

PANORAMA DAS PESQUISAS SOBRE DADOS GOVERNAMENTAIS ABERTOS NO PERÍODO 2014 A 2024: UMA ANÁLISE DE CONTEÚDO

**Renato Machado de Godoy
João Paulo Calembo Batista Menezes
João Cesar de Souza Ferreira
Marcio Coutinho de Souza**

Universidade Federal dos Vales do Jequitinhonha e Mucuri (UFVJM), Teófilo Otoni – MG, Brasil

Os dados governamentais abertos (DGA) têm se estabelecido como elemento central na modernização da gestão pública. Este estudo investiga como estão caracterizadas as publicações científicas sobre DGA nas bases *Web of Science* e *Scopus*. A metodologia empregou análise de conteúdo em 332 artigos no período de 2014 a 2024, utilizando o software IRaMuTeQ. Os resultados revelaram uma estrutura temática dividida em sete classes principais, a saber: métricas de qualidade e aplicações na saúde (Classe 6), modelos teóricos de avaliação (Classe 7), desafios de interoperabilidade tecnológica (Classe 2), forte foco na governança local e engajamento cívico (Classe 3), diferenças entre estratégias governamentais regionais (Classe 5), dificuldades de implementação (Classe 4) e aplicações práticas em contextos urbanos (Classe 1). A análise de similitude identificou ‘*data*’ e ‘*open_government_data*’ como termos centrais, nos quais orbitam discussões sobre governança, transparência e aplicações práticas. O principal achado da pesquisa é a existência de uma estrutura que conecta harmoniosamente os aspectos técnicos, sociais e práticos dos DGA, demonstrando que o sucesso das iniciativas depende dessa integração. O estudo contribui para a compreensão do panorama da pesquisa em DGA, oferecendo perspectivas para o desenvolvimento de políticas públicas mais efetivas.

Palavras-chave: dados governamentais abertos; análise de conteúdo; transparência; governança pública; IRaMuTeQ.

PANORAMA DE LAS INVESTIGACIONES SOBRE DATOS GUBERNAMENTALES ABIERTOS EN EL PERÍODO 2014 A 2024: UN ANÁLISIS DE CONTENIDO

Los datos gubernamentales abiertos (DGA) se han establecido como un elemento central en la modernización de la gestión pública. Este estudio investiga cómo están caracterizadas las publicaciones científicas sobre DGA en las bases Web of Science y Scopus. La metodología empleó análisis de contenido en 332 artículos en el período de 2014 a 2024, utilizando el software IRaMuTeQ. Los resultados revelaron una estructura temática dividida en siete clases principales, a saber: métricas de calidad y aplicaciones en la salud (Clase 6), modelos teóricos de evaluación (Clase 7), desafíos de interoperabilidad tecnológica (Clase 2), fuerte foco en la gobernanza local y el compromiso cívico (Clase 3), diferencias entre estrategias gubernamentales regionales (Clase 5), dificultades de implementación (Clase 4) y aplicaciones prácticas en contextos urbanos (Clase 1). El análisis de similitud identificó 'data' y 'open_government_data' como términos centrales, en los cuales orbitan discusiones sobre gobernanza, transparencia y aplicaciones prácticas. El principal hallazgo de la investigación es la existencia de una estructura que conecta armoniosamente los aspectos técnicos, sociales y prácticos de los DGA, demostrando que el éxito de las iniciativas depende de esa integración. El estudio contribuye a la comprensión del panorama de la investigación en DGA, ofreciendo perspectivas para el desarrollo de políticas públicas más efectivas.

Palabras clave: datos gubernamentales abiertos; análisis de contenido; transparencia; gobernanza pública; IRaMuTeQ.

OVERVIEW OF RESEARCH ON OPEN GOVERNMENT DATA FROM 2014 TO 2024: A CONTENT ANALYSIS

Open government data (OGD) has become a central element in the modernization of public administration. This study investigates the characteristics of scientific publications on OGD in the Web of Science and Scopus databases. The methodology involved a content analysis of 332 articles from 2014 to 2024, using the IRaMuTeQ software. The results revealed a thematic structure divided into seven main classes: quality metrics and healthcare applications (Class 6), theoretical evaluation models (Class 7), technological interoperability challenges (Class 2), a strong focus on local governance and civic engagement (Class 3), differences between regional government strategies (Class 5), implementation difficulties (Class 4), and practical applications in urban contexts (Class 1). A similitude analysis identified 'data' and 'open_government_data' as central terms, around which discussions on governance, transparency, and practical applications revolve. The research's main finding is the existence of a framework that harmoniously connects the technical, social, and practical aspects of OGD, demonstrating that the success of these initiatives depends on this integration. The study contributes to the understanding of the OGD research landscape, offering perspectives for the development of more effective public policies.

Keywords: open government data; content analysis; transparency; public governance; IRaMuTeQ.

1. INTRODUÇÃO

Em 1948, o acesso à informação foi reconhecido pela Assembleia Geral das Nações Unidas como um direito humano universal. Assim, todas as pessoas têm o direito fundamental de coletar, transmitir e publicar informações (United Nations, 1948).

Desde 2007, diversos movimentos de dados abertos foram fortalecidos ao redor do mundo, ganhando destaque com iniciativas como a reunião em Sebastopol, Califórnia, que definiu princípios dos dados abertos; e os memorandos presidenciais de 2009 nos Estados Unidos da América (EUA). Esses movimentos têm como principais objetivos a transparência e a reutilização de dados. Nesse contexto, visando possibilitar a transparência e condução aberta de suas ações, os governos viabilizaram o acesso às suas informações.

A crescente demanda por transparência e acessibilidade dos dados na administração pública, especialmente no contexto da aplicação de tecnologia da informação para gestão eficiente, exige que esses dados sejam dispostos com o máximo de transparência, assegurando a acessibilidade e operacionalidade dos sistemas de dados abertos.

Attard *et al.* (2015) destacam que o valor dos dados abertos nos governos compreende a importância de garantir que essas informações sejam acessíveis a todos, superando barreiras tecnológicas, legais e de acessibilidade. Trata-se de uma questão de promover transparência e empoderar os cidadãos.

Os dados governamentais abertos (DGA) são informações geradas ou coletadas por instituições públicas, disponibilizadas de forma acessível e reutilizável, com liberdade para uso e redistribuição, promovendo transparência, inclusão e inovação (Vetrò *et al.*, 2016). Podem ser livremente acessadas, modificadas e compartilhadas por qualquer cidadão ou organização, sem restrições, e usadas para diversas finalidades, como a produção de conhecimento, desenvolvimento de estudos, elaboração de reportagens, criação de aplicativos de serviços e softwares cívicos, entre outras aplicações que possam beneficiar a sociedade.

Além disso, existem muitos benefícios decorrentes da divulgação de dados pelos governos, como maior transparência, responsabilidade, ações anticorrupção, confiança, aumento da participação e melhoria e geração de novas atividades econômicas (Kleiman *et al.*, 2024).

Diante desse cenário, este artigo busca responder à seguinte questão de pesquisa: Como estão caracterizadas as publicações científicas sobre dados governamentais abertos (DGA) nas bases *Web of Science* e *Scopus*? Para tanto, o objetivo geral foi mapear tendências, categorias temáticas e principais contribuições dessas publicações no período de 2014 a 2024.

Esta análise identificou quais foram as tendências de publicação, bem como explorou qualitativamente os temas abordados nos artigos mais citados, utilizando ferramentas como o software *Interface de R pour les Analyses Multidimensionnelles de Textes et de Questionnaires*

(IRaMuTeq) para gerar estatísticas textuais; a Classificação Hierárquica Descendente (CHD); Análise Fatorial de Correspondências (AFC); e Análise de Similitude, com o propósito de demonstrar as principais contribuições e limitações do estudo, e, se possível, apresentar sugestões para pesquisas futuras.

2. REFERENCIAL TEÓRICO

2.1 Dados governamentais abertos

A definição de ‘aberto’ estabelece critérios (veja Figura 1) que os dados precisam cumprir. Esses critérios mostram como os dados podem ser usados, reutilizados e redistribuídos de forma gratuita (Attard *et al.*, 2015; Possamai; Souza, 2020).

Figura 1 – Princípios dos dados abertos



Fonte: elaboração própria com suporte de inteligência artificial, com base em dados de Attard *et al.*, (2015) e Possamai e Souza (2020).

A abertura de dados governamentais visa promover transparência, participação e inovação, trazendo benefícios, como maior responsabilidade, combate à corrupção, confiança pública e estímulo à economia (Kleiman *et al.*, 2024).

