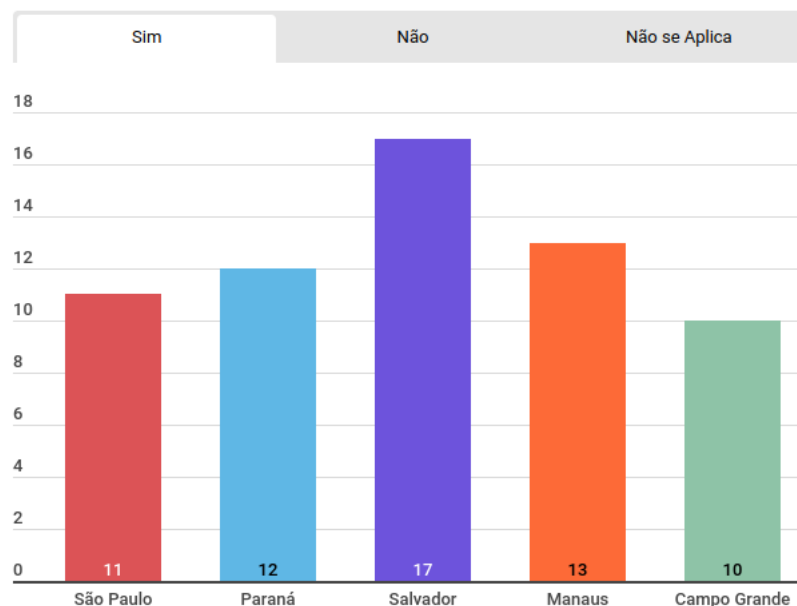


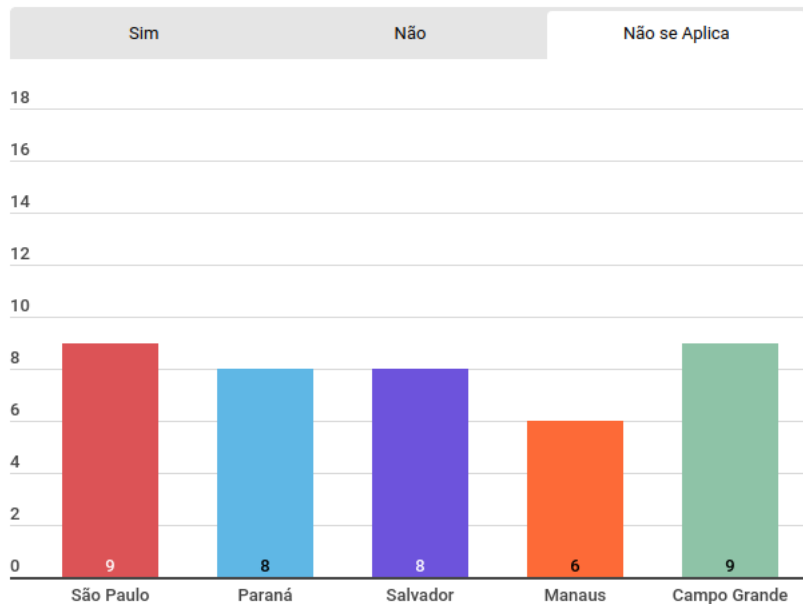
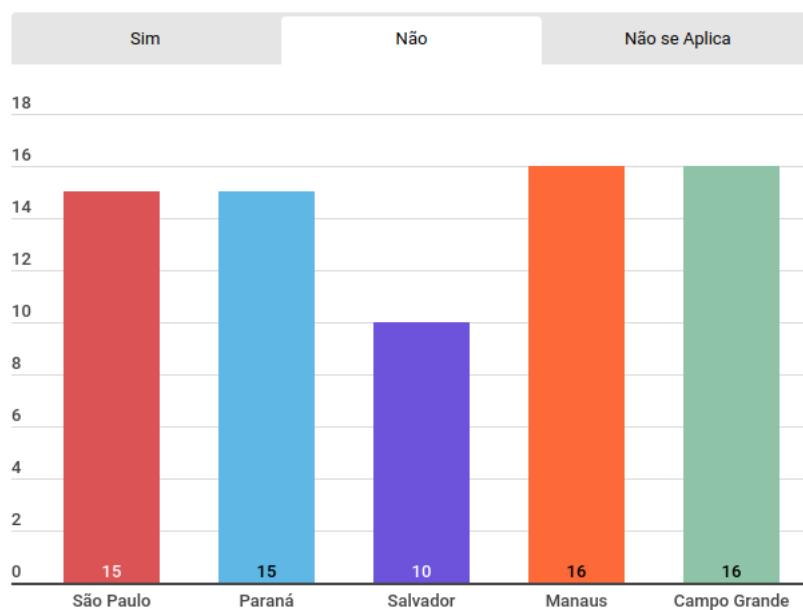
A análise dos dados do Programa de Transferência de Recursos Financeiros revelou que 31,43% das boas práticas propostas para dados abertos são seguidas pela plataforma, sendo 25,71% não aplicáveis e 42,86% não cumpridas.

