



Parecer B de *Reúso de dados abertos do governo brasileiro: atualização da métrica DGABr*

Parecerista Vinícius Medina Kern

<https://doi.org/10.1590/1808-5245.30.138332B>

Artigo avaliado: NASCIMENTO SILVA, Patrícia. Reúso de dados abertos do governo brasileiro: atualização da métrica DGABr. **Em Questão**, Porto Alegre, v. 30, e-138332, 2024. DOI: <https://doi.org/10.1590/1808-5245.30.138332>

PARECER

Completo em: 2024-03-08 12:09 AM

Recomendação: Rejeitar

1. Adequação ao perfil editorial da Revista Em Questão:*

Excelente

2. Relevância do tema:*

Excelente

3. Originalidade na abordagem do tema:*

Bom

4. Contribuição para a área da Ciência da Informação:*

Ruim

5. Considerações a respeito da relevância, originalidade e contribuição para a área do conhecimento:

O artigo precisa ser reescrito para ser adequado em relevância, originalidade e contribuição para a ciência da informação. O tópico tem grande relevância e tempestividade.

6. Qualidade e pertinência do referencial teórico:*

Bom

7. Considerações a respeito da qualidade e pertinência do referencial teórico:

A literatura usada parece adequada, mas não tenho condições de avaliar categoricamente a atualidade e abrangência do referencial, por não ser perito no tópico.

8. Pertinência e adequação dos procedimentos metodológicos:*

Bom

9. Clareza na apresentação e detalhamento dos procedimentos metodológicos:*

Ruim

10. Considerações a respeito dos procedimentos metodológicos:

A seção dos Procedimentos tem clareza e conteúdo bons, em contraste com as seções anteriores.

11. Consistência e discussão dos resultados e coerência das conclusões:*

Bom

12. Considerações a respeito dos resultados e das conclusões:

Resultados apresentados com bom detalhamento. As conclusões podem melhorar se forem baseadas na interpretação dos resultados.

13. Qualidade da argumentação (clareza, concisão, objetividade), correção textual e estrutura do texto:*

Ruim

14. Considerações a respeito da linguagem e redação do texto:*

A comunicação da pesquisa é problemática. Problemas de sintaxe, alguns dos quais levam a problemas de semântica. Problemas de estilo, resultando em baixa legibilidade. Baixa coesão. Erros fatuais. Mais detalhes no Detalhamento.

15. Qual a sua recomendação sobre o aceite do artigo?*

Não aceite.

16. Especificar as alterações sugeridas e/ou justificar a não aceitação.*

O artigo apresenta uma segunda versão da métrica DGABr (2018) para avaliar o potencial de reúso dos dados governamentais abertos no Brasil. Identifica mudanças contextuais que levam a alterações na métrica original. No entanto, a comunicação da pesquisa tem muitos problemas, pelo que não considero possível apenas fazer correções. Me parece necessário escrever outro artigo, pelo menos quanto às seções 1, 2 e 3, bem como ao Resumo.

COMENTÁRIOS PRINCIPAIS

- O artigo trata de um tema muito relevante e tempestivo, de grande interesse para o público de *Em Questão*.
- O principal problema é o uso da linguagem, que não favorece o sucesso do futuro artigo (i.e., ser lido, citado, usado). Os desajustes de sintaxe/semântica e estilo do texto resultam em falta de clareza e coesão no Resumo e 3 primeiras seções. Há algumas afirmações errôneas. A escolha do que dizer ou omitir também não favoreceu a compreensão. Há exemplos desses problemas na seção "Detalhamento", abaixo.
- Sugiro um título mais representativo do artigo, por exemplo: "Atualização da métrica DGABr: avaliando o potencial de reúso de dados governamentais abertos no Brasil".
- O Resumo é extenso (250 palavras), mas omite aspectos essenciais. A parte introdutória é excessiva (145 p., abre demais o contexto - movimento do acesso aberto, números de normas que poderiam ficar para a Introdução) e não revela a lacuna de conhecimento, a motivação para a pesquisa (O que há ou não na métrica original que motiva esta segunda versão?). As demais partes têm tamanho razoável. A parte metodológica classifica a pesquisa (dispensável, mas é tradição na área), mas não menciona o método de análise. Os resultados parecem bem escritos. A parte conclusiva poderia ser melhor: o que há nos resultados que permite essa expectativa (opinião) de que a v. 2.0 da especificação seja referência? Os leitores podem construir sua opinião se os fatos são dados.

