MENSURANDO A SALIÊNCIA FISCAL NA COMUNICAÇÃO INSTITUCIONAL: DESENVOLVIMENTO E VALIDAÇÃO DE UM INDICADOR BASEADO EM MINERAÇÃO TEXTUAL PARA A SEF-MG

MEASURING FISCAL SALIENCE IN INSTITUTIONAL COMMUNICATION: DEVELOPMENT AND VALIDATION OF AN INDICATOR BASED ON TEXTUAL MINING FOR SEF-MG

https://doi.org/10.63330/aurumpub.069-037

Jean Lucas Viana Almeida
Mestrando
Universidade Federal dos Vales do Jequitinhonha e Mucuri
E-mail: jeanlucasviana@hotmail.com

Geórgia Fernandes Barros
Doutora
Universidade Federal dos Vales do Jequitinhonha e Mucuri
E-mail: georgia.barros@ufvjm.edu.br

Michel Cândido de Souza
Doutor
Universidade Federal dos Vales do Jequitinhonha e Mucuri
E-mail: michel.cands@gmail.com

Resumo

Este estudo desenvolveu um Indicador de Saliência Fiscal (ISF) da comunicação institucional da Secretaria de Estado de Fazenda de Minas Gerais (SEF/MG), utilizando técnicas de mineração textual. A pesquisa possui abordagem quantitativa, caráter descritivo e delineamento documental, tendo como corpus notícias institucionais publicadas entre 2007 e 2022. A metodologia compreendeu auditoria e reconstrução do corpus textual, identificação de zeros artificiais, elaboração de um dicionário lexical estruturado em cinco dimensões da gestão fiscal e processamento dos dados no software IRaMuTeQ. A Classificação Hierárquica Descendente e a análise de similitude evidenciaram elevada consistência entre as classes lexicais, as redes semânticas e as dimensões teóricas do dicionário, validando o indicador proposto. Os resultados demonstram que a comunicação institucional reflete mudanças nas prioridades da gestão fiscal, especialmente em temas relacionados ao endividamento, arrecadação, governança e recuperação fiscal. O ISF constitui uma contribuição metodológica inédita para mensurar a saliência fiscal da comunicação pública e oferece um instrumento replicável para estudos sobre governança, transparência e administração pública.

Palavras-chave: Administração pública; Comunicação institucional; Gestão fiscal; Mineração textual; Saliência fiscal.

Abstract

This study developed a Fiscal Salience Indicator (FSI) for the institutional communication of the State Department of Finance of Minas Gerais (SEF/MG), using text mining techniques. The research has a quantitative approach, descriptive nature, and documentary design, with a corpus consisting of institutional news published between 2007 and 2022. The methodology included auditing and reconstructing the textual corpus, identifying artificial zeros, creating a lexical dictionary structured into five dimensions of fiscal management, and processing the data using the IRaMuTeQ software. Descending Hierarchical Classification and similarity analysis showed high consistency between lexical classes, semantic networks, and the theoretical dimensions of the dictionary, validating the proposed indicator. The results demonstrate that institutional communication reflects changes in fiscal management priorities, especially on topics related to debt and revenue, governance and fiscal recovery. The ISF represents a groundbreaking methodological contribution to measure the fiscal salience of public communication and provides a replicable tool for studies on governance, transparency, and public administration.

Keywords: Fiscal management; Fiscal salience; Institutional communication; Public administration; Text mining.

1 INTRODUÇÃO

A crescente complexidade da gestão das finanças públicas tem ampliado a necessidade de instrumentos capazes de compreender como os governos comunicam suas prioridades fiscais. Nesse contexto, a comunicação institucional constitui um importante mecanismo de transparência e prestação de contas, além de refletir as estratégias adotadas pelos órgãos públicos diante das mudanças do ambiente econômico.

Os avanços da mineração textual possibilitam transformar documentos oficiais em indicadores capazes de mensurar padrões discursivos da administração pública. Entretanto, a construção desses indicadores exige bases documentais consistentes, pois falhas na organização do corpus e lacunas de informação podem comprometer a representatividade dos resultados e introduzir vieses analíticos.

Diante desse cenário, este estudo propõe uma metodologia para desenvolver e validar o Indicador de Saliência Fiscal (ISF), construído a partir das notícias oficiais da Secretaria de Estado de Fazenda de Minas Gerais (SEF/MG). A proposta integra auditoria do corpus documental, elaboração de um dicionário

lexical fiscal, normalização das frequências textuais e tratamento dos períodos sem comunicação institucional.

O objetivo deste artigo é desenvolver um indicador de saliência fiscal da comunicação institucional baseado em mineração textual de documentos oficiais da SEF/MG. Especificamente, busca-se construir e auditar a base documental das notícias oficiais (2007–2022), elaborar um dicionário lexical fiscal estruturado em dimensões temáticas, operacionalizar o ISF trimestral com tratamento dos períodos sem comunicação e avaliar evidências de validade de conteúdo e de critério do indicador.

A principal contribuição do estudo consiste na proposição de um procedimento metodológico reproduzível para construção de indicadores de saliência fiscal aplicados à comunicação institucional de organizações públicas. A metodologia amplia a confiabilidade das análises baseadas em mineração textual e pode ser adaptada a diferentes contextos governamentais, contribuindo para pesquisas em gestão pública, transparência e comunicação institucional.

2 REFERENCIAL TEÓRICO

2.1 SECRETARIA DA FAZENDA DE MINAS GERAIS

A administração fazendária em Minas Gerais possui uma trajetória histórica, acompanhando as transformações políticas, econômicas e institucionais do estado ao longo dos séculos (AFFEMG, 1997; Batista, 2001 e SEF, 1991).

Nesse processo, a Secretaria de Estado de Fazenda de Minas Gerais consolidou-se como um órgão fundamental para a organização financeira estadual, passando por constantes adaptações até alcançar os atuais modelos de modernização administrativa, voltados à reestruturação da área tributária e ao aperfeiçoamento das técnicas de planejamento e gestão pública (Batista, 2001; SEF, 2026).

A Secretaria de Estado de Fazenda de Minas Gerais (SEF/MG) é responsável pela gestão das finanças públicas estaduais e pela arrecadação dos tributos que sustentam o funcionamento do estado. Entre suas principais atribuições estão a arrecadação de impostos como o ICMS, o IPVA e o ITCMD, além da fiscalização tributária e da autorização para emissão de notas fiscais. Dessa forma, a SEF/MG exerce papel estratégico na manutenção da capacidade financeira do estado e na viabilização das políticas públicas (SEF, 2026).

Além da função arrecadatória, a Secretaria também participa diretamente do planejamento e da execução do orçamento estadual em conjunto com a Secretaria de Estado de Planejamento e Gestão (Seplag-MG). Esse processo é estruturado por instrumentos legais e gerenciais que orientam a arrecadação e os gastos públicos, destacando-se o Plano Plurianual de Ações Governamentais (PPAG), a Lei de Diretrizes Orçamentárias (LDO) e a Lei Orçamentária Anual (LOA). Esses mecanismos permitem maior

organização, transparência e eficiência na administração dos recursos públicos estaduais (Agência Minas, 2026).

2.2 DÍVIDA PÚBLICA E LEI DE RESPONSABILIDADE FISCAL

A dívida pública ocupa posição central nas discussões sobre equilíbrio fiscal e desenvolvimento econômico, especialmente diante das constantes transformações do cenário econômico nacional e internacional. Sua gestão exige planejamento financeiro, mecanismos institucionais de controle e transparência capazes de assegurar a sustentabilidade das contas públicas. Nesse contexto, compreender os fatores que influenciam o endividamento estatal é essencial para analisar os desafios da administração pública e seus impactos sobre a economia.

Em síntese, a dívida pública é, do ponto de vista jurídico e econômico, o conjunto de compromissos financeiros que o estado assume com outras pessoas ou instituições para conseguir recursos e manter suas atividades quando suas receitas habituais não bastam. Para entender como essa dívida deve ser classificada legalmente, é necessário consultar a Lei nº 4.320/1964, que define as regras gerais do Direito Financeiro e organiza o passivo público (Amorim *et al*., 2026).

A gestão da dívida pública constitui um tema complexo, envolvendo dimensões macroeconômicas, jurídicas e fiscais. O endividamento público corresponde às obrigações financeiras assumidas pelo estado por meio de operações de crédito, como empréstimos e financiamentos, com compromisso de pagamento futuro do principal acrescido de juros (Ferreira, 2019). Assim, a dívida pública representa um importante instrumento de financiamento governamental, cuja utilização requer planejamento e disciplina fiscal para evitar desequilíbrios nas contas públicas.