Gray (2014) aponta várias evidências que contribuíram para o surgimento do paradigma de abertura de dados durante os anos 1990 e início dos anos 2000, particularmente na Europa e nos EUA. Algumas das principais contribuições incluem:

- a) **Debates sobre informação do setor público:** nos anos 1990 e início dos anos 2000, políticas e debates em torno da disponibilização de informação do setor público foram fundamentais. As discussões sobre direitos autorais governamentais e modelos de licenciamento, como o caso da falta de direitos autorais sobre informações federais nos EUA, desempenharam um papel importante na abertura de dados públicos;
- b) **Diretiva de Informação do Setor Público (PSI) na Europa:** a Europa implementou a PSI, que procurava maximizar a reutilização de dados públicos e minimizar as restrições sobre eles. Embora inicialmente focada em esclarecer as condições de reutilização, a diretiva estabeleceu precedentes para iniciativas posteriores que promoviam uma maior abertura de dados;
- c) **EUA como pioneiro em dados abertos:** o governo dos EUA, particularmente por meio de iniciativas como a do *Office of Management and Budget*, começou a promover a reutilização irrestrita de informações públicas, apoiando o princípio de que o governo deveria facilitar o acesso a esses dados, sem cobrar além do custo de disseminação;
- d) **Aumento do valor econômico dos dados geoespaciais:** tanto nos EUA quanto na Europa, houve um aumento no reconhecimento do valor econômico dos dados geoespaciais. Dessa forma, políticas em torno da infraestrutura de dados espaciais foram vistas como fundamentais para permitir inovações comerciais, incentivando a abertura de dados nessa área.

O movimento pela abertura de dados governamentais deu um salto importante em 2007, com uma reunião em Sebastopol, Califórnia, onde especialistas trabalharam juntos para definir o conceito de dados abertos, auxiliando na formulação de políticas públicas voltadas à transparência (Gray, 2014).

A partir disso, a abertura de dados passou a ser uma prioridade em agendas de organizações internacionais, como o Grupo dos Oito (G8), que reúne as maiores economias do mundo; o Grupo dos Vinte (G20), que inclui as principais economias avançadas e emergentes; a Organização para a Cooperação e Desenvolvimento Econômico (OCDE); e o Banco Mundial, instituição financeira internacional que oferece assistência financeira e técnica a países em desenvolvimento (Gray, 2014).

Em 2009, o presidente dos EUA, Barack Obama, assinou três memorandos presidenciais. Dois deles tratavam do governo aberto, do qual os dados abertos são um dos pilares. Esses memorandos presidenciais colocaram explicitamente a cultura de código aberto no centro da ação pública, reafirmando seus princípios fundamentais: transparência, participação e colaboração (Chignard, 2013).

Em 2011, foi lançado o *Open Government Partnership* (OGP, 2011), com oito países fundadores, entre eles o Brasil, que visava aumentar a participação cívica e o acesso à informação governamental (Brasil, 2020).

O compromisso com dados abertos foi reafirmado em 2013 na cúpula do G8, com a *Carta Aberta de Dados* (Open Data Charter, 2013), que definiu padrões para a abertura de dados respeitando leis de propriedade intelectual e privacidade.

No Brasil, as iniciativas começaram em 2005 e ganharam força em 2009 com a lei complementar nº 131 (Brasil, 2009), que exigia a publicação de informações detalhadas sobre receitas e despesas públicas.

A Lei de Acesso à Informação (Brasil, 2011) foi um marco que regulamentou o acesso a dados abertos (Albano; Craveiro; Carossi, 2023). Em 2016, o governo implementou a Política de Dados Abertos, pondo em prática a Infraestrutura Nacional de Dados Abertos (INDA) e o Plano de Dados Abertos (PDA) (Brasil, 2016).

2.2 Governança e transparência governamental

Governança pública refere-se à implementação de práticas estratégicas de liderança e controle que capacitam os gestores e *stakeholders* de uma organização pública a avaliar as necessidades e condições atuais, direcionar suas ações e monitorar suas operações. Esse processo visa ampliar as probabilidades de sucesso na entrega de resultados efetivos aos cidadãos, tanto em termos de serviços quanto de políticas públicas (Brasil, 2020).

Governança pública abrange diversos aspectos da gestão, incluindo transparência, prestação de contas, ética, integridade, legalidade e participação social nas decisões (Pascoal; Oliveira, 2019).

Attard *et al.* (2015) e Macedo e Lemos (2021) apontam que a implementação de uma governança de dados abertos pode gerar diversos benefícios, como a eficiência operacional, a transparência, a participação cidadã e a qualidade dos serviços públicos.

No Brasil, a legislação que rege a governança de dados abertos foi significativamente fortalecida pela Lei de Acesso à Informação (Brasil, 2011) e pelo Decreto nº 8.777/2016, que instituiu a Política de Dados Abertos do Poder Executivo federal. Esse decreto estabelece diretrizes claras para a publicação de dados governamentais como dados abertos, promovendo uma administração pública mais transparente e participativa (Brasil, 2016).

O Decreto nº 9.903/2019 instituiu a Política de Dados Abertos do Poder Executivo Federal, enfatizando a gestão coordenada pela Controladoria-Geral da União e a obrigatoriedade de disponibilização dos dados em formatos que permitam o uso eficiente pela sociedade e outros órgãos governamentais (Brasil, 2019).

Um fator fundamental da governança de dados governamentais é a transparência, o que envolve possibilitar que a sociedade acesse informações atualizadas acerca das operações, estruturas, processos decisórios, resultados e desempenho do setor público. Ela se manifesta na disponibilização de informações que interessem às partes envolvidas, inclusive sob a forma de dados abertos – conforme estabelecem os artigos 3º, incisos I e II, 5º, 8º e 10 da Lei nº 12.527/2011 –, e não se limita apenas às informações exigidas por leis ou regulamentos específicos (Brasil, 2020).

Além disso, a transparência e a acessibilidade dos dados são imperativos, democráticos e legais, e contribuem para uma governança aberta e participativa, permitindo às partes interessadas um melhor acesso às informações e utilização de maneira produtiva (Reinsberg; Swedlund, 2023).

De acordo com Gurstein (2011), a transparência pode ser compreendida em diferentes níveis, incluindo a disponibilização ativa de dados, a publicação de informações relevantes e a implementação de mecanismos que facilitem o acesso e a compreensão dos dados públicos. Gurstein (2011) ainda argumenta que o uso dos dados abertos pode empoderar cidadãos ao proporcionar acesso direto a informações antes inacessíveis.

3. METODOLOGIA

Para atender ao objetivo deste estudo, realizou-se uma de análise de conteúdo. As buscas foram executadas no período de 29 de setembro de 2024 a 09 de novembro de 2024 nas bases de dados *Scopus* e *Web of Science*. Essas bases são reconhecidas pela qualidade e abrangência de seu acervo científico (Costa; Silva; Assunção, 2023; Martín-Martín *et al.*, 2019).

3.1 Procedimentos de busca

Para garantir a consistência e reproduzibilidade da pesquisa, estabeleceu-se um protocolo de busca sistemático em ambas as bases. O termo '*open government data*' foi selecionado como descritor principal, sendo aplicado de forma abrangente para capturar publicações que mencionassem esse conceito em seus elementos fundamentais (título, resumo e palavras-chave). A revisão de literatura, como sugerido por Snyder (2019) e Webster e Watson (2002), permite mapear e sintetizar o conhecimento acumulado e identificar lacunas a serem exploradas.

Nas bases *Scopus* e *Web of Science*, a estratégia de busca foi configurada para identificar publicações que contivessem o termo exato '*open government data*' em qualquer um dos três campos mencionados (título, resumo ou palavras-chave). A busca foi delimitada ao período de 2014 a 2024, focando exclusivamente em artigos científicos disponíveis em acesso aberto. Essa estratégia resultou em 289 documentos na base *Scopus* e 223 na base *Web of Science*, que atenderam a todos os critérios estabelecidos. Excluíram-se documentos que não fossem artigos (como revisões, capítulos de livros, anais de congressos), publicações anteriores a 2014, artigos não disponíveis em acesso aberto e duplicatas entre as bases.

A escolha por uma estratégia de busca abrangente, sem restrições por área específica do conhecimento, fundamenta-se em quatro aspectos principais:

Primeiro, essa abordagem permite capturar uma diversidade de temas e perspectivas que poderiam ser omitidos com um escopo mais restrito, enriquecendo a análise exploratória. Segundo, evita-se o viés de categoria, que poderia surgir ao limitar as áreas de conhecimento, especialmente considerando as intersecções entre Administração Pública e outros campos das Ciências Sociais. Terceiro, essa amplitude metodológica permite identificar a interdisciplinaridade característica

da área de Administração Pública, revelando importantes intersecções com outras disciplinas. Por fim, a não restrição do escopo maximiza a quantidade de dados disponíveis para análise, proporcionando uma visão mais completa da literatura sobre DGA.

A escolha por artigos em acesso aberto visa garantir a acessibilidade integral dos materiais para análise detalhada e permite que futuros pesquisadores possam replicar o estudo com acesso completo às mesmas fontes (Tennant *et al.*, 2016). O recorte temporal de dez anos (2014-2024) foi estabelecido para capturar a evolução recente do campo de DGA, período que coincide com importantes desenvolvimentos na área.

3.2 Eliminação de artigos duplicados

Para assegurar a consistência das referências bibliográficas e evitar duplicidades, foi desenvolvido um *script* Python denominado ‘mesclar_bib.py’, disponível publicamente em https://github.com/engodoy/dga_analise_conteudo, e ilustrado na Figura 2.

Figura 2 – Processo de eliminação de artigos duplicados



Fonte: elaboração própria com suporte de inteligência artificial.