No entanto, esta é uma avaliação específica de um único município do país. Com esses pontos isolados, é difícil afirmar se a proporcionalidade está dentro do esperado. Por isso, a fim de quantificar a eficácia dos dados disponibilizados, avaliamos a disponibilidade de dados relativos à educação em outros municípios. É importante salientar que o programa PTRF é exclusivo do município de São Paulo, não existindo em outras cidades. Realizamos uma análise preliminar em diversas capitais municipais brasileiras e constatamos que, além da forma de disponibilização ser muito diferente, alguns sistemas apresentam falhas de funcionalidade. Com isso em mente, selecionamos quatro capitais específicas, uma para cada região do Brasil (desconsiderando o município de São Paulo da região Sudeste): Curitiba (Sul), Campo Grande (Centro-Oeste), Salvador (Nordeste) e Manaus (Norte).

Para cada município, extraiu-se ou avaliou-se os dados exibidos relativos às escolas municipais e à educação, utilizando as métricas do W3C como base. A planilha com a avaliação completa está disponível no Github [5], sob o nome “boasPraticas.odt”.

A seguir, destacamos a pontuação obtida por cada município em cada categoria. Abaixo, apresentamos os gráficos da pontuação de cada cidade.





Os gráficos apresentados evidenciam que o município de São Paulo não demonstra um bom desempenho em transparência quando comparado a outras capitais

brasileiras, pois está em quarto lugar na pontuação de “Sim”, atrás de Paraná, Manaus e Salvador, sendo que este último atende a quase 50% dos 35 requisitos. Em relação à classificação “Não”, São Paulo tende a ficar na média das outras cidades.

Vale ressaltar algumas características observadas ao avaliar os portais de transparência. A primeira delas é que, além destes cinco municípios, os dados abertos das principais capitais sempre atendem a alguns pontos específicos do W3C: fornecimento de acesso em tempo real, fornecimento de dados atualizados e representação de dados que sejam independentes de localidade.

Acredita-se que esses pontos sempre serão atendidos (ou é esperado que sejam) devido a uma das principais leis que regulam a acessibilidade dos dados governamentais, a Lei Nº 12.527, conhecida como Lei de Transparência. Nesta lei, é especificado que os dados devem ser exibidos em tempo real (Art. 8, Inciso 4º), de forma atualizada (Art. 8, Inciso 3º) e em linguagem de fácil compreensão (Art. 5º).

Outra característica notada é que a disponibilização de dados por meio de API não era atendida em nenhum portal de transparência, inclusive nos cinco avaliados acima. Como consequência, os pontos de fornecimento de documentação completa da API, evitar alterações que afetem o funcionamento da API e usar padrões web como base para construção de API são considerados não aplicáveis.

Acreditamos que, por ser a API uma abordagem mais técnica de disponibilização de dados, essa implementação pode não ser tão conhecida para leigos, como arquivos nos formatos .CSV, .XLSX e outros. Essa pode ser uma explicação plausível para a ausência de dados disponíveis por meio de uma API.