A crise financeira internacional de 2007–2008 intensificou o debate sobre o endividamento público e impôs novos desafios às políticas econômicas. Entre seus efeitos, destacou-se a migração de capitais das economias centrais para as periféricas, motivada pelas baixas taxas de juros, pelas reduzidas perspectivas de crescimento e pelo aumento da dívida nas economias mais afetadas (Cunha; Garcia, 2012). Esse movimento alterou a dinâmica financeira internacional e ampliou a vulnerabilidade fiscal de diversos países emergentes.

No Brasil, esse cenário contribuiu para a ampliação da dívida pública e reforçou a necessidade de mecanismos voltados ao controle fiscal, destacando-se a Lei de Responsabilidade Fiscal (LRF), que estabeleceu regras para disciplinar o endividamento dos entes federativos (Tesouro Nacional, 2025). Dessa forma, a gestão responsável da dívida tornou-se condição indispensável para preservar a sustentabilidade das finanças públicas, manter a capacidade de investimento estatal e promover um desenvolvimento econômico equilibrado.

Nesse contexto, a comunicação institucional da Secretaria de Estado de Fazenda de Minas Gerais (SEFMG) assume papel estratégico ao divulgar informações relacionadas à dívida pública. A recorrência desse tema decorre de exigências legais, necessidades de gestão fiscal e arrecadatória e do compromisso com a transparência e a prestação de contas à sociedade.

2.3 EQUILÍBRIO FISCAL

O Planejamento Orçamentário Anual – PLOA – é o instrumento de planejamento que detalha a previsão dos recursos a serem arrecadados, juntamente com sua destinação, de forma anual. Assim, ao analisar receitas e despesas, o PLOA busca manter o equilíbrio das contas públicas e atender às prioridades do governo.

A Lei Orçamentária Anual (LOA) é uma ferramenta de planejamento e programação que deve estar alinhada ao Plano Plurianual (PPA) e à Lei de Diretrizes Orçamentárias (LDO) (Brasil, 2025). Ela é obrigatória e deve seguir as normas constitucionais, como a Lei de Responsabilidade Fiscal. Isso mostra que o orçamento é uma expressão prática do planejamento governamental, traduzindo as ações previstas em lei para o período de um ano.

Segundo o Ministério do Planejamento e Orçamento (MPO) o orçamento público é o planejamento que mostra, ano a ano, quanto dinheiro o governo espera arrecadar com impostos e outras receitas, e em que vai gastar esse dinheiro. Ao reunir receitas e despesas, o orçamento ajuda a manter as contas públicas em equilíbrio e mostra quais são as prioridades do governo para a sociedade. (Brasil, 2024).

As transformações sociais impõem maiores flexibilidades orçamentárias. Elas demostram com profundidade as evoluções, heterogêneas, que a sociedade vêm passando no contexto histórico. Dessa forma, as questões públicas orçamentárias devem ser um processo dinâmico e adaptável. Neuhold, Pozzer e Pellejero (2025) descrevem que, o planejamento se torna extremamente relevante em ambientes que se caracterizam pelas constantes mudanças e reivindicação sociais.

As mudanças sociais são complexas, o que aumenta a importância do planejamento e da previsão em nível governamental. Isso mostra que, diante de uma sociedade em constante transformação, o planejamento se torna essencial para orientar ações e políticas públicas Comparato (1995). Por isso, o planejamento orçamentário precisa ser dinâmico e capaz de se ajustar a esse contexto, o que aumenta a importância da previsão e da revisão de estimativas em nível governamental (Cardoso Jr., 2014).

2.4 TEORIA DA CONTINGÊNCIA

A Teoria da Contingência representa uma das abordagens mais relevantes para a compreensão das organizações diante das constantes mudanças do ambiente econômico, tecnológico e institucional.

Sua principal contribuição está na ideia de que não existe um único modelo de gestão capaz de atender todas as organizações de maneira universal, uma vez que as estruturas administrativas precisam adaptar-se às circunstâncias e às condições externas em que estão inseridas. Dessa forma, a teoria passou a oferecer uma visão mais dinâmica e flexível da administração, considerando a influência do ambiente sobre o desempenho organizacional (Sobrinho e Callado, 2025).

Os estudos desenvolvidos por pesquisadores como Joan Woodward, Tom Burns, G. M. Stalker, Paul Lawrence e Jay Lorsch, entre as décadas de 1950 e 1960, foram fundamentais para o desenvolvimento da Teoria da Contingência. Esses autores demonstraram que a performance econômica das organizações era fortemente influenciada por fatores como a estrutura organizacional, o nível tecnológico empregado e as condições do ambiente externo (Fagundes, Petri, Lavarda, Rodrigues, Lavarda e Soller, 2011). Assim, passou-se a compreender que diferentes contextos exigem diferentes formas de organização e gestão.

No âmbito da Administração Pública, a aplicação da Teoria da Contingência ganha relevância diante das incertezas provocadas pelas transformações econômicas, políticas e sociais. A abordagem contingencial defende que não existem princípios administrativos absolutos, pois tudo depende das circunstâncias enfrentadas pelas organizações (Chiavenato, 1979).

Nesse sentido, a teoria estabelece uma relação funcional entre as condições ambientais e as técnicas administrativas mais adequadas para o alcance dos objetivos institucionais. As variáveis do ambiente são compreendidas como fatores independentes, enquanto as práticas administrativas assumem o papel de variáveis dependentes, ajustando-se às necessidades impostas pelo contexto organizacional (Ferreira, 2002).

2.5 COMUNICAÇÃO INSTITUCIONAL COMO OBJETO DE MENSURAÇÃO

A saliência fiscal refere-se à forma como os tributos são percebidos pelos contribuintes, influenciando sua legitimidade e produzindo efeitos econômicos e políticos. Maior visibilidade tende a ampliar a percepção de risco e volatilidade, enquanto menor saliência pode elevar a incerteza sobre os impactos fiscais (Shabbir, 2023).

A comunicação institucional constitui um instrumento estratégico de transparência, prestação de contas e divulgação das prioridades governamentais, refletindo as escolhas relacionadas às políticas públicas e aos objetivos da administração (Miola; Marques, 2024). Com a expansão das tecnologias digitais, os meios eletrônicos ampliaram o acesso às informações públicas e consolidaram novos canais de interação entre Estado e sociedade (De Holanda, 2026).

Além de fortalecer a imagem organizacional, a comunicação estratégica integra públicos, evidencia resultados e amplia a credibilidade das organizações públicas (Capriotti; Losada-Díaz *et al*., 2023; Capriotti; Oliveira *et al*., 2023; Capriotti; Zeler, 2023; Salas-Canales, 2021; Spiazzi; Pozobon, 2024). Os

temas, a frequência e a linguagem das comunicações oficiais revelam prioridades institucionais e decisões administrativas, constituindo uma importante fonte para compreender a atuação governamental (Viola *et al*., 2021; Miola; Marques, 2024).

A recorrência de determinados assuntos ao longo do tempo permite identificar mudanças de foco e padrões de atuação institucional. Nesse sentido, documentos oficiais podem ser utilizados como bases empíricas para construir indicadores quantitativos, convertendo informações textuais em séries temporais capazes de representar a intensidade da atenção institucional dedicada a temas específicos (Silva; Hayashi; Hayashi, 2011).

Essa abordagem aproxima a comunicação institucional dos estudos sobre indicadores sintéticos, que permitem agregar diferentes dimensões em uma única medida analítica, favorecendo o monitoramento de processos complexos, a avaliação de políticas públicas e a comparação entre diferentes períodos e contextos (Corrêa *et al*., 2014; Schumann; Moura, 2015). Assim, a frequência temática das comunicações oficiais passa a representar evidências quantitativas das prioridades do Estado, fundamentando a utilização do corpus da Secretaria de Estado de Fazenda de Minas Gerais na construção do Indicador de Saliência Fiscal.

As teorias da incerteza econômica explicam como alterações no ambiente macroeconômico influenciam as decisões fiscais, enquanto a Teoria da Contingência sustenta que as organizações públicas adaptam suas prioridades às mudanças do ambiente. Complementarmente, a comunicação institucional evidencia essas prioridades, e a mineração textual fornece os procedimentos para transformá-las em indicadores quantitativos baseados em grandes volumes de documentos (Tozzi, 2024; Rocha, 2026).

Dessa forma, o Indicador de Saliência Fiscal representa mais do que uma medida da frequência de termos fiscais. Fundamentado teoricamente, o indicador traduz as mudanças na saliência da agenda fiscal da Secretaria de Estado de Fazenda de Minas Gerais diante das contingências econômicas e institucionais, permitindo sua aplicação em análises econométricas e na avaliação da dinâmica das finanças públicas estaduais.