Esse *script* foi especialmente projetado para realizar a unificação de arquivos de referências no formato BibTeX provenientes das bases *Scopus* e *Web of Science*, implementando um processo de identificação e remoção de duplicatas. O *script* opera por meio de um algoritmo de duas etapas para identificação de duplicatas. Primeiramente, utiliza o *Digital Object Identifier* (DOI) como identificador principal, após normalizar os DOIs para garantir consistência na comparação (removendo espaços e padronizando maiúsculas/minúsculas). Para entradas sem DOI, o *script* emprega uma estratégia secundária baseada na combinação do título e autores da publicação, também normalizados para maximizar a precisão da identificação.

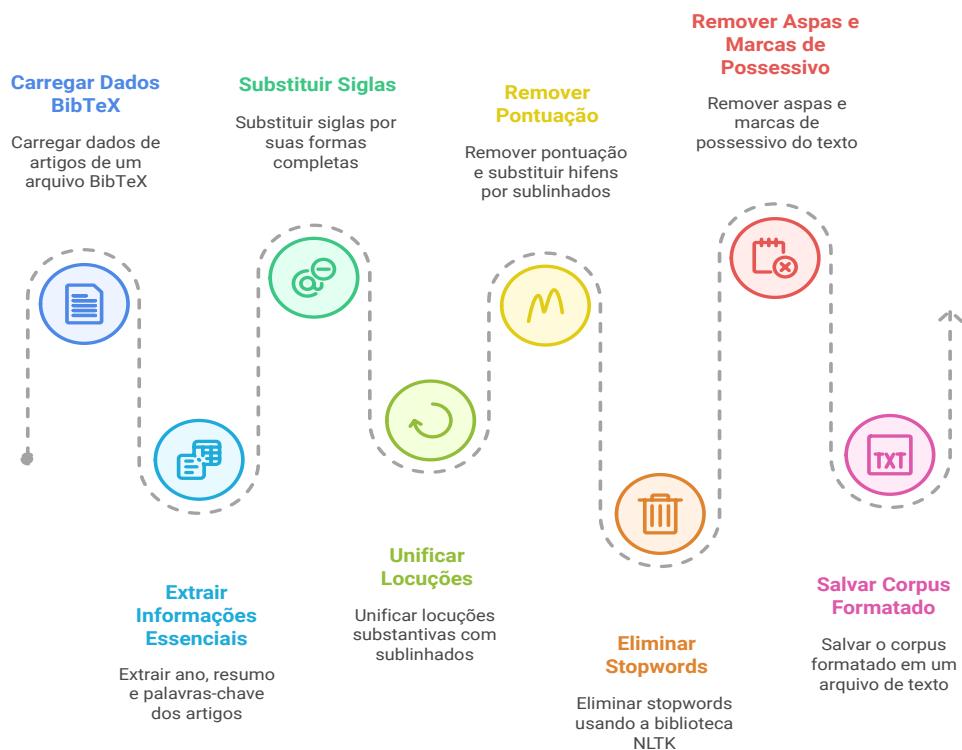
Em casos de duplicatas – quando o mesmo artigo é encontrado em ambas as bases –, o *script* mantém preferencialmente a entrada proveniente da primeira base de dados (definida como prioritária no momento da execução), garantindo consistência na seleção. O processo é totalmente automatizado e gera um novo arquivo BibTeX unificado, preservando a formatação original das entradas e incluindo estatísticas sobre o processo de unificação, como o número total de referências únicas e a quantidade de entradas identificadas por DOI versus título/autores.

Após a execução do *script*, do total inicial de 512 referências (289 da *Scopus* e 223 da *Web of Science*), foram identificadas 332 referências únicas, todas elas contendo DOI válido. A identificação de duplicatas exclusivamente via DOI, sem necessidade de recorrer à comparação por título e autores, indica alta qualidade e consistência dos metadados nas publicações selecionadas. Essa redução de 512 para 332 referências (aproximadamente 35%) demonstra uma sobreposição significativa entre as bases, justificando a importância do processo de unificação para evitar duplicidades na análise.

A execução desse processo de unificação foi fundamental para garantir a integridade do *corpus* de análise, permitindo identificar e eliminar sobreposições entre as bases de forma sistemática e reproduzível.

3.3 Preparação do *corpus* textual

Para realizar uma análise textual estruturada e garantir a consistência semântica do *corpus*, foi necessário preparar os resumos dos artigos selecionados de forma adequada para o *software* IRaMuTeQ, que auxilia na classificação e interpretação de grandes volumes de dados textuais. O objetivo desse processamento foi assegurar que o *corpus* estivesse padronizado e isento de elementos que pudesse distorcer a análise lexicométrica, como variações de siglas, pontuações e palavras irrelevantes. Esse processo pode ser visualizado na Figura 3.

Figura 3 – Preparação do *corpus* textual

Fonte: elaboração própria com suporte de inteligência artificial.

Utilizando como referências Camargo e Justo (2021) e Salviati (2017), foi desenvolvido um *script* Python¹, com o nome de ‘extrair_resumos_iramuteq.py’, para adequar os textos para serem processados pelo IraMuTeQ, conforme descrito abaixo.

A primeira etapa do *script* envolve o carregamento dos dados a partir do arquivo BibTeX unificado. Usando a biblioteca ‘bibtexparser’, o script acessa e extrai informações essenciais, como ano de publicação, resumo e palavras-chave dos artigos.

Para garantir consistência, o *script* realiza a substituição de siglas e abreviações comuns (como “AI” para “*artificial_intelligence*” e “OD” para “*open_data*”) por suas versões completas. Além disso, locuções substantivas recorrentes (por exemplo, “*public sector*” e “*data privacy*”) são convertidas para uma única expressão unificada com sublinhado (ex.: “*public_sector*”).

O *script* remove pontuações e substitui hifens por sublinhados nas palavras compostas. Essa etapa é necessária para eliminar caracteres que podem ser interpretados erroneamente pelo IRaMuTeQ, além de garantir que termos compostos sejam tratados como uma unidade lexical.

As chamadas *stopwords* (palavras de uso frequente, mas sem carga semântica relevante, como artigos, preposições e conjunções) são removidas utilizando a biblioteca ‘*nltk*’, com o pacote

¹ Acessível em https://github.com/rengodoy/dga_analise_conteudo

de palavras em inglês. O *script* exclui essas palavras, deixando apenas os termos que carregam significado relevante para o contexto.

O *script* também remove aspas e marcas de possessivo do texto. Esse tratamento é feito para evitar interferências nos agrupamentos lexicais e nas contagens de frequência, uma vez que esses elementos gramaticais não são úteis para a análise temática.

Após o processamento, o *script* salva o *corpus* formatado em um arquivo de texto compatível com o IRaMuTeQ. Cada resumo é precedido por um cabeçalho que indica o ano de publicação, facilitando a organização e a segmentação do texto no software.

3.4 Análise de conteúdo

Segundo Bardin (2011), a análise de conteúdo é um conjunto de ferramentas metodológicas que estão constantemente sendo refinadas e aprimoradas, são aplicadas a uma variedade de “discursos” (abrangendo tanto conteúdo quanto contexto). O elemento unificador dessas diversas técnicas – que vão desde o cálculo de frequências, fornecendo dados quantitativos, até a identificação de estruturas que podem ser traduzidas em modelos – é uma interpretação rigorosamente controlada, baseada em dedução: a inferência.

Para estabelecer categorias de análise *a priori*, selecionamos os dez artigos mais citados entre os documentos recuperados nas bases de dados Scopus e Web of Science. Essa seleção permite identificar temas centrais e recorrentes na literatura sobre *open government data*, o que auxilia na definição das categorias iniciais para a análise de conteúdo. Assim, a partir de uma leitura dos artigos, foram definidas as sete categorias de análise *a priori* (Quadro 1), com o objetivo de verificar a presença ou ausência de cada categoria no conteúdo analisado.

Quadro 1 – Categorias de análise *a priori* e definições conceituais

Categoria <i>a priori</i>	Definição conceitual
Impacto Econômico dos DGA	Refere-se aos efeitos econômicos gerados pela disponibilização de DGA, incluindo o estímulo à criação de negócios, a eficiência em setores específicos e a geração de empregos. A Comissão Europeia, por exemplo, projeta que a Estratégia de Dados Abertos traga cerca de US\$ 53 bilhões anuais à economia da União Europeia (Jetzek; Avital; Bjorn-Andersen, 2014). Ahmadi Zeleti <i>et al.</i> (Ahmadi Zeleti; Ojo; Curry, 2016) também discutem modelos de negócios baseados em DGA.
Impacto Social dos DGA	Examina como os DGA podem ajudar na solução de problemas sociais e na melhoria da qualidade de vida, incluindo avanços na saúde pública, educação e sustentabilidade ambiental. Jetzek <i>et al.</i> (2014) mencionam o uso de DGA para promover eficiência energética pela empresa Opower.
Transparência e Accountability	Envolve o uso de DGA para fortalecer a transparência governamental e promover mecanismos de <i>accountability</i> , como a divulgação de gastos públicos e processos de decisão. Gonzalez-Zapata e Heeks (2015) destacam a transparência como um dos principais objetivos de iniciativas de dados abertos.