- A Introdução é pouco efetiva em seus 9 parágrafos, pelo precário encadeamento de ideias (ver Detalhamento).
- Na seção 2, foi difícil acompanhar e entender as perspectivas, dimensões, pesos e níveis. "Peso" não me parece uma designação adequada, pois se trata de um valor ou escore escolhido entre 1 e 3. Neste trecho, me parece que nível e peso foram confundidos, aumentando a dificuldade para ler: "Cada perspectiva reúne dimensões para avaliação dos DGA e para avaliar as dimensões foram criados seis níveis, de 0 a 5, sendo o peso 0 de menor valor e 5 o de maior valor". [Mantive o que escrevi para mostrar minha confusão na leitura] Ao ler o Quadro 1, verifiquei que os "pesos" são mesmo pesos. Tudo melhoraria se o Quadro viesse para o início da seção 2, bastando um "conforme o Quadro I" para o leitor ler e entender, em vez de, nesta versão do artigo, tentar acompanhar o texto (sem sucesso, no meu caso) e só no final ver o esquema.
- A seção metodológica cumpre seu papel e o Quadro 2 é muito útil - seria melhor aparecer no começo da seção. O nome comercial da ferramenta Rayyan e o fato de ser "utilizada mundialmente para revisões bibliográficas" não diz o que ela é e faz. As marcas passam (Altavista era um motor de busca dominante na web que precedeu o Google...). Percebo, no penúltimo parágrafo da seção, que o método para atingir o objetivo 2 parece ter sido o levantamento.
- A seção de Resultados me pareceu bem mais clara (parece que começa o texto de outro/a redator). Se o artigo todo tivesse este nível de qualidade de texto, haveria pouco a mudar. Há problemas pontuais, como mencionar as especificações Open Data Charter e Data on the Web Best Practices sem apontar fonte (que é dada em parágrafo mais à frente).
- As seções seguintes aos Resultados também são de leitura possível, com problemas pontuais. No final das Considerações Finais, a conclusão deve derivar dos resultados da pesquisa. O que significam os resultados que cumprem os objetivos 1 e 2? Responder a isso vai permitir alterar sutilmente (porém essencialmente) o último parágrafo.
- Em suma, o tópico do artigo é meritório, a métrica é importante para a situação dos DGA no Brasil, mas o texto do artigo é inadequado, por isso recomendei a rejeição.

DETALHAMENTO

- Estas anotações devem ser vistas como apontamentos sobre os tipos de problemas encontrados no texto e não como "correção" (copy-editing).

- O Resumo apresenta vários tipos de problema encontrados no artigo (escolhas do que detalhar ou omitir, sintaxe, semântica, estilo): Começa falando do movimento de acesso aberto (AA), contexto geral demais para o Resumo deste artigo, em vez de ir direto ao contexto geral dos dados abertos governamentais e ao contexto específico da métrica e sua necessidade para avaliar instâncias desses dados. Afirma que o AA iniciou "com dados de pesquisa" (foi pela serials crisis, artigos e não dados). A primeira sentença tem 23 palavras (2 linhas) entre o sujeito e seu verbo - problema clássico de texto complicado (não necessariamente complexo) de entender. Em "Inclinada em suprir", a regência não funciona, além de ser uma sentença iniciada por predicado (a inversão sintática raramente funciona bem). Além disso, o texto usa palavras mais complicadas que o necessário e dificulta ainda mais a leitura pela falta da crase (concernente à): "Inclinada em suprir uma lacuna concernente a avaliação dos dados abertos publicados". Se o texto fosse claro, não seria preciso dizer que DGABr foi construída "como resultado de uma pesquisa acadêmica" (e deixa supor, mas não esclarece se são os mesmos autores do artigo/métrica anterior). A abertura dos dados via LAI não veio "apesar de previsto na Constituição de 1988", mas por causa desta, para regulamentar o texto constitucional. O Resumo omite informações essenciais: Quais são "as cinco perspectivas da métrica"? Que métrica a DGABr é, o que/como se calcula? Cada perspectiva é dividida em dimensões? Aparentemente, são 5 perspectivas e 28 dimensões (mas não é possível saber lendo o Resumo). Talvez algumas dessas dúvidas (mas não todas) possam ser respondidas a partir da Introdução, para manter o Resumo conciso, com o mínimo necessário para resumir o artigo (comunicar claramente e em poucas palavras o contexto e a definição do problema, a abordagem metodológica, os resultados e as conclusões).