3 METODOLOGIA

Este capitulo descreve, em detalhe, o procedimento utilizado para construir a base textual trimestral empregada nas análises exploratórias da comunicação fiscal da secretaria de estado de fazenda de minas gerias (SEF/MG). O objetivo metodológico foi transformar notícias oficiais publicadas no portal da SEF/MG em indicadores temporais de comunicação institucional, com periodicidade trimestral, preservando uma distinção fundamental entre volume de comunicação e composição fiscal do discurso. O produto final desta etapa é o arquivo 07_base_modelavel.csv, que contém variáveis textuais, indicadores de qualidade da informação, dummies de regime e defasagens necessárias para modelagem posteriores.

A construção da base foi organizada em sete etapas: i) definição do universo documental e auditoria do *corpus* original; ii) coleta de links oficiais no portal da SEF/MG; iii) extração e limpeza dos textos das notícias; iv) agregação temporal e comparação com o *corpus* original; v) construção dos indicadores textuais por dicionário; vi) montagem da base trimestral; e vii) construção da base modelável. As subseções seguintes descrevem cada etapa.

3.1 PROCEDIMENTOS METODOLÓGICOS

3.1.1 Definição do universo documental

O universo documental de interesse é composto por notícias publicadas no portal oficial da SEF/MG no período de 2007 a 2022. Cada notícia é tratada como uma unidade documental individual. Seja:

$$D = \{d_i\}_{i=1}^{N}; \quad (1)$$

O conjunto de documentos oficiais extraídos do portal, em que d_i representa a notícia e N é o número total de notícias válidas no período analisado. Após a etapa de coleta, extração textual e filtragem temporal, obteve-se:

$$N = 562 \quad (2)$$

Notícias oficiais com o texto extraído e data atribuída dentro do intervalo 2007-2022. Essas notícias são a base documental principal do estudo.

O *corpus* originalmente utilizado no artigo de referência, construído em formato compatível com o IRaMuTeQ, foi documental. Denota-se esse *corpus* por:

$$J = \{j_m\}_{n=1}^{M}, \qquad M = 236 \quad (3)$$

A comparação entre $\mathcal{D}$ e $\mathcal{J}$ indicou que o *corpus* original continha 236 textos, enquanto a base oficial reconstruída continha 562 notícias. A cobertura aproximada do corpus original em relação à base oficial foi, portanto:

$$Cobertura_j = \frac{M}{N} = \frac{236}{562} \approx 0{,}420 \quad (4)$$

Assim, o *corpus* original cobre aproximadamente 42% das notícias oficiais reconstruídas. Essa evidência motivou a decisão de utilizar o *corpus* original apenas como referência documental e comparativa, preservando a base temporal principal.

3.1.2 Coleta dos links oficiais

A coleta foi realizada diretamente no portal oficial da SEF/MG, em páginas de notícias organizadas por ano e paginação. Para cada ano a no conjunto:

$$\mathcal{A} = \{2007,2008, \dots ,2022\}, \quad (5)$$

Foram ativados os filtros anuais disponíveis no portal e coletados os links associados a notícias. A coleta foi feita com navegação automatizada, salvando também cópias HTML das páginas visitadas, de modo a permitir auditoria posterior dos links encontrados.

Seja La o conjunto bruto de links extraídos para o ano a. Cada elemento de La contém, quando disponível, as seguintes informações: texto da âncora, título, endereço absoluto, endereço bruto, página de origem, URL do índice anual e possíveis datas identificadas no próprio link. Formalmente:

$$L_a = \{l_{a,r}\}_{r=1,}^{R_a}\} \quad (6)$$

Em que R é o número de links coletados para o ano a antes dos procedimentos de deduplicação e validação.

A deduplicação foi feita principalmente pelo endereço da notícia. Seja $h(\ell)$ o endereço padronizado do link ℓ . o conjunto deduplicado por ano é definido por:

$$L_a^* = \{l \in L_a : h(l) \quad (7)$$

A união dos links deduplicados em todos os anos fornece o conjunto preliminar de notícias candidatas:

$$L^* = \bigcup_{a \in \mathcal{A}} L_a^* \quad (8)$$

Na etapa de reconstrução dos links, todos os arquivos HTML salvos foram processados para garantir que a base final de links não dependesse apenas dos objetos gerados em memória durante a navegação. Essa etapa produziu o arquivo 02_links_oficiais.csv, utilizado como entrada para a extração dos textos.

3.1.3 Extração textual das notícias

Cada link válido foi acessado individualmente para extração do conteúdo textual, título e data da notícia, a função de extração textual pode ser representada por:

$$\mathrm{E}(u_i) = (\tau_i, T_i, x_i, \delta_i) \qquad (9)$$

Em que τ é o título extraído, T é o texto integral extraído, x é a data atribuída a notícia e δ é um indicador de sucesso da extração.

A data final da notícia foi obtida por regra hierárquica de precedência. Foram consideradas diferentes fontes de data: data identificada na página, data presente no texto, data inferida da URL e data preliminar extraída na etapa de coleta de links. A data final é definida por:

$$x_i = \mathrm{coalesce}(x_i^{página}, x_i^{texto}, x_i^{url}, x_i^{preliminar}) \qquad (10)$$

Em que coalesce retorna o primeiro valor não ausente segundo a ordem de prioridade especificada. A prioridade atribuída à data da página e à data textual decorre do fato de que alumas URLs podem conter pastas ou datas inconsistentes com a data efetiva de publicação.

Após a extração foram mantidas apenas notícias com data final dentro do intervalo de análise:

$$2007 \leq \mathrm{ano}(x_i,) \leq 2022. \qquad (11)$$

O texto de cada notícia foi preservado em forma integral e também convertido para uma forma normalizada para fins busca lexical. Essa normalização teve o objetivo de reduzir variações superficiais, tais como diferenças entre maiúscula e minúsculas e acentuação, permitindo a comparação consistente com os termos do dicionário.

Seja T o texto original da notícia$_i$ e $\tilde{T}$ sua versão normalizada. A transformação pode ser representada genericamente por:

$$\sim T_i = \mathcal{N}(T_i,) \qquad (12)$$

Em que $\mathcal{N}(\cdot)$ representa a função de normalização textual. O número de palavras da notícia$i$ é denotado por:

$$W_i = \#\ \{\text{tokens em } \sim T_i\} \qquad (13)$$

A extração final resultou em 562 notícias válidas e 247.979 palavras no total:

$$\sum_{i=1}^{562} W_i = 247.979 \qquad (14)$$

3.1.4 Indexação temporal trimestral

Cada notícia foi associada a um trimestre com base em sua data final x. Seja q(i) a função que associa a notícia i ao trimestre correspondente. O conjunto de trimestre analisados é:

$$Q = \{2007T1,2007T2,...,2022T4\} \qquad (15)$$

Como o período compreende 16 anos e cada ano possui 4 trimestres, o número total de trimestres é:

$$|Q| = 16 \times 4 = 64. \qquad (16)$$

Para cada trimestre $q \in \mathcal{Q}$, define-se o conjunto de notícias publicadas naquele trimestre como:

$$\mathcal{D}_2 = \{d_i \in \mathcal{D}: q(i) = q\} \qquad (17)$$

O número de notícias no trimestre q é dado por:

$$\boldsymbol{N}_q = |\mathcal{D}_q| \qquad (18)$$

O volume textual trimestral é definido como a soma do número de palavra de todas as notícias publicadas no trimestre:

$$\boldsymbol{W}_q = \sum_{i:d_i \in \mathcal{D}_q} W_i \qquad (19)$$

Quando não há notícias em determinado trimestre, tem-se $\boldsymbol{N}_q = 0$ e $\boldsymbol{W}_q = 0$. Na base oficial reconstruída, forma identificadas 51 trimestres com comunicação e 13 trimestre sem comunicação:

$$\sum_{q\in Q} 1(\boldsymbol{N}_q > 0) = 51\,, \quad \sum_{q\in Q} 1(\boldsymbol{N}_q = 0) = 13 \qquad (20)$$

3.1.5 Auditoria do *corpus* original e identificação de zeros artificiais

O *corpus* original em formato IRaMuTeQ foi parseado, por meio dos marcadores textuais de notícia. Essa etapa permitiu reconstruir a distribuição temporal do *corpus* original e compará-la com a distribuição obtida no portal oficial.