Categoría a priori	Definição conceitual
Participação Cidadã	Explora como os DGA podem incentivar o engajamento da sociedade em decisões governamentais, incluindo consultas públicas e o desenvolvimento de aplicativos cívicos. Jetzek <i>et al.</i> (2014) ressaltam a participação cidadã como meio de gerar valor social por meio dos DGA.
Qualidade dos DGA	Avalia aspectos relacionados à precisão, acessibilidade e atualização dos dados abertos. Vetrò <i>et al.</i> (2016) propõem um <i>framework</i> para medir a qualidade dos DGA, enfatizando a importância de indicadores objetivos para garantir sua eficácia.
Desafios e Barreiras na Implementação de DGA	Discute os desafios técnicos, culturais e legais na implementação de políticas de dados abertos, como resistência à mudança e falta de capacitação técnica, legais, organizacionais e culturais, como a resistência à mudança por parte de servidores públicos (Safarov; Meijer; Grimmelikhuijsen, 2017). Estudos de Hossain <i>et al.</i> (2016) e Safarov <i>et al.</i> (2017) destacam as principais dificuldades na utilização de DGA.
Fatores que Influenciam a Adoção e Utilização de DGA	Investiga elementos que impactam a adoção dos DGA, como predisposição dos usuários e incentivos institucionais. Zuiderwijk <i>et al.</i> (2015) utilizam a Teoria Unificada de Aceitação e Uso da Tecnologia (UTAUT) para explorar os preditores de aceitação de tecnologias baseadas em DGA.

Fonte: elaboração própria a partir de Jetzek; Avital; Bjorn-Andersen (2014), Ahmadi Zeleti *et al.* (2016), Gonzalez-Zapata e Heeks (2015), Vetrò *et al.* (2016), Safarov; Meijer; Grimmelikhuijsen, (2017), Hossain *et al.* (2016) e Zuiderwijk *et al.* (2015).

Neste estudo realizou-se a análise de CHD, que é uma técnica de análise de dados que permite dividir o texto em segmentos de texto e classificar esses segmentos com base em padrões de frequência de palavras (Camargo; Justo, 2021).

O IRaMuTeQ utiliza um algoritmo de classificação para dividir o *corpus* (ou texto completo) em segmentos de texto, agrupando-os em classes distintas. Cada classe é caracterizada por um conjunto de palavras que ocorrem com maior frequência do que seria esperado ao acaso (Batista; Brandalise, 2023). Esse processo é denominado classificação descendente, porque o programa começa com o texto como um todo e, em seguida, o divide em segmentos menores ou classes (Camargo; Justo, 2013).

Depois de agrupados, o IRaMuTeQ pode exibir um dendrograma, que é um tipo de diagrama em árvore que ilustra a distribuição e a relação entre as classes. Usa-se também a AFC, um método estatístico que permite visualizar a relação entre as palavras e as variáveis em um gráfico bidimensional, fornecendo uma representação visual do agrupamento e da distância entre as variáveis. No contexto da sua análise do *corpus* textual, a AFC fornece uma maneira de visualizar as associações entre os temas tratados no *corpus* e as palavras ativas presentes em cada classe (Martins *et al.*, 2022).

Cada quadrante do gráfico representa uma das classes que foram identificadas na CHD. As classes que estão próximas uma da outra no gráfico (por exemplo, as que estão no mesmo quadrante ou em quadrantes adjacentes) tendem a compartilhar temas semelhantes ou relacionados, enquanto

as classes que estão mais distantes uma da outra (por exemplo, as que estão em quadrantes opostos) tendem a ter temas distintos ou opostos (Martins *et al.*, 2022).

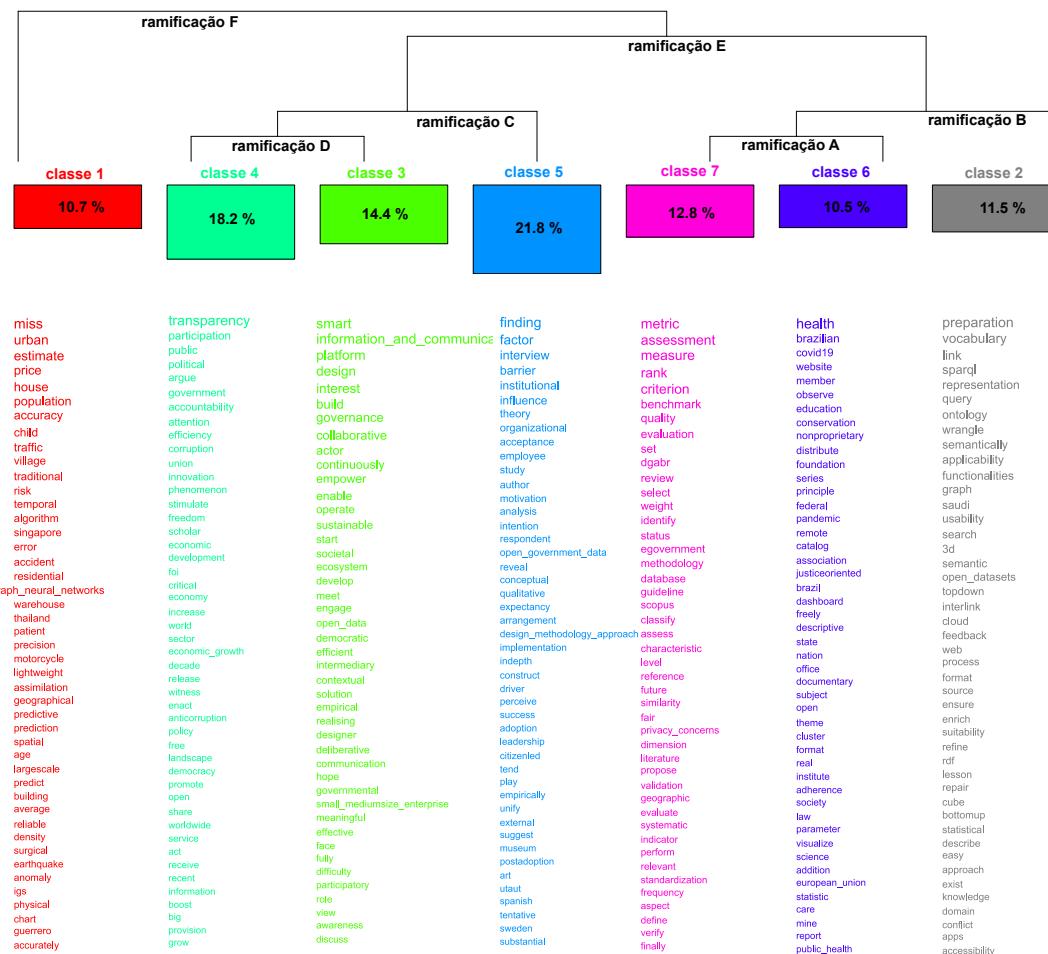
A análise de similitude se fundamenta na teoria dos grafos, uma área da matemática que auxilia no estudo das conexões entre elementos de um modelo matemático. No software IRaMuTeQ, essa análise produz um gráfico que ilustra a interligação das palavras presentes no corpo de texto. Isso possibilita inferir a estrutura textual e os temas de destaque, com base na ocorrência das palavras (Salviati, 2017).

4. RESULTADOS E DISCUSSÃO

O *corpus* textual foi construído por 332 resumos de artigo, que foi o resultado da unificação das duas bases de dados consultadas excluindo a duplicidade de artigos, como descrito na metodologia.

4.1 Análise de Classificação Hierárquica Descendente (CHD)

Esses 332 resumos de artigo foram separados em 1066 segmentos de texto (ST), com o aproveitamento de 1031 ST, ou seja, 96,71%. Emergiram 41472 ocorrências (palavras, formas ou vocabulários), sendo 4879 palavras distintas e 2071 com uma única ocorrência. Categorizou-se o resultado em sete classes: classe 1, com 110 ST (10,67%); classe 2, com 119 ST (11,54%); classe 3, com 149 ST (14,45%); classe 4, com 188 ST (18,23%); classe 5, com 225 ST (21,82%); classe 6, com 108 ST (10,48%); e classe 7, com 132 ST (12,80%), como ilustrado na Figura 4.

Figura 4 – Dendrograma do *corpus* textual IRaMuTeq

Fonte: elaboração própria.

Destaca-se que essas 7 classes estão divididas em seis ramificações (A, B, C, D, E e F) do *corpus* total em análise. A ramificação A, denominada Qualidade e Avaliação dos DGA e Saúde, compreende as classes 6 e 7. A classe 6, Métricas e Avaliação de Qualidade e Saúde, explora metodologias de otimização da qualidade dos dados abertos, com destaque para aplicações em saúde pública e em situações de crise, como a pandemia de covid-19.

A classe 7, Metodologias e Modelos Teóricos, foca no desenvolvimento de *frameworks* analíticos, *benchmarks* e modelos de maturidade para avaliar o progresso das iniciativas de dados abertos. Eles refletem um olhar técnico e operacional, voltado à consolidação de práticas consistentes e à melhoria da confiabilidade e usabilidade dos dados abertos.