- Se a métrica foi criada em 2018, seria bom saber se foi aceita, usada ... mas o Resumo fala sobre aceitação bem no fim ("...continue..."; implica que vem sendo, mas não diz).
- Se "esta pesquisa objetivou revisar a DGABr", qual teria sido a lacuna que levou a empreender uma nova versão? Tinha defeitos, surgiram mudanças contextuais etc.? Fica-se sem saber.
- No trecho "à luz de seus fundamentos, permitindo a atualização", de qual atualização se trata? Se é da métrica, é sua atualização.
- A informação na Introdução é relevante, porém, organizada de uma forma que não se pode entender a partir da leitura, restando ao leitor fazer as conexões cognitivas ou esperar encontrá-las nas seções seguintes. Diferentemente do Resumo, que começa por um contexto genérico demais, a Introdução vai direto ao contexto específico (a métrica), sem menção ao contexto geral (dados abertos governamentais). Daí pula para o movimento do acesso aberto (a meu ver desnecessário), volta para expandir (em relação ao Resumo) o contexto geral, "enxerta" um trecho sobre Ciência da Informação com pouca conexão com o restante da Introdução (baixa coesão), volta novamente ao contexto geral, passa ao específico, apresenta, agora sim, a motivação para empreender uma revisão da métrica (mas em uma única sentença e sem citação/referência às mudanças contextuais) e o objetivo da pesquisa (precedido por uma pergunta de pesquisa cuja resposta não leva a atingir o objetivo).
- Mau uso de citação: Segundo o 1º parágrafo da Introdução, a DGABr foi publicada em 2018 conforme Nikifova (2023). Não parece ser isso, mas é o que o texto diz ao fazer afirmações cujo único embasamento possível é essa única citação no fim do parágrafo.
- Na Introdução, falar em "novo "petróleo" da Era de Dados" não informa quem já entende a figura de linguagem e deixa a ver navios quem não entende que história é essa de "petróleo".
- Se DGABr foi "um dos primeiros instrumentos para avaliar o potencial de reúso dos dados governamentais abertos disponibilizados no Brasil", deve

haver menção a outros instrumentos pioneiros. Sem isso, a afirmação carece de evidência.

- Há problemas de coesão do texto, como no início da p. 4: o volume de DGA aumentou muito, com vários formatos e padrões... A métrica DGABr objetiva avaliar o potencial de reúso dos DGA publicados pelo governo brasileiro. O que uma coisa tem a ver com a outra?

- A motivação da pesquisa finalmente é apresentada na p. 4 (tem que estar no Resumo também), embora também de forma problemática (ao mencionar "novos princípios como o FAIR", precisa dar uma fonte).

- Logo a seguir, é apresentada uma pergunta de pesquisa-que não combina com o objetivo que vem na sequência.

- Problemas de sintaxe/semântica (um exemplo): "compartilhamento pelas mesmas licenças que foi apresentada" --> "compartilhamento pelas mesmas licenças EM que foi apresentada", ou "compartilhamento pelas mesmas licenças que foRAM apresentadaS", ou ...?

- Ainda sobre sintaxe, há vários casos de falta de crase (embora eu considere um problema muito menor do que crase onde não deve existir).

- "A métrica DGABr tem como objetivo avaliar ..." O fato de esta informação contextual essencial (e a lacuna de conhecimento exposta a seguir) estar lá na p. 4 (3a da Intro) mostra que o texto introdutório anterior é excessivo. Introduzir é apresentar o contexto geral, específico, a lacuna de conhecimento e a definição do problema de pesquisa/desenvolvimento, completando com uma transição para o restante do artigo. Muitos defendem a "Introdução de 5 parágrafos", o que não é uma "lei" da redação científica, mas sugere o quanto os 9 longos parágrafos poderiam ser mais sucintos, principalmente pela eliminação de informações não-essenciais para a Introdução.

- Eliminar texto desnecessário melhora a legibilidade: "o motivo da revisão é devido aos oito" --> "a revisão é devida aos oito"

- Mistura todo X parte: "as legislações foram atualizadas" --> "a legislação foi atualizada" ou "as leis foram atualizadas" ou...

- Falta citação: "O ePING é o padrão de interoperabilidade do Governo Eletrônico, foi criado em 2014 e atualizado em 2018".

- Sobre a mania/tradição na área de classificar (mal) a pesquisa, sem preparação para isso. Os manuais metodológicos omitem que só discutem a pesquisa estritamente científica e omitem a tecnológica (grande parte da pesquisa na área), o que leva os novos cientistas a classificar mal sua pesquisa tecnológica (é o caso aqui; o principal resultado é um artefato) embora use métodos científicos para coletar e analisar dados e atingir os primeiros objetivos. A confusão sobre a tipologia da pesquisa pode afetar a composição do objetivo e das conclusões...
- Há espaçamento excedente (" ") por digitação e não por justificação do texto).
- Há 4 referências distintas a Silva, P. N. Todas usam grafias diferentes.
- As referências não estão em ordem alfabética, bem como há URLs com link vivo e outras como texto comum (heterogeneidade)

Recomendação: REJEITAR