Seja N_q^j o número de notícias do *corpus* original atribuídas ao trimestre q, e N_q^D o número de notícias da base oficial reconstruída no mesmo trimestre. A diferença absoluta entre as duas bases é:

$$\Delta_q = N_q^D - N_q^j \qquad (21)$$

A cobertura trimestral do *corpus* original pode ser definida como:

$$\boldsymbol{k}_q = \{\frac{N_q^j}{N_q^D}\text{, se } N_q^D > 0\ \}$$
$$\boldsymbol{k}_q = \{\text{NA, se } N_q^D = 0\ \} \qquad (22)$$

Distinguem-se dois fenômenos: i) zeros artificiais do *corpus* original: trimestre em que o *corpus* original (M = 236 textos) não continha nenhuma notícia, embora o portal registrasse publicações. Identificaram-se [n] ocorrências, corrigidas pela reconstrução.

$$Z_q^{art} = 1\ (N_q^j = 0 \cap N_q^D > 0) \qquad (23)$$

ii) Trimestres sem comunicação na base reconstruída: trimestres em que o portal não registra nenhuma notícia ($N_q = 0$), em número de 13. Estes podem refletir ausência efetiva de publicação ou limitações de preservação do acervo e, por isso, não são tratados como saliência fiscal nula.

O número total de zeros artificiais é:

$$Z^{art} = \sum_{q\in Q} Z_q^{art} = 13 \qquad (24)$$

Esse resultado foi metodologicamente central. Ele demostra que a ausência de notícias no corpus original nem sempre correspondia à ausência de notícias no portal oficial. Por essa razão, o corpus original não foi utilizado como série temporal principal. Sua função foi restrita à comparação e à validação qualitativa do processo de reconstrução documental.

3.1.6 Construção do dicionário fiscal

A mensuração da composição do discurso foi feita por meio de dicionário lexical. O dicionário foi organizado em dimensões temáticas associadas à situação fiscal, endividamento, renegociação, gestão financeira e governança fiscal.

Seja k o número de dimensões lexicais do dicionário. Para cada dimensão $k \in \{1, \dots, K\}$, define-se um conjunto de termos ou expressões:

$$v_k = \{v_{k1}, v_{k2},..., v_{kj_k}\} \qquad (25)$$

Em que j_k é o número de termos pertencentes a dimensão k.

A contagem de termos da dimensão k na notícia i é representada por:

$$C_{ik} = \sum_{j=1}^{jk} C_{ij} \qquad (26)$$

Em que C_{ij} é a frequência do termo V_{kj} na versão normalizada da notícia i.

Quando o termo é uma palavra simples, a contagem pode ser representada como:

$$C_{ij} = \sum_{r=1}^{W_i} 1(W_{ir} = v_{kj}) \qquad (27)$$

Em que W_{ir} é o token r da notícia i. Quando o item lexical é uma expressão composta, a ocorrência é identificada na sequência normalizada de texto.

O score lexical da dimensão k na noticiai é definido como a razão entre o número de ocorrências dos termos da dimensão e o total de palavras da notícia:

$$S_{ik} = \frac{C_{ik}}{W_i}, \qquad W_i > 0 \qquad (28)$$

Essa normalização por número de palavras é essencial, pois evita que notícias mais longas tenham scores maiores apenas por possuírem maior extensão textual.

3.1.7 Agregação dos scores por trimestre

Os indicadores trimestrais foram construídos agregando as contagens lexicais antes do cálculo das proporções. Para cada dimensão k e trimestre q, a contagem agregada é:

$$C_{qk} = \sum_{i:d_i \in \mathcal{D}_q} C_{ik} \qquad (29)$$

O score trimestral da dimensão k é então:

$$\boldsymbol{S}_{qk} = \{\frac{C_{qk}}{W_q}, \text{ se } W_q > 0 \}$$
$$\boldsymbol{S}_{qk} = \{0, \text{ se } W_q = 0 \} \qquad (30)$$

A escolha de agregar contagem antes de dividir pelo total de palavras faz com que o score trimestral seja ponderado pelo volume textual de cada notícia. Assim, uma notícia muita curta não recebe o mesmo peso de uma notícia substancialmente mais longa.

O ISF do trimestre q é a participação dos termos do dicionário fiscal no total de palavras publicadas:

$$ISF_q = \frac{C_q}{W_q}$$

Em que $C_q = \sum_k \sum_j C_{q,kj}$ é o número de ocorrências dos termos do dicionário (união das K dimensões, sem dupla contagem de termos presentes em mais de uma dimensão) e W_q é o total de palavras das notícias do trimestre. Os subíndices por dimensão, $ISF_{q,K}$, são calculados da mesma forma e usados em análises complementares.

Além das dimensões específicas, foi construído um score fiscal agregado restrito. Seja R o conjunto de termos pertencentes ao dicionário fiscal restrito. A contagem restrita da notícia i é:

$$C_i^R = \sum_{v \in \mathrm{R}} freq(\mathrm{v}, \sim T_i) \qquad (31)$$

No trimestre q, a contagem fiscal restrita agregada é:

$$C_q^R = \sum_{i:d_i \in \mathcal{D}_q} C_i^R \qquad (32)$$

O score fiscal total restrito trimestral é definido como:

$$S_q^R = \{\frac{C_q^R}{W_q}, \text{ se } W_q > 0 \}$$
$$S_q^R = \{0, \text{ se } W_q = 0\} \quad (33)$$

Esse indicador é a principal medida de saliência fiscal do discurso. Ele mede a participação relativa do discurso. Ele mede a participação relativa dos termos fiscais restritos no conjunto de palavras publicadas pelo SEF/MG no trimestre.

A etapa da construção dos índices textuais gerou 64 observações trimestrais, 562 notícias e 247.979 palavras. O score fiscal restrito médio foi aproximadamente 0,0293, com valor máximo de aproximadame 0,0849.

3.1.8 Construção da base temporal trimestral

A base temporal trimestral foi organizada com uma linha para cada trimestre q ∈ Q. As variáveis básicas são:

$$(q, N_q, W_q, S_q^R, S_{q1},..., S_{qk}) \quad (34)$$

A média de notícias por trimestre na base final foi.

$$\tilde{N} = \frac{1}{64} \sum_{q \in Q} N_q \approx 8{,}78 \quad (35)$$

A média de palavras por trimestre foi aproximadamente:

$$\sim W = \frac{1}{64} \sum_{q \in Q} W_q \approx 3.875 \quad (36)$$

Essas médias incluem os trimestres sem comunicação, nos quais $\boldsymbol{N_q} = 0$ e $\boldsymbol{W_q} = 0$.

3.1.9 Problemas dos zeros e distinção entre ausência de comunicação e ausência de termos fiscais

Um ponto metodológico importante é que o valor zero do score fiscal bruto pode ter duas interpretações distintas. Em um trimestre q, pode ocorrer.

$$S_q^R = 0 \quad (37)$$

Por pelo menos dois motivos:

1) Não houve nenhuma notícia no trimestre, isto é, $W_q = 0$;
2) Houve notícias, mas não foram identificados termos do dicionário fiscal restrito, isto é, $W_q > 0$ e $C_q^R = 0$.

Essas duas situações não são substantivamente equivalentes. A primeira representa ausência de comunicação; a segunda representa comunicação sem saliência fiscal mensurada pelo dicionário. Por essa razão, a base modelável separa o score bruto de versões condicionais do score.

O score fiscal bruto é definido como:

$$S_{q,bruto}^R = S_q^R \qquad (38)$$

O score fiscal condicional à existência de comunicação é definido como:

$$S_{q,cond}^R = \{S_q^R, \text{ se } W_q > 0 \}$$
$$S_{q,cond}^R = \{\text{NA, se } W_q = 0\} \quad (39)$$

Essa transformação impede que trimestres sem qualquer texto sejam tratados como trimestres nos quais a SEF/MG efetivamente comunicou, mas sem conteúdo fiscal.

Também foi criada uma versão mais conservadora do score, condicionada a um volume textual mínimo de 500 palavras:

$$S_{q,500}^R = \{S_q^R, \text{ se } W_q \geq 0 \}$$
$$S_{q,500}^R = \{\text{NA, se } W_q < 0\} \qquad (40)$$

Essa variável reduz a influência de trimestres com comunicação textual muito pequena, nos quais poucos termos podem alterar fortemente o score por causa do denominador reduzido.

3.1.10 Transformações de volume comunicacional

Além dos scores fiscais, foram construídos transformações logarítmicas do volume de comunicação. O logaritmo do número de notícias é:

$$l_{N,q} = \log (N_q + 1) \qquad (41)$$

O acréscimo de uma unidade evita indefinição em trimestres com N_q= 0. De modo análogo, o logaritmo do volume de palavras é:

$$l_{W,q} = \log (W_q + 1) \qquad (42)$$

Essas variáveis capturam a margem extensiva da comunicação institucional. Enquanto S_q^R mede a saliência fiscal relativa do discurso, $l_{N,q}$ e $l_{W,q}$ medem a intensidade de publicação e o volume textual.