Já a ramificação B, chamada Interoperabilidade e Integração de Dados, liga a ramificação A, classes 6 e 7, à classe 2. A classe 2, Tecnologias e Interoperabilidade, mostra os desafios e as soluções para integrar dados em diferentes contextos, como os padrões de metadados, amplia a

abordagem técnica e introduz elementos para a integração e reutilização de dados abertos, o que melhora a acessibilidade e a eficiência na dispersão de informações.

Aponta-se que a ramificação D, intitulada Impactos Sociais e Governança Regional, abrange as classes 3 e 5. A classe 5, Estratégias Governamentais e Diferenças Regionais, examina as políticas e culturas que moldam a adoção de dados abertos em diferentes regiões, com destaque para estudos comparativos entre países e contextos locais.

Por outro lado, a classe 3, Governança Local e Engajamento Cívico, enfatiza o papel dos dados abertos na democratização das decisões e na melhoria da participação cidadã em níveis locais. Evidencia, assim, as interações entre os dados abertos e os atores sociais, destacando tanto os benefícios quanto os desafios de sua implementação.

A ramificação C, Críticas e Barreiras Sistêmicas, integra o agrupamento 5-3 à classe 4. A classe 4, Dificuldades e Impactos na Sociedade, oferece uma perspectiva crítica dos desafios estruturais enfrentados pelas iniciativas de dados abertos, como desigualdades no acesso digital e limitações na capacidade de promover mudanças. Adiciona-se, assim, uma dimensão reflexiva sobre os limites e as implicações práticas da adoção de dados abertos.

Na ramificação E, chamada Domínio Técnico-Social Integrado, os subcorpos anteriores consolidam as dimensões técnicas e sociais em um conjunto, que reflete a interconexão entre práticas tecnológicas avançadas e os impactos sociais das iniciativas de dados abertos, destacando o potencial transformador da transparência e da inovação nos contextos público e privado.

Por fim, aponta-se que o subcorpo independente, denominado Foco Urbano e Aplicações Práticas, é representado pela classe 1, que se destaca por seu enfoque nas aplicações práticas de dados abertos em contextos urbanos, como o planejamento de infraestrutura inteligente e a gestão de serviços públicos. Embora conectado ao restante do dendrograma pela ramificação F em um nível mais elevado, ele atua como um pilar autônomo, representando o uso tangível e aplicado dos dados abertos em cenários específicos e independentes.

As correlações entre as classes identificadas na análise *a priori* e aquelas derivadas da análise do IRaMuTeQ revelam uma convergência significativa nos temas de impacto social e econômico dos DGA. Além disso, há uma convergência nos desafios de implementação dos DGA. Identificaram-se, ainda, categorias como Impacto Econômico e Impacto Social dos DGA, que encontram eco nas Classes 5 e 6, as quais abordam, respectivamente, estratégias regionais e qualidade dos dados com enfoque em saúde pública. Essas interseções sugerem que os DGA podem não apenas impulsionar a economia regional, como também promover melhorias substanciais na qualidade de vida, especialmente quando aplicados a serviços essenciais, como saúde.

Além disso, a Classe 3, que trata da Governança Local e Engajamento Cívico, está fortemente alinhada com a categoria de Participação Cidadã da análise *a priori*. Ambas destacam o papel crucial dos DGA em fomentar a participação dos cidadãos e fortalecer a governança local, oferecendo

aos cidadãos ferramentas e informações necessárias para se envolverem ativamente nas decisões governamentais. Esse ponto é particularmente importante quando se observa a crescente demanda por maior transparência e *accountability*, aspectos que também são abordados pela Classe 4.

4.2 Análise Fatorial de Correspondência (AFC)

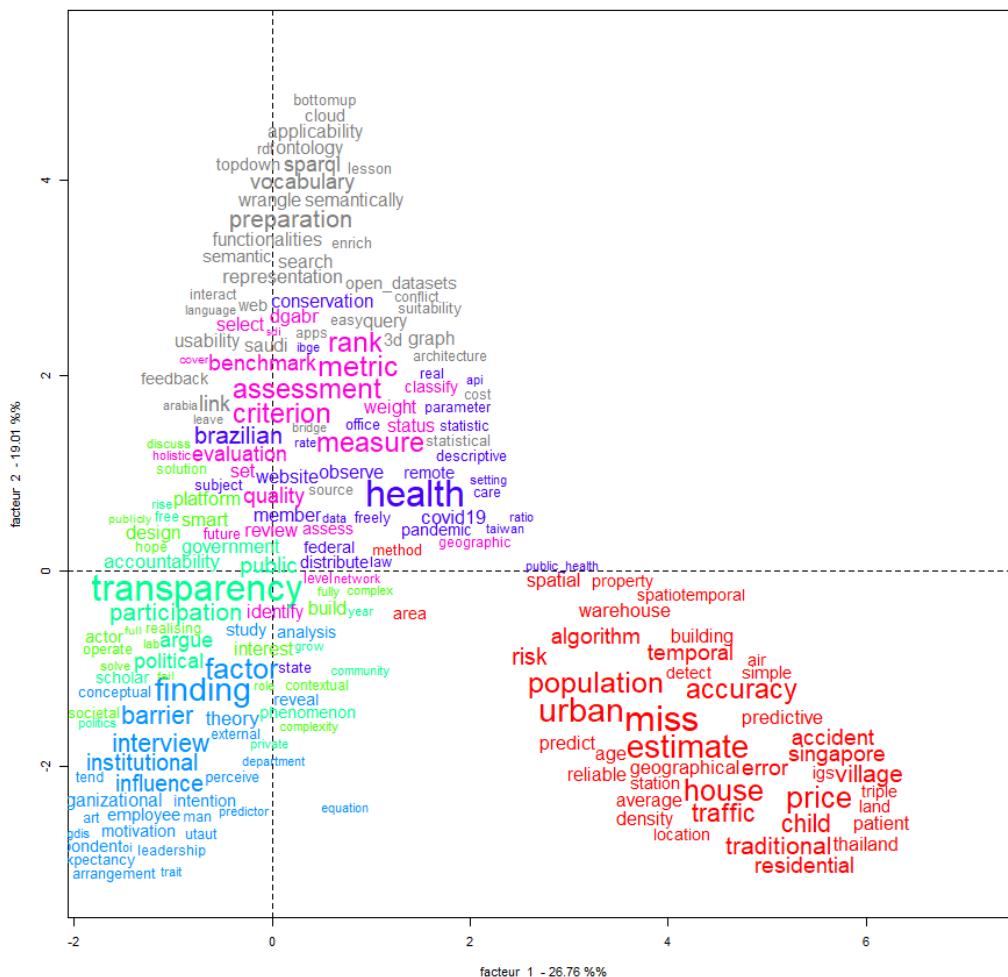
A AFC, Figura 5, proporciona uma visualização bidimensional que ajuda a compreensão das relações temáticas no campo dos DGA. A partir da AFC, foi possível observar como as palavras do *corpus* se associam, sendo representadas em um plano cartesiano. Essa visualização permite compreender melhor as relações entre os temas e como eles se conectam no contexto dos DGA.

As classes 6 (Métricas e Avaliação de Qualidade e Saúde) e 7 (Metodologias e Modelos Teóricos) demonstram proximidade no gráfico, sugerindo uma relação temática importante. Ambas estão ligadas ao uso e à qualidade dos DGA, mas cada uma tem sua ênfase específica. A classe 6 destaca o impacto dos DGA em áreas como saúde pública, especialmente no Brasil, evidenciado por termos como '*health*', '*brazilian*' e '*covid19*'.

Já a classe 7 se concentra em métricas e avaliação, refletindo uma preocupação com a medição da qualidade e a definição de padrões, com palavras como '*metric*', '*assessment*' e '*benchmark*'. Essa proximidade sugere que o sucesso das aplicações, em saúde por exemplo, depende diretamente de uma análise cuidadosa da qualidade dos dados.

Por outro lado, as classes 3 (Governança Local e Engajamento Cívico) e 4 (Dificuldades e Impactos na Sociedade) também aparecem próximas, indicando uma forte conexão entre os temas de governança colaborativa e práticas democráticas. A classe 3, com termos como '*platform*', '*design*' e '*collaboration*', aborda a construção de sistemas inteligentes e colaborativos para dados abertos. A classe 4, por sua vez, foca na transparência e participação cidadã, com destaque para palavras como '*transparency*', '*participation*' e '*accountability*'. Juntas, essas classes reforçam a ideia de que práticas colaborativas e transparentes caminham lado a lado, promovendo confiança e engajamento público.

Figura 5 – AFC das palavras ativas mais frequentes em cada classe



Fonte: elaboração própria.

O agrupamento de palavras em azul claro e verde, na parte inferior esquerda, reflete a proximidade temática entre as classes 3 e 4. Palavras como ‘*platform*’, ‘*design*’ e ‘*governance*’ demonstram a interseção entre governança local, engajamento cívico e participação democrática. Ao mesmo tempo, essas classes mantêm uma correlação com palavras em azul escuro da classe 5 (Estratégias Governamentais e Diferenças Regionais), como ‘*factor*’ e ‘*institutional*’, que tratam das barreiras institucionais que influenciam a implementação dos DGA. Observa-se que essa conexão reforça a superação de desafios organizacionais e culturais, para que práticas de governança aberta e participação se tornem viáveis.