Capítulo 5

Conclusão

Através da análise isolada da transparência de dados do PTRF, disponibilizados pela Secretaria da Educação de São Paulo, é possível concluir que os responsáveis por desenhar a implementação de acessibilidade de dados através dos portais de transparência tiveram o cuidado, ou pelo menos uma preocupação, na forma como estes dados são disponibilizados, visto que quase um terço das boas práticas adotadas pelo W3C são seguidas. No entanto, ainda há muitos pontos a serem melhorados, principalmente no quesito de disponibilização dos dados de outras formas, por exemplo através de uma API ou em outros formatos abertos.

Além disto, ao realizar a análise comparativa com outros municípios brasileiros, constata-se que a transparência adotada por São Paulo não atende tantos requisitos quanto as capitais avaliadas. Por exemplo, o portal de transparência de Salvador, possui uma aba onde é possível baixar os metadados descritivos e estruturais, em outras palavras, há disponível uma documentação que explica quais são os dados disponibilizados. Do ponto de vista prático, é algo simples de ser feito e auxilia usuários leigos que buscam informação dos recursos destinados as escolas, mas que não foi ainda implementado no município de São Paulo. Por isso, é reiterado que há pontos que devem ser melhorados para que a transparência dos dados governamentais sejam mais acessíveis e entendíveis pelos paulistanos.

Bibliografia

- [1] Glenn Bowen. “Document Analysis as a Qualitative Research Method”. Em: *Qualitative Research Journal* 9 (ago. de 2009), pp. 27–40. DOI: [10.3316/QRJ0902027](https://doi.org/10.3316/QRJ0902027). URL: <https://biotap.utk.edu/wp-content/uploads/2019/02/document-analysis.pdf>.
- [2] Prefeitura de Curitiba. *Portal de Transparência*. URL: <https://www.transparencia.curitiba.pr.gov.br/sgp/despesasorcamentarias.aspx>.
- [3] Educacenso. *Educacenso*. URL: <https://censobasico.inep.gov.br/censobasico/#/>.
- [4] ENAP. *Receitas e Despesas*. URL: <https://transparencia.educacao.sp.gov.br/Home/PrestacaoContas>.
- [5] Governo Aberto. URL: <https://github.com/swy-326/GovernoAberto>.
- [6] Inep. *Inep catálogo de Escolas*. URL: <https://inepdata.inep.gov.br/analytics/saw.dll?Dashboard>.
- [7] Darryn Lavery, Gilbert Cockton e Malcolm Atkinson. “Heuristic Evaluation. Usability Evaluation Materials.” Em: (mai. de 1996), pp. 249–256. URL: <https://concretaux.com/wp-content/uploads/nielsenheuristicsCHI.pdf>.
- [8] Manaus Prefeitura Municipal. *Portal de Transparência Municipal*. URL: <https://transparencia.manaus.am.gov.br/transparencia/v2/#/home>.
- [9] PatioDigital. *Patio Digital*. URL: <https://patiodigital.sme.prefeitura.sp.gov.br/>.
- [10] Campo Grande Prefeitura. *Consulta de Despesas*. URL: https://transparencia.campogrande.ms.gov.br/consulta-de-despesas/?consulta_post=despesas&draw=1&data-inicial=01%5C%2F01%5C%2F2023&data-final=21%5C%2F06%5C%2F2023&orgao=30&categoria=&nome-fornecedor=&fornecedor=.
- [11] Salvador Prefeitura. *Transparência Salvador*. URL: <http://www.transparencia.salvador.ba.gov.br/#/RealizacaoDespesa>.

-
- [12] QEDU. *Qedu*. URL: <https://qedu.org.br/>.
- [13] Prefeitura de São Paulo. *Programa de Transferência de Recursos Financeiros - PTRF*. URL: http://dados.prefeitura.sp.gov.br/pt_PT/dataset/programa-de-transferencia-de-recursos-financeiros-ptrf.
- [14] Portal da Transparência. *Portal da Transparência*. URL: <https://portaldatransparencia.gov.br/funcoes/12-educacao?ano=2023>.
- [15] W3C. *Best Practices Summary*. URL: <https://www.w3.org/TR/dwbp/#bp-summary>.
- [16] W3C. *Data on the Web Best Practices*. URL: <https://www.w3.org/TR/dwbp/>.