A distinção é relevante porque um aumento da presença fiscal no discurso pode ocorrer por diferentes mecanismos. Pode haver aumento do volume total de comunicação, aumento da proporção de termos fiscais ou ambos. Algebricamente, o número total de termos fiscais restritos no trimestre pode ser decomposto como:

$$C_q^R = S_q^R \times W_q \qquad (43)$$

Assim, variações em C_q^R podem decorrer tanto de variações em S_q^R quanto de variações em W_q:

$$\Delta C_q^R \approx W_q\, \Delta S_q^R + S_q^R\, \Delta W_q \qquad (44)$$

Essa decomposição mostra por que é importante distinguir volume textual e intensidade fiscal do discurso.

3.1.11 Indicadores binários de presença fiscal e baixa informação

Foi construída uma variável indicadora para a presença de termos fiscais restritos no trimestre:

$$F_q = 1\ (C_q^R > 0) \qquad (45)$$

A versão condicional à existência de comunicação é:

$$F_{q,cond} = \{F_q, \text{ se } W_q > 0\}$$
$$F_{q,cond} = \{\text{NA, se } W_q = 0\} \qquad (46)$$

Também foi criado um indicador de baixa informação textual:

$$B_q = 1(N_q > 0 \cap W_q < 500 \qquad (47)$$

Na base final, foram identificadas 7 trimestres com baixa informação textual. Esses casos não foram necessariamente excluídos da base bruta, mas foram marcados para permitir diagnósticos e especificações alternativas.

3.1.12 Dummy de quebra informacional em 2017

A inspeção da série de notícias e palavras revelou mudanças expressivas no padrão de disponibilidade ou volume textual a partir de 2017. Essa alteração pode refletir mudança de organização do portal, política de comunicação, preservação do acervo digital ou mudança institucional mais ampla. Para modelar essa quebra, foi criada a dummy:

$$D_{2017,q} = 1(\text{ano(q)} \geq 2017 \qquad (48)$$

Essa variável não deve ser interpretada automaticamente como mudança econômica ou fiscal. Sua função principal é controlar uma quebra observada na série textual, separando, na medida do possível, variações de comunicação associadas ao conteúdo substantivo de variações associadas à disponibilidade ou organização do acervo.

4 RESULTADOS E DISCUSSÃO

4.1 DICIONÁRIO LEXICAL FISCAL

Para mensurar a saliência fiscal do discurso institucional, foi elaborado um dicionário lexical organizado em cinco dimensões temáticas, contemplando aspectos relacionados ao endividamento público, gestão financeira, renegociação de passivos e governança fiscal, conforme apresentado no Quadro 3. A estrutura do dicionário permitiu construir indicadores textuais representativos da comunicação institucional da SEF/MG, posteriormente agregados em frequência trimestral para utilização nas análises qualitativas e econométricas.

Quadro 1- Dimensões do dicionário lexical utilizado na construção do indicador textual

Dimensão	Objetivo analítico	Principais termos
Situação fiscal e endividamento	Identificar referências à condição financeira do Estado e ao comportamento do endividamento público.	dívida, déficit, caixa, despesa, orçamento, equilíbrio, fiscal, endividamento
Pagamento e renegociação	Captar elementos relacionados à administração dos passivos financeiros e às estratégias de regularização das obrigações públicas.	pagamento, acordo, parcela, prazo, desconto, multa, renegociação, quitação, vencimento
Gestão administrativa e financeira	Identificar o gerenciamento das receitas públicas e dos principais instrumentos de execução financeira do Estado.	receita, arrecadação, recursos, folha, investimento, ICMS, IPVA
Governança fiscal ampla	Mensurar referências às instituições responsáveis pela formulação, coordenação e execução das políticas fiscais estaduais.	governo, estado, União, Assembleia, Secretaria, Fazenda, regime, RRF, Confaz
Governança fiscal restrita	Mensurar especificamente o discurso relacionado à recuperação fiscal e aos mecanismos institucionais de ajuste das contas públicas.	Assembleia, recuperação fiscal, adesão, Confaz, Conselho de Contribuintes

Fonte: Elaborado pelo autor (2026)

A utilização do dicionário lexical estruturado em cinco dimensões amplia a capacidade analítica do indicador ao incorporar diferentes aspectos da gestão fiscal. Em vez de mensurar apenas referências ao endividamento, o ISF contempla elementos relacionados à arrecadação, pagamento, governança institucional e mecanismos de recuperação fiscal, proporcionando uma representação mais abrangente da comunicação oficial.

A construção do indicador também reduz a subjetividade inerente às análises qualitativas da comunicação pública. Como os termos foram selecionados a partir do referencial teórico e organizados em dimensões conceituais, a mensuração passa a ser realizada por critérios replicáveis e transparentes, fortalecendo a validade de conteúdo da medida desenvolvida (Silva; Hayashi; Hayashi, 2011).

A estrutura do indicador permite sua utilização em análises longitudinais e em estudos comparativos sobre comunicação institucional. Como a metodologia é baseada em mineração textual e frequência relativa dos termos, sua aplicação pode ser reproduzida em outras secretarias estaduais, órgãos federais ou instituições públicas que disponibilizem séries históricas de documentos oficiais.

A validade empírica do ISF é reforçada pelos resultados da Classificação Hierárquica Descendente e da análise de similitude. A elevada correspondência entre as classes identificadas pelo IRaMuTeQ, a estrutura das redes semânticas e as dimensões propostas no dicionário lexical demonstra que o indicador representa adequadamente os principais eixos temáticos da comunicação fiscal da SEF/MG.

4.2 DISCUSSÃO DO DICIONÁRIO LEXICAL FISCAL

A estrutura do dicionário lexical fiscal demonstra elevada coerência com o referencial teórico ao representar, em dimensões distintas, os principais eixos da gestão das finanças públicas. A organização dos termos aproxima a mineração textual dos fundamentos da administração pública, permitindo que o discurso

institucional da Secretaria de Estado de Fazenda seja interpretado como evidência das prioridades governamentais (Miola; Marques, 2024; Corrêa *et al*., 2014).

A dimensão Situação Fiscal e Endividamento sintetiza o núcleo conceitual da gestão fiscal ao reunir termos como dívida, déficit, orçamento, despesa e equilíbrio. Esses vocábulos refletem os instrumentos utilizados pelo estado para preservar a sustentabilidade financeira, em consonância com a literatura que atribui ao controle do endividamento papel central na manutenção da capacidade de investimento e da estabilidade das contas públicas (Ferreira, 2019).

Além de representar indicadores financeiros, essa dimensão expressa a preocupação institucional com a responsabilidade fiscal. A recorrência desses termos na comunicação oficial demonstra que a gestão da dívida ultrapassa sua dimensão contábil e passa a constituir um elemento estratégico da administração pública, associado à transparência, ao planejamento e à prestação de contas exigidos pela Lei de Responsabilidade Fiscal (Brasil, 2025).

A dimensão Pagamento e Renegociação amplia essa interpretação ao incorporar termos relacionados à administração dos passivos públicos. Palavras como acordo, parcela, renegociação, quitação e vencimento representam mecanismos utilizados para preservar a liquidez financeira e reorganizar obrigações fiscais diante de restrições orçamentárias, evidenciando respostas institucionais às mudanças do ambiente econômico (Ferreira, 2019).

Sob a ótica da Teoria da Contingência, essa dimensão evidencia que as estratégias administrativas não permanecem estáticas ao longo do tempo. A intensificação de vocábulos relacionados à renegociação tende a refletir momentos em que o ambiente econômico exige adaptações institucionais, confirmando que a gestão pública ajusta seus instrumentos conforme as condições impostas pelo contexto externo (Chiavenato, 1979; Ferreira, 2002).

A dimensão Gestão Administrativa e Financeira representa o funcionamento cotidiano da Secretaria da Fazenda ao concentrar termos associados à arrecadação, receitas, investimentos e tributos estaduais. Sua composição está diretamente alinhada às atribuições legais da SEF/MG, demonstrando que a comunicação institucional também desempenha função informativa sobre a execução das atividades fazendárias e tributárias (SEF, 2026).

Essa dimensão também reforça a relação entre planejamento orçamentário e gestão financeira discutida no referencial teórico. A arrecadação constitui a principal fonte de financiamento das políticas públicas e, portanto, sua recorrência no discurso institucional revela a preocupação permanente com a manutenção do equilíbrio entre receitas e despesas previsto pelos instrumentos de planejamento governamental, como PPA, LDO e LOA (Brasil, 2024; Brasil, 2025).