A classe 1 (Foco Urbano e Aplicações Práticas), em vermelho, ocupa uma posição isolada no gráfico, refletindo seu caráter específico e técnico e seu isolamento temático. Concentra-se no uso dos DGA em análises preditivas e desafios urbanos, como planejamento de tráfego, estimativas populacionais e questões habitacionais, evidenciado por termos como ‘*urban*’, ‘*predictive*’, ‘*population*’ e ‘*traffic*’. Aponta-se como os DGA podem ser aplicados de maneira

direcionada na solução de problemas específicos, como planejamento de tráfego, estimativas populacionais e gestão urbana. Embora isolada, a classe 1 se conecta indiretamente às classes técnicas, destacando como o rigor na preparação e avaliação dos dados contribui para a efetividade dessas aplicações práticas.

Cabe salientar que a análise das cores também permite observar como as classes técnicas, representadas por palavras em azul e roxo, interagem com as classes sociais e práticas, destacadas por verde, vermelho e azul claro. A interseção entre palavras como '*health*' (classe 6, roxo) e '*brazilian*' (classe 7, azul), ambas relacionadas à saúde pública, reflete uma complementaridade entre abordagens técnicas e contextos específicos de aplicação. Da mesma forma, a correlação entre '*participation*' (classe 3, verde) e '*transparency*' (classe 4, verde claro) reforça a ideia de que os DGA não apenas reforçam a governança, mas também dependem de estruturas sociais inclusivas para atingir seus objetivos.

Constatou-se que há diversidade nas abordagens aos DGA, com temas voltados para questões técnicas e específicas, como saúde e urbanismo, e, em contrapartida, temas mais amplos, ligados à governança, participação e avaliação de dados. Essa interação entre os diferentes grupos reforça a importância de uma abordagem integrada, em que a qualidade técnica dos dados, sua transparência e sua aplicação prática se complementam.

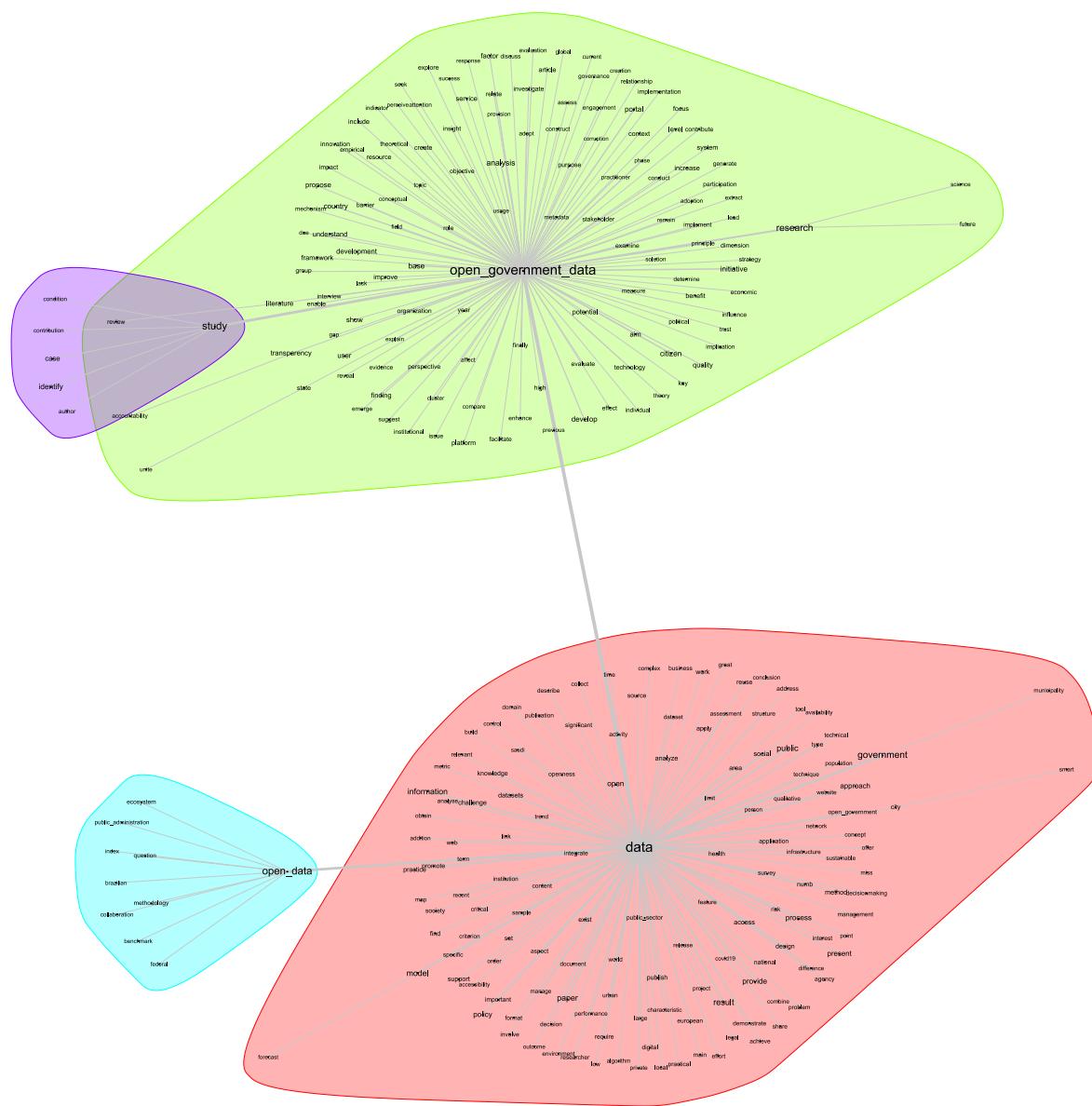
4.3 Análise de Similitude

O produto da Análise de Similitude é ilustrado como um gráfico, que revela as correlações e conexões entre as palavras existentes no corpo de texto, com base nas palavras ativas identificadas (Camargo; Justo, 2021).

A análise de similitude aplicada ao *corpus* textual sobre DGA, representada na Figura 6, revela a estrutura semântica subjacente do texto, destacando as interconexões entre os principais termos e conceitos por meio de grafos. Esse método, baseado na co-ocorrência de palavras e suas frequências, permite mapear os núcleos discursivos que organizam o campo temático.

No grafo gerado, os termos centrais emergem como pivôs do discurso, enquanto os clusters que orbitam esses termos evidenciam os temas que estruturam o debate. De maneira complementar, a análise demonstra como os DGA são abordados em múltiplos aspectos, desde as dimensões técnicas até as aplicações práticas e seus impactos sociais.

Figura 6 – Grafo de Similitude



Fonte: elaboração própria.

Para melhorar a legibilidade e reduzir a complexidade visual do gráfico, adotou- se, como ponto de corte, o número de ocorrências de cada palavra equivalente a 25 vezes ou mais, excluindo as palavras que não corroboram a análise do gráfico e marcando as opções de comunidade e halo na geração do gráfico.

O termo '*data*' surge como um núcleo central no grafo, evidenciando seu papel fundamental na construção do discurso. Ao seu redor, concentram-se palavras relacionadas a governança, políticas públicas e uso de dados em contextos sociais. Termos como '*government*', '*public*', e '*information*' indicam um enfoque significativo nos desafios e oportunidades associados à transparéncia governamental e à democratização da informação. Essa rede reflete a visão de que

os DGA são um recurso essencial para melhorar processos de tomada de decisão, promover *accountability* e aumentar a participação pública.

Ressalta-se que o termo '*open_government_data*' forma outro ponto de agrupamento no grafo, com conexões que destacam seu papel em aplicações específicas e estratégias de implementação. Palavras como '*transparency*', '*accountability*' e '*participation*' aparecem como colunas associadas a esse conceito, sugerindo que o *corpus* enfatiza os DGA como ferramentas para fortalecer práticas democráticas e criar espaços de engajamento cívico. A proximidade entre esses termos sugere que o uso de dados abertos vai além de uma abordagem técnica e alcança implicações sociais e políticas.

Adicionalmente, o grafo identifica *clusters* mais especializados que se conectam aos núcleos centrais. Um desses *clusters*, associado ao termo '*study*', destaca o uso dos DGA em contextos de pesquisa acadêmica e inovação. Palavras como '*research*', '*case*', e '*author*' revelam um interesse significativo em explorar os DGA como objeto de estudo e como fonte de dados para análises científicas. Esse agrupamento reflete a relevância dos DGA para sustentar avanços no conhecimento e na elaboração de novos modelos analíticos.

Outro agrupamento, relacionado ao termo '*open_data*', está conectado a discussões mais técnicas e práticas. Palavras como '*methodology*', '*benchmark*' e '*collaboration*' sugerem um enfoque na avaliação de qualidade e na interoperabilidade dos dados abertos, evidenciando a necessidade de infraestrutura sólida e práticas colaborativas para garantir que os DGA sejam acessíveis, confiáveis e úteis em diversos contextos.

Além disso, o grafo apresenta um *cluster* específico associado a aplicações práticas em contextos urbanos e demográficos. Palavras como '*municipality*', '*traffic*' e '*population*' demonstram a utilização de DGA para resolver problemas concretos, como planejamento urbano, estimativas populacionais e gestão de serviços públicos. Esse agrupamento reforça a visão de que os DGA não são apenas uma ferramenta estratégica de governança, mas também possuem impacto direto em soluções práticas.

As conexões observadas refletem uma dinâmica integrada e complementar entre as dimensões técnicas, sociais e aplicadas dos DGA. Os núcleos discursivos, como '*data*' e '*open_government_data*', funcionam como pontos de convergência, conectando elementos que vão desde o rigor metodológico até a democratização do acesso e uso de informações. Essa integração reforça que o campo dos DGA combina uma base técnica com ambições de impacto social e inovação, adaptando-se a diferentes demandas e contextos.

Dessa forma, a análise de similitude expõe como os DGA operam como um elemento central em discursos que conectam tecnologias, governança e sociedade. A representação gráfica não apenas organiza as relações semânticas do *corpus*, mas também evidencia as áreas em que o impacto dos DGA é mais importante, destacando os desafios e potencialidades desse campo.

5. CONSIDERAÇÕES FINAIS

Este estudo buscou compreender como os DGA têm sido utilizados, mapeando suas principais aplicações, desafios e contribuições para a sociedade e a governança pública. O referencial teórico destacou a centralidade dos DGA no fortalecimento da transparência, da *accountability* e do engajamento cívico, revelando que a abertura de dados não é apenas uma questão técnica, mas também um instrumento de transformação social. Essa análise teórica permitiu integrar perspectivas sobre governança, inovação e participação cidadã, situando os DGA como um tema que conecta tecnologias emergentes a demandas sociais e econômicas.

Por meio da análise de conteúdo, revelou-se a existência de sete classes temáticas principais, que organizam as discussões acadêmicas sobre DGA. Elas abrangem desde aspectos técnicos, como interoperabilidade e qualidade dos dados, inclusive dimensões sociais, como governança participativa até impactos econômicos, demonstrando a complexidade e a interconexão dos debates.

Os resultados indicaram que as publicações científicas sobre DGA orbitam em torno de núcleos semânticos bem definidos, representados por termos centrais como '*data*' e '*open_government_data*'. Essas palavras-chave articulam discussões que conectam governança, transparência, inovação e aplicações práticas. A análise de similitude evidenciou que, embora existam *clusters* temáticos distintos – como os voltados para saúde pública, planejamento urbano e desenvolvimento de *frameworks* analíticos –, eles estão interligados, reforçando a importância da abordagem multidisciplinar no campo dos DGA.

Este estudo identificou uma estrutura temática que conecta aspectos técnicos, sociais e práticos dos DGA, mostrando que o sucesso dessas iniciativas depende de um equilíbrio entre qualidade técnica, governança e impacto social. Esse equilíbrio é relevante em um contexto em que os dados não são usados somente para suportar os processos administrativos, mas também como base para a transformação digital dos serviços públicos, o engajamento cidadão e a formulação de políticas mais inclusivas.

Conclui-se que os DGA são, além de um recurso tecnológico, um componente importante para a construção de uma governança mais transparente e participativa. Ao mapear as discussões acadêmicas propondo o avanço do conhecimento, este estudo reforça o papel estratégico dos DGA no enfrentamento dos desafios contemporâneos da administração pública e no fortalecimento das democracias.

Cabe destacar, como limitação deste estudo, a ausência de análises qualitativas aprofundadas de casos emblemáticos, bem como a não delimitação específica ao contexto latino-americano ou do Sul Global. Embora os resultados tragam uma visão abrangente, tais recortes poderiam enriquecer

a compreensão sobre especificidades regionais e trajetórias históricas distintas. Recomenda-se que pesquisas futuras explorem essa dimensão comparativa.

Como apporte para pesquisas futuras, sugere-se ainda: a) aprofundar investigações sobre o impacto dos DGA na transformação digital dos serviços públicos, especialmente em países em desenvolvimento; b) realizar estudos comparativos entre políticas de DGA em diferentes contextos e culturas institucionais; c) desenvolver *frameworks* para avaliação da efetividade dos DGA em fomentar a participação cidadã e a inovação social, bem como do impacto dessas iniciativas na redução de desigualdades digitais e no fortalecimento da confiança pública.

REFERÊNCIAS

AHMADI ZELETI, Fatemeh; OJO, Adegboyega; CURRY, Edward. Exploring the economic value of open government data. **Government Information Quarterly**, [s. l.], v. 33, n. 3, p. 535-551, 2016. DOI: <https://doi.org/10.1016/j.giq.2016.01.008>.

ALBANO, Cláudio Sonáglio; CRAVEIRO, Gisele Da Silva; CAROSSI, Daniel Fernando. Oferta e demanda de dados abertos governamentais: um estudo em universidades federais brasileiras. **Revista de Gestão e Secretariado (Management and Administrative Professional Review)**, [s. l.], v. 14, n. 9, p. 15191-15213, 2023. DOI: <https://doi.org/10.7769/gesec.v14i9.2794>.

ATTARD, Judie; ORLANDI, Fabrizio; SCERRI, simon; AUER, Sören. A systematic review of open government data initiatives. **Government Information Quarterly**, [s. l.], v. 32, n. 4, p. 399-418, 2015. DOI: <https://doi.org/10.1016/j.giq.2015.07.006>.

BARDIN, Laurence. **Análise de conteúdo**. 3a reimpressão da 1a Edição de 2011 ed. [S. l.]: Edições 70, 2011.

BATISTA, Ruhama Ariella Sabião; BRANDALISE, Mary Ângela Teixeira. A utilização do software Iramuteq na análise de dados textuais em revisão sistemática de literatura. **Roteiro**, [s. l.], v. 48, 2023. DOI: <https://doi.org/10.18593/r.v48.32352>. Disponível em: <https://periodicos.unoesc.edu.br/roteiro/article/view/32352>. Acesso em: 11 nov. 2024.

BRASIL. **Decreto no 8.777, de 11 de maio de 2016**. Institui a Política de Dados Abertos do Poder Executivo Federal. Brasília, 2016. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2016/decreto/D8777.htm. Acesso em: 18 maio 2024.

BRASIL. **Decreto no 9.903, de 8 de julho de 2019**. Altera o Decreto no 8.777, de 11 de maio de 2016, que institui a Política de Dados Abertos do Poder Executivo Federal, para dispor sobre a gestão e os direitos de uso de dados abertos. Brasília, 2019. Disponível em: https://www.planalto.gov.br/ccivil_03/_Ato2019-2022/2019/Decreto/D9903.htm#art1. Acesso em: 18 maio 2024.

BRASIL. **Lei Complementar nº 131**. 27 maio 2009. Disponível em: https://www.planalto.gov.br/ccivil_03/leis/lcp/lcp131.htm. Acesso em: 14 set. 2024.

BRASIL. **Lei no 12.527, de 18 de novembro de 2011**. Regula o acesso a informações previsto no inciso XXXIII do art. 5o, no inciso II do § 3o do art. 37 e no § 2o do art. 216 da Constituição Federal; altera a

Lei no 8.112, de 11 de dezembro de 1990; revoga a Lei no 11.111, de 5 de maio de 2005, e dispositivos da Lei no 8.159, de 8 de janeiro de 1991; e dá outras providências. (LAI). LAI, Brasília, 2011. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/l12527.htm. Acesso em: 18 maio 2024.

BRASIL. Referencial básico de governança aplicável a organizações públicas e outros entes jurisdicionados ao TCU. 3. ed. Brasília: TCU, Secretaria de Controle Externo da Administração do Estado – SecexAdministração, 2020. Disponível em: <https://portal.tcu.gov.br/lumis/portal/file/fileDownload.jsp?fileId=8A81881F7595543501762EB92E957799>. Acesso em: 11 jun. 2024.

CAMARGO, Brígido V.; JUSTO, Ana M. IRAMUTEQ: um software gratuito para análise de dados textuais. **Temas em Psicologia**, [s. l.], v. 21, n. 2, p. 513-518, 2013. DOI: <https://doi.org/10.9788/TP2013.2-16>.

CAMARGO, Brigido V.; JUSTO, Ana M. **Tutorial para uso do software IRAMUTEQ (Interface de R pour les Analyses Multidimensionnelles de Textes et de Questionnaires)**. [S. l.: s. n.], 2021. Disponível em: http://iramuteq.org/documentation/fichiers/Tutorial%20IRaMuTeQ%20em%20portugues_22.11.2021.pdf. Acesso em: 26 maio 2023.

CHIGNARD, Simon. **A brief history of Open Data**. [S. l.], 2013. Disponível em: <https://www.paristechreview.com/2013/03/29/brief-history-open-data/>. Acesso em: 14 set. 2024.

COSTA, Danilo da; SILVA, Gustavo Javier Castro; ASSUNÇÃO, Maria Aparecida de. **Scopus vs. Web of Science**: uma avaliação comparativa das principais bases de dados para a pesquisa acadêmica. [s. l.], 2023. DOI: <https://doi.org/10.5281/ZENODO.10777653>. Disponível em: <https://zenodo.org/doi/10.5281/zenodo.10777653>. Acesso em: 19 jul. 2024.