A dimensão Governança Fiscal Ampla desloca a análise do aspecto operacional para o ambiente institucional onde as decisões fiscais são formuladas. A presença de termos como ‘Governo’, ‘Estado’,

‘União’, ‘Assembleia’, ‘Secretaria’ e ‘Confaz’ demonstra que a política fiscal resulta da interação entre diferentes órgãos e níveis de governo, evidenciando a natureza institucional da gestão das finanças públicas.

Essa dimensão dialoga diretamente com a literatura sobre comunicação institucional, segundo a qual os conteúdos divulgados pelos órgãos públicos representam manifestações das prioridades organizacionais. Assim, ao enfatizar instituições responsáveis pela formulação e coordenação das políticas fiscais, a comunicação oficial revela como o estado estrutura seus processos decisórios e fortalece sua legitimidade perante a sociedade (Miola; Marques, 2024).

A dimensão Governança Fiscal Restrita apresenta elevada especificidade analítica ao concentrar termos relacionados ao Regime de Recuperação Fiscal e aos mecanismos formais de ajuste das contas públicas. Diferentemente da dimensão anterior, seu foco recai sobre instrumentos institucionais voltados à reorganização financeira do estado, permitindo identificar períodos em que o discurso institucional privilegia estratégias de recuperação fiscal.

Do ponto de vista da Teoria da Contingência, essa dimensão representa uma manifestação concreta da adaptação organizacional. A intensificação da comunicação relacionada ao Regime de Recuperação Fiscal tende a ocorrer em cenários de maior restrição financeira, indicando que a instituição ajusta sua agenda comunicacional às demandas impostas pelo ambiente econômico e pelas exigências legais (Chiavenato, 1979; Fagundes *et al*., 2011).

Em conjunto, as cinco dimensões apresentam complementaridade conceitual e metodológica. Enquanto as três primeiras representam aspectos operacionais da gestão fiscal, as duas últimas capturam a dimensão institucional da governança pública. Essa organização amplia a capacidade explicativa do indicador de saliência fiscal ao contemplar simultaneamente processos financeiros, administrativos e decisórios presentes na comunicação institucional da SEF/MG.

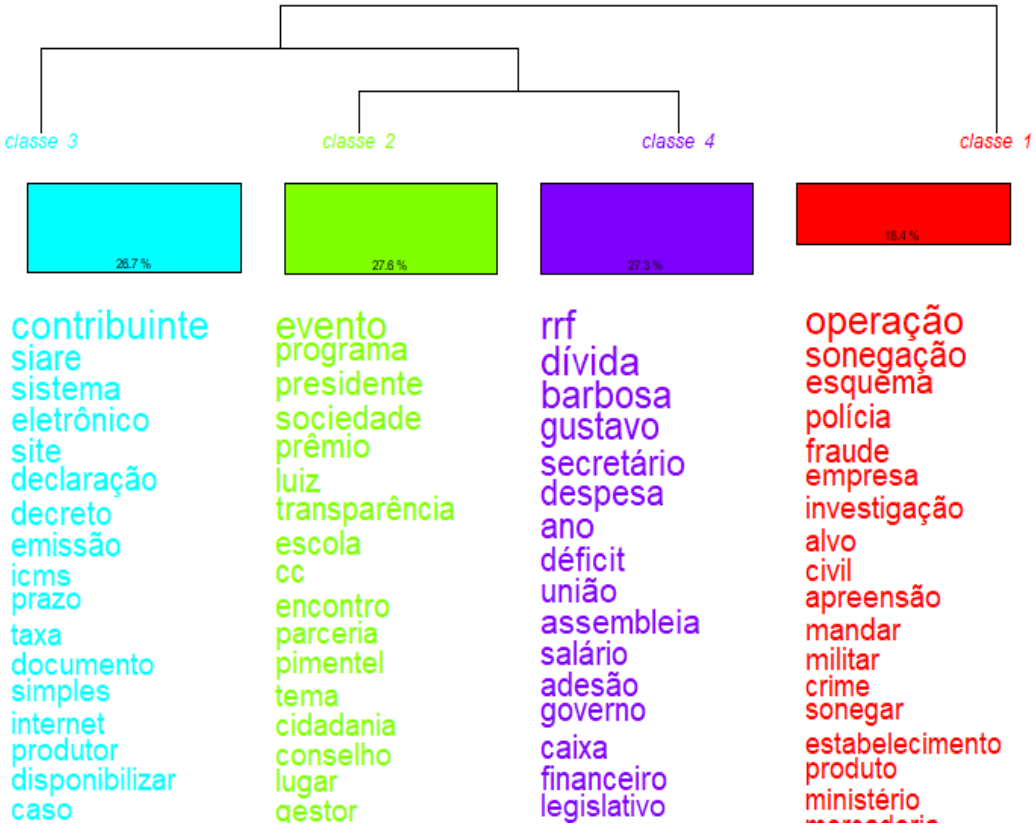
Sob a perspectiva metodológica, a segmentação do vocabulário em dimensões reduz a subjetividade inerente às análises qualitativas de textos institucionais. Em vez de interpretar palavras isoladamente, o estudo estrutura categorias fundamentadas na literatura sobre dívida pública, equilíbrio fiscal, comunicação institucional e Teoria da Contingência, fortalecendo a validade de conteúdo do indicador desenvolvido.

Por fim, o dicionário lexical fiscal não constitui apenas um conjunto de palavras-chave, mas um modelo analítico capaz de operacionalizar conceitos teóricos em medidas quantitativas. Essa característica aproxima a comunicação institucional da produção de indicadores sintéticos, permitindo transformar prioridades discursivas em evidências empíricas sobre a dinâmica da gestão fiscal estadual ao longo do período analisado (Corrêa *et al*., 2014; Silva; Hayashi; Hayashi, 2011).

4.3 CLASSIFICAÇÃO HIERÁRQUICA DESCENDENTE

A Classificação Hierárquica Descendente (CHD) foi utilizada para identificar os principais agrupamentos lexicais presentes na comunicação institucional da Secretaria de Estado de Fazenda de Minas Gerais. A técnica permite segmentar o corpus em classes semanticamente homogêneas, evidenciando os temas predominantes e possibilitando interpretar como o discurso institucional se organiza em torno das diferentes funções da gestão fiscal.

Figura 1 – Análise CHD (2007/2022)



Fonte: Software IRaMuTeQ a partir dos dados da pesquisa.

A Classificação Hierárquica Descendente (CHD) demonstra que o discurso institucional da Secretaria de Estado de Fazenda de Minas Gerais está organizado em quatro classes temáticas relativamente distintas, indicando que a comunicação oficial não se concentra exclusivamente na administração tributária, mas incorpora dimensões de fiscalização, governança, programas institucionais e recuperação fiscal. Essa segmentação confirma que a comunicação institucional expressa diferentes prioridades administrativas ao longo do tempo, conforme discutem Miola e Marques (2024).

A Classe 4 (27,3%) constitui o núcleo da gestão fiscal e apresenta elevada convergência com as dimensões Situação Fiscal e Endividamento e Governança Fiscal Restrita do dicionário lexical. A predominância de termos como ‘RRF’, ‘dívida’, ‘déficit’, ‘despesa’, ‘adesão’, União e Assembleia demonstra que a comunicação institucional enfatiza mecanismos de equilíbrio das contas públicas e

negociações federativas. Esse resultado está alinhado ao referencial sobre dívida pública e responsabilidade fiscal, que atribui à gestão do endividamento papel central na sustentabilidade financeira do estado (Ferreira, 2019; Brasil, 2025).

Sob a perspectiva da Teoria da Contingência, a concentração desses vocábulos evidencia que a SEF/MG adapta seu discurso às restrições impostas pelo ambiente econômico. A recorrência do Regime de Recuperação Fiscal demonstra que a organização modifica suas prioridades comunicacionais conforme as exigências institucionais e financeiras, corroborando a proposição de que estruturas administrativas respondem às contingências ambientais (Chiavenato, 1979; Ferreira, 2002).

A Classe 1 (18,4%) reúne termos como ‘operação’, ‘sonegação’, ‘fraude’, ‘polícia’, ‘investigação’ e ‘apreensão’, caracterizando um eixo voltado à fiscalização tributária e ao combate às irregularidades fiscais. Embora essa dimensão não apareça explicitamente como categoria independente no dicionário lexical, ela complementa a dimensão Gestão Administrativa e Financeira, evidenciando que a arrecadação depende simultaneamente de mecanismos de controle, fiscalização e conformidade tributária.

Essa classe amplia a interpretação do conceito de governança fiscal ao demonstrar que a sustentabilidade das contas públicas não depende apenas da arrecadação, mas também da capacidade institucional de reduzir perdas decorrentes da evasão fiscal. Dessa forma, a comunicação oficial reforça sua função de prestação de contas e de demonstração da capacidade estatal de proteger as receitas públicas, conforme discutido na literatura sobre comunicação institucional (Miola; Marques, 2024).

A Classe 2 (27,6%) apresenta termos relacionados a ‘programa’, ‘transparência’, ‘sociedade’, ‘escola’, ‘parcerias’ e ‘gestores’, evidenciando uma dimensão voltada às ações institucionais e ao relacionamento entre governo e sociedade. Essa configuração aproxima-se da dimensão Governança Fiscal Ampla, pois demonstra que a atuação da Secretaria ultrapassa a administração financeira e incorpora iniciativas de educação fiscal, transparência e participação social.

Esse resultado dialoga diretamente com o referencial sobre comunicação institucional, segundo o qual os órgãos públicos utilizam seus canais oficiais para divulgar resultados, fortalecer a legitimidade institucional e ampliar a transparência das políticas públicas. Assim, a comunicação da SEF/MG evidencia que a gestão fiscal também envolve processos educativos e mecanismos de aproximação com a sociedade (Capriotti *et al*., 2023; Miola; Marques, 2024).

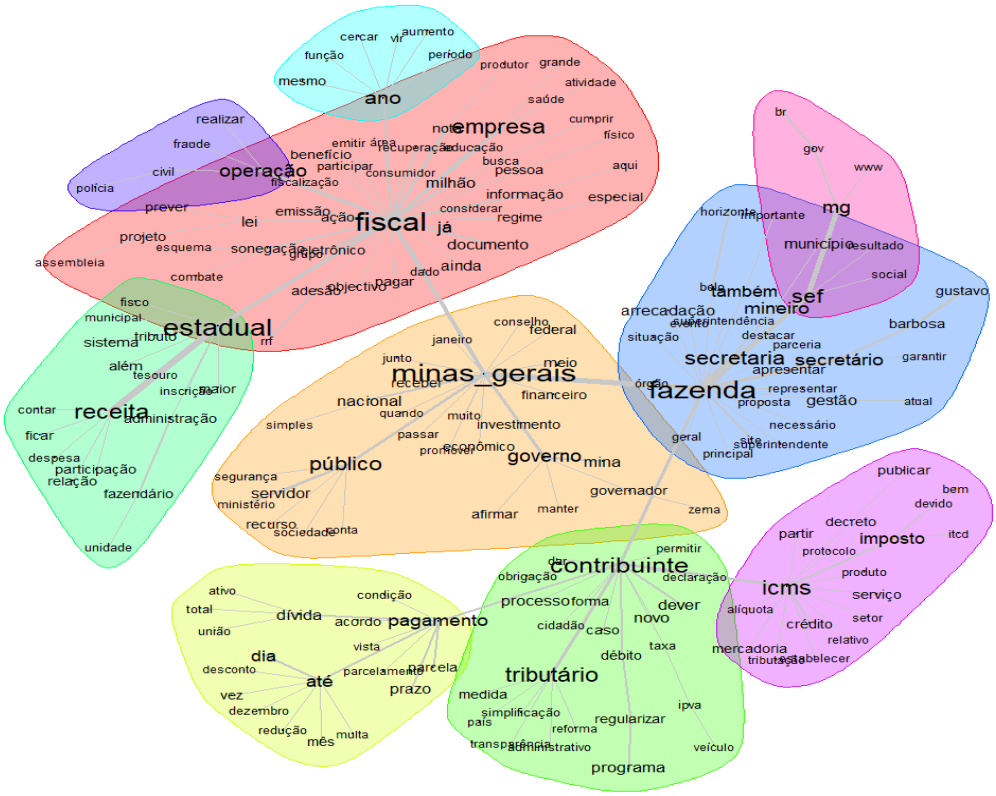
A Classe 3 (26,7%) concentra vocábulos como ‘contribuinte’, ‘SIARE’, ‘sistema eletrônico’, ‘declaração’, ‘ICMS’, ‘emissão’ e ‘documento’, representando a dimensão operacional da administração tributária. Sua proximidade com os termos previstos na dimensão Gestão Administrativa e Financeira confirma que grande parte da comunicação institucional está voltada ao atendimento ao contribuinte e à operacionalização dos serviços fazendários.

Essa predominância evidencia que a comunicação institucional possui função essencialmente instrumental, orientando contribuintes e reduzindo assimetrias de informação. Sob esse aspecto, a literatura destaca que a comunicação pública constitui mecanismo estratégico para aumentar a eficiência administrativa e facilitar a interação entre estado e sociedade (Salas-Canales, 2021; Spiazzi; Pozobon, 2024).

4.4 DISCUSSÃO INTEGRADA DA ANÁLISE DE SIMILITUDE

Complementarmente à CHD, a análise de similitude foi empregada para identificar a estrutura de coocorrência entre os termos do corpus textual. Diferentemente da classificação em classes, essa técnica demonstra as conexões existentes entre os vocábulos, permitindo compreender a organização semântica da comunicação institucional e verificar a consistência das dimensões definidas no dicionário lexical fiscal.

Figura 2 – Árvore de similitude



Fonte: Software IRaMuTeQ a partir dos dados da pesquisa (2026)

A análise integrada da árvore de similitude reforça os resultados da CHD ao revelar que o termo fiscal ocupa posição central na estrutura lexical do corpus, funcionando como principal elemento articulador da comunicação institucional. A partir desse núcleo irradiam-se conexões com ‘empresa’,

‘estadual’, ‘secretaria’, ‘fazenda’, ‘Minas Gerais’, ‘receita’, ‘contribuinte’, ‘tributário’ e ‘ICMS’, indicando elevada integração entre gestão tributária, governança institucional e administração financeira.

A centralidade do termo fiscal confirma a consistência do dicionário lexical desenvolvido, pois praticamente todas as dimensões propostas convergem para esse núcleo semântico. As conexões entre ‘receita’, ‘ICMS’, ‘contribuinte’ e ‘tributário’ corroboram a dimensão Gestão Administrativa e Financeira, enquanto as ligações com ‘secretaria’, ‘fazenda’, ‘Minas Gerais’ e ‘governo’ reforçam a dimensão Governança Fiscal Ampla.

Outro aspecto relevante é a formação de comunidades lexicais relativamente independentes, como os agrupamentos relacionados à recuperação fiscal (‘dívida’, ‘pagamento’, ‘prazo’, ‘acordo’), aos tributos (‘ICMS’, ‘imposto’), à arrecadação (‘receita’) e às operações fiscais (‘operação’, ‘fraude’). Essa organização demonstra que o discurso institucional apresenta elevada especialização temática, característica esperada em comunicações produzidas por um órgão fazendário especializado.

Sob a perspectiva da Teoria da Contingência, a árvore de similitude demonstra que essas comunidades lexicais permanecem conectadas por um mesmo núcleo institucional, indicando que as diferentes funções da Secretaria se articulam para responder às demandas impostas pelo ambiente econômico e fiscal. Assim, fiscalização, arrecadação, recuperação fiscal e comunicação ao contribuinte não representam atividades isoladas, mas componentes complementares de uma estratégia institucional adaptativa (Chiavenato, 1979; Ferreira, 2002).

Em conjunto, a CHD e a árvore de similitude validam empiricamente a estrutura do dicionário lexical fiscal. As classes identificadas pelo IRaMuTeQ apresentam elevada correspondência com as dimensões teóricas propostas, indicando que o vocabulário selecionado possui consistência conceitual e capacidade de representar os principais eixos da comunicação institucional da SEF/MG.

Consequentemente, os resultados reforçam a validade de conteúdo do Indicador de Saliência Fiscal, pois a análise de similitude confirmou empiricamente que os termos que compõem o dicionário lexical do Indicador de Saliência Fiscal não estão dispersos ou desconectados, mas organizados em uma rede semântica coerente e compatível com o construto teórico de gestão fiscal.

Também demonstram que a mineração textual foi capaz de transformar o discurso institucional em uma medida representativa das prioridades fiscais do governo estadual, pois preservou a estrutura semântica do discurso institucional ao convertê-lo em uma medida quantitativa, permitindo que o Indicador de Saliência Fiscal represente, de maneira válida e sistemática, a evolução da agenda fiscal governamental ao longo do período analisado.

5 CONCLUSÃO

O presente estudo desenvolveu um indicador de saliência fiscal da comunicação institucional da Secretaria de Estado de Fazenda de Minas Gerais, demonstrando que técnicas de mineração textual podem ser utilizadas para transformar documentos oficiais em indicadores quantitativos capazes de representar prioridades institucionais ao longo do tempo.

A principal contribuição metodológica consistiu na reconstrução e auditoria do corpus documental, procedimento que ampliou significativamente a representatividade da base de dados e reduziu distorções decorrentes de lacunas documentais. A identificação dos chamados zeros artificiais fortaleceu a confiabilidade das séries temporais e permitiu distinguir ausência de comunicação de ausência de conteúdo fiscal.

Outro avanço relevante foi a elaboração de um dicionário lexical estruturado em cinco dimensões temáticas, fundamentadas no referencial sobre dívida pública, equilíbrio fiscal, governança e comunicação institucional. Essa organização reduziu a subjetividade inerente às análises textuais e proporcionou maior validade conceitual ao indicador desenvolvido.

Os resultados obtidos pela Classificação Hierárquica Descendente demonstraram que a comunicação institucional da SEF/MG se organiza em diferentes eixos temáticos relacionados à recuperação fiscal, fiscalização tributária, administração financeira e governança pública. Essa estrutura demostra que o discurso institucional reflete as múltiplas funções desempenhadas pela administração fazendária estadual.

A análise de similitude complementou esses resultados ao revelar elevada integração entre os principais termos fiscais, indicando que conceitos relacionados à arrecadação, tributação, recuperação fiscal e governança institucional formam uma estrutura semântica consistente. Essa convergência reforça empiricamente a adequação do dicionário lexical utilizado na construção do indicador.

Do ponto de vista teórico, os resultados corroboram a Teoria da Contingência ao demonstrar que a comunicação institucional acompanha as mudanças do ambiente econômico e institucional. As variações observadas no discurso oficial refletem adaptações organizacionais frente às necessidades de gestão das finanças públicas, evidenciando que a comunicação também constitui um mecanismo de resposta às contingências fiscais.

Como contribuição científica, o estudo amplia a literatura ao propor um procedimento metodológico reproduzível para construção de indicadores baseados em comunicação institucional. Diferentemente dos indicadores tradicionais fundamentados exclusivamente em variáveis econômicas, o indicador desenvolvido incorpora informações textuais capazes de captar mudanças qualitativas nas prioridades governamentais.

Sob a perspectiva aplicada, o Indicador de Saliência Fiscal representa uma ferramenta potencial para o monitoramento contínuo da comunicação institucional, podendo subsidiar avaliações sobre transparência, governança e comportamento organizacional em diferentes órgãos públicos. Sua estrutura metodológica também permite adaptação para outras instituições e diferentes contextos administrativos.

Como limitação, a pesquisa utiliza exclusivamente documentos oficiais produzidos pela SEF/MG, representando apenas a comunicação institucional do órgão. Dessa forma, o indicador não mensura diretamente a execução das políticas fiscais nem a percepção da sociedade sobre as ações governamentais.

Pesquisas futuras poderão integrar o indicador de saliência fiscal a modelos econométricos e séries macroeconômicas, permitindo investigar sua capacidade de antecipar mudanças no comportamento fiscal, no endividamento público e na dinâmica econômica regional. Essa ampliação poderá consolidar o indicador como instrumento de apoio à formulação e avaliação de políticas públicas.

REFERÊNCIAS BIBLIOGRÁFICAS

AFFEMG, Associação dos Funcionários Fiscais do Estado de Minas Gerais. **Os fiscais de Minas**: a história da AFFEMG (1950-1997). Belo Horizonte, Dossiê: Agência de Investigação Histórica, 1997.

BATISTA JR, Onofre Alves. **O Poder de Polícia Fiscal**. Ed. Mandamentos. Belo Horizonte, 2001.

BRASIL. Câmara dos Deputados. Consultoria de Orçamento e Fiscalização Financeira. **Entenda o Orçamento**. Brasília. Disponível em: http://www2.camara.leg.br/atividade-legislativa/orcamentobrasil/entenda/cartilha/cartilha.pdf. Acesso em: 06 out. 2024.
Acesso em: 06 outubro 2024.

BRASIL. Câmara dos Deputados. **Portal institucional da Câmara dos Deputados**: propostas legislativas, notícias e agenda parlamentar. Brasília, 2026. Disponível em: https://www2.camara.leg.br/. Acesso em: 16 abr. 2026.

BRASIL. Ministério da Fazenda. Tesouro Nacional. **Portal do Tesouro Nacional – Informações, Serviços e Transparência Fiscal**. Disponível em: https://www.tesourotransparente.gov.br/. Acesso em: 22 jul. 2026.

CAPRIOTTI, P., Losada-Dí-az, J.-C., e Martinez-Gras, R. (2023). Evaluating the content strategy developed by universities on social media. **Profesional de La Información**, 32(2), Article 2. https://doi.org/10.3145/epi.2023.mar.10. Acesso em: 12 jul. 2026.

CAPRIOTTI, P., Martínez-Gras, R., e Zeler, I. (2023). Does universities' posting strategy influence their social media engagement? An analysis of the top-ranked higher education institutions in different countries. **Higher Education Quarterly**, 77(4), 911–931. https://doi.org/10.1111/hequ.12439. Acesso em: 12 jul. 2026.

CAPRIOTTI, P., Oliveira, A., e Zeler, I. (2023). Analysis of universities' integrated communication strategies on Twitter-X. **Profesional de la información**, 32(6), Article 6. https://doi.org/10.3145/epi.2023.nov.16. Acesso em: 12 jul. 2026.

CAPRIOTTI, P., e Zeler, I. (2023). Analysing effective social media communication in higher education institutions. **Humanities and Social Sciences Communications**, 10(1), 1–13. https://doi.org/10.1057/s41599-023-02187-8. Acesso em: 12 jul. 2026.

CARVALHO DE AMORIM, Carina; SILVA VALDÃO, Eldys; DA SILVA OLIVEIRA, Francielle; LOPES DA CUNHA, Isaias; ALMEIDA-SANTOS, Paulo Sérgio. Dívida pública dos estados brasileiros e a Lei de Responsabilidade Fiscal (2000–2025): uma análise da sustentabilidade fiscal. **Revista Controle - Doutrina e Artigos**, Fortaleza, CE, Brasil, v. 24, n. 2, p. 411–439, 2026. DOI: 10.32586/rcda.v24i2.1105. Disponível em: https://revistacontrole.tce.ce.gov.br/index.php/RCDA/article/view/1105.. Acesso em: 5 out. 2026.

CHIANENATO I. Novas abordagens na teoria administrativa. **Rev adm empres** [Internet]. 1979May;19(2):27–42. Available from: https://doi.org/10.1590/S0034-75901979000200002. Acesso em: 9 jul. 2026.

COMPARATO, Fábio Konder. Organização constitucional da função planejadora. **Desenvolvimento Econômico e Intervenção do Estado Na Ordem Constitucional**: Estudos Jurídicos em Homenagem Ao Professor Washington Peluso Albino de Souza. Tradução. Porto Alegre: Fabris, 1995. . . Acesso em: 16 abr. 2026.

CORRÊA, C. A. R et al. a sociedade da informação e do conhecimento e os estados brasileiros. **Informação e Informação**, Londrina, v. 19, n. 1, p. 31-54,jan./abr. 2014.

CUNHA, P. M. D.; GARCIA, M. G. A gerência recente do endividamento público brasileiro. **Brazilian Journal of Political Economy**, 32, p. 260-281, 2012.

DE HOLANDA, Cínthia Maria Silva; DOS SANTOS, Raimundo Nonato Macedo. Análise métrico-temática da produção normativa de Leis Municipais: contribuições para a governança urbana baseada em dados. **ConCI: Convergências em Ciência da Informação**, v. 9, 2026.

FAGUNDES, J. A., PETRI, M., LAVARDA, R. B., RODRIGUES, M. R., Lavarda, C. E. F., e Soller, C. C. (2011). Estrutura Organizacional E Gestão Sob a Ótica Da Teoria Da Contingência. **Gestão e Regionalidade**, 26(78). https://doi.org/10.13037/gr.vol26n78.792. Acesso em: 06 jun. 2026.

FERREIRA, F. G., 2019, A regra de ouro como limite constitucional ao endividamento público no Brasil. **Blucher Open Access**. 607-640.

FERREIRA, A. A.; REIS, A. C. F. e PEREIRA M. I. (2002). **Gestão Empresarial**: de Taylor aos nossos dias. São Paulo: Pioneira.

MINAS GERAIS. Governo do Estado. **Portal de Dados Abertos do Governo de Minas Gerais**. Disponível em: https://dados.mg.gov.br/. Acesso