GONZALEZ-ZAPATA, Felipe; HEEKS, Richard. The multiple meanings of open government data: understanding different stakeholders and their perspectives. **Government Information Quarterly**, [s. l.], v. 32, n. 4, p. 441-452, 2015. DOI: <https://doi.org/10.1016/j.giq.2015.09.001>.

GRAY, Jonathan. Towards a genealogy of Open Data. **SSRN Electronic Journal**, [s. l.], 2014. DOI: <https://doi.org/10.2139/ssrn.2605828>. Disponível em: <http://www.ssrn.com/abstract=2605828>. Acesso em: 28 abr. 2024.

GURSTEIN, Michael B. Open Data: empowering the empowered or effective data use for everyone? **First Monday**, [s. l.], 2011. DOI: <https://doi.org/10.5210/fm.v16i2.3316>. Disponível em: <https://firstmonday.org/ojs/index.php/fm/article/view/3316>. Acesso em: 31 maio 2024.

HOSSAIN, Mohammad Alamgir; DWIVEDI, Yogesh K.; RANA, Nripendra P. State-of-the-art in open data research: insights from existing literature and a research agenda. **Journal of Organizational Computing and Electronic Commerce**, [s. l.], v. 26, n. 1–2, p. 14-40, 2016. DOI: <https://doi.org/10.1080/10919392.2015.1124007>.

JETZEK, Thorhildur; AVITAL, Michel; BJORN-ANDERSEN, Niels. Data-driven innovation through open government data. **Journal of Theoretical and Applied Electronic Commerce Research**, [s. l.], v. 9, n. 2, p. 100-120, 2014. DOI: <https://doi.org/10.4067/S0718-18762014000200008>.

KLEIMAN, Fernando *et al.* Understanding civil servants' intentions to open data: factors influencing behavior to disclose data. **Information Technology and People**, [s. l.], v. 37, n. 2, p. 729-752, 2024. DOI: <https://doi.org/10.1108/ITP-12-2020-0893>.

MACEDO, Dirceu Flávio; LEMOS, Daniela Lucas da Silva. Government open data: initiatives and challenges in data opening in Brazil and other international spheres; [Dados abertos governamentais: iniciativas e desafios na abertura de dados no Brasil e outras esferas internacionais]. *AtoZ*, [s. l.], v. 10, n. 2, p. 14-26, 2021. DOI: <https://doi.org/10.5380/atoz.v10i2.77737>.

MARTÍN-MARTÍN, Alberto; ORDUNA-MALEA, Enrique; THELWALL, Mike; LÓPEZ-CÓZAR, Emílio Delgado. **Google Scholar, Web of Science, and Scopus**: a systematic comparison of citations in 252 subject categories. [S. l.]: arXiv, 2019. DOI: <https://doi.org/10.1016/j.joi.2018.09.002>. Disponível em: <http://arxiv.org/abs/1808.05053>. Acesso em: 11 nov. 2024.

MARTINS, Kaique Nascimento *et al.* O software IRaMuTeQ como recurso para a análise textual discursiva. **Revista Pesquisa Qualitativa**, [s. l.], v. 10, n. 24, p. 213-232, 2022. DOI: <https://doi.org/10.33361/RPQ.2022.v10.n.24.383>.

OPEN GOVERNMENT PARTNERSHIP. Open Government Declaration, 2011. Disponível em: <https://www.opengovpartnership.org/process/joining-ogp/open-government-declaration/>. Acesso em: 14 set. 2024.

OPEN DATA CHARTER. [S. l.], 2013. Disponível em: <https://www.gov.uk/government/publications/open-data-charter>. Acesso em: 14 set. 2024.

PASCOAL, Márcia Néa Oliveira; OLIVEIRA, Oderlene Vieira de. Práticas de governança pública adotadas pela administração pública federal brasileira. **Administração Pública e Gestão Social**, v. 11, n. 2, p. 215-231, 2019. DOI: <https://doi.org/10.21118/apgs.v11i2.5431>.

POSSAMAI, Ana Júlia; SOUZA, Vitoria Gonzatti de. Transparência e dados abertos governamentais: possibilidades e desafios a partir da Lei de Acesso à Informação. **Administração Pública e Gestão Social**, [s. l.], v. 12, 2020. DOI: <https://doi.org/10.21118/apgs.v12i2.5872>. Disponível em: <https://periodicos.ufv.br/apgs/article/view/5872>. Acesso em: 3 out. 2023.

REINBERG, Bernhard; SWEDLUND, Haley. How transparent are aid agencies to their citizens? Introducing the Citizen Aid Transparency Dataset. **Journal of International Development**, [s. l.], v. 35, n. 7, p. 2177-2212, 2023. DOI: <https://doi.org/10.1002/jid.3762>.

SAFAROV, Iqbal; MEIJER, Albert; GRIMMELIKHUIJSEN, Stephan. Utilization of open government data: a systematic literature review of types, conditions, effects and users. **Information Polity**, [s. l.], v. 22, n. 1, p. 1-24, 2017. DOI: <https://doi.org/10.3233/IP-160012>.

SALVIATI, Maria Elisabeth. **Manual do Aplicativo Iramuteq (versão 0.7 Alpha 2 e R Versão 3.2.3)**. [S. l.: s. n.], 2017. Disponível em: <http://iramuteq.org/documentation/fichiers/manual-do-aplicativo-iramuteq-par-maria-elisabeth-salviati>. Acesso em: 26 maio 2023.

SNYDER, Hannah. Literature review as a research methodology: an overview and guidelines. **Journal of Business Research**, [s. l.], v. 104, p. 333-339, 2019. DOI: <https://doi.org/10.1016/j.jbusres.2019.07.039>.

TENNANT, Jonathan P. *et al.* The academic, economic and societal impacts of Open Access: an evidence-based review. **F1000Research**, [s. l.], v. 5, p. 632, 2016. DOI: <https://doi.org/10.12688/f1000research.8460.3>.

UNITED NATIONS. **Universal Declaration of Human Rights**. Paris, 1948. The Office of the High Commissioner for Human Rights (OHCHR). Disponível em: <https://www.ohchr.org/en/human-rights/universal-declaration/translations/english>. Acesso em: 28 set. 2024.

VETRÒ, Antonio *et al.* Open data quality measurement framework: definition and application to Open Government Data. **Government Information Quarterly**, [s. l.], v. 33, n. 2, p. 325-337, 2016. DOI: <https://doi.org/10.1016/j.giq.2016.02.001>.

WEBSTER, Jane; WATSON, Richard T. Analyzing the past to prepare for the future: writing a literature review. **MIS Quarterly**, [s. l.], v. 26, n. 2, p. xiii–xxiii, 2002.

ZUIDERWIJK, Anneke; JANSSEN, Marijn; DWIVEDI, Yogesh K. Acceptance and use predictors of open data technologies: drawing upon the unified theory of acceptance and use of technology. **Government Information Quarterly**, [s. l.], v. 32, n. 4, p. 429-440, 2015. DOI: <https://doi.org/10.1016/j.giq.2015.09.005>.

Renato Machado de Godoy

 <https://orcid.org/0009-0005-1232-5894>

Mestrando em Administração Pública pela Universidade Federal dos Vales do Jequitinhonha e Mucuri (UFVJM). Especialista em Redes de Computadores pela Pontifícia Universidade Católica de Minas Gerais (PUC Minas). Bacharel em Ciência da Computação pela Universidade Federal de Ouro Preto (UFOP). Servidor do Instituto Federal de Minas Gerais (IFMG).

renato.godoy@gmail.com

João Paulo Calembo Batista Menezes

 <https://orcid.org/0000-0002-4739-839X>

Doutor em Administração pela Universidade Federal de Minas Gerais (UFMG). Mestre em Contabilidade, Fiscalidade e Finanças pela Universidade de Lisboa (ULisboa). Especialista em Ciência de Dados e Big Data (PUCMinas). Bacharel em Contabilidade (PUCMinas), Administração (PUCMinas) e Economia (Faculdade Arnaldo). Professor Adjunto na Universidade Federal dos Vales do Jequitinhonha e Mucuri (UFVJM).

joao.calembo@ufvjm.edu.br

João Cesar de Souza Ferreira

 <https://orcid.org/0000-0002-3705-8746>

Doutor em Administração pela Universidade Municipal de São Caetano do Sul (USCS). Mestre em Gestão de Instituições Educacionais pela Universidade Federal dos Vales do Jequitinhonha e Mucuri (UFVJM). Bacharel em Administração pela Fundação Educacional Nordeste Mineiro (FENM). Professor Assistente na UFVJM.

joao.cesar@ufvjm.edu.br

Marcio Coutinho de Souza

 <https://orcid.org/0000-0002-4238-1572>

Doutor e Mestre em Engenharia de Produção pela Unimep. Bacharel em Administração pela Univila. Professor na Universidade Federal dos Vales do Jequitinhonha e Mucuri (UFVJM/DEPAD) e membro do Núcleo de Estudos e Pesquisa em Gestão e Desenvolvimento Regional (NEGED) da UFVJM.

marcio.souza@ufvjm.edu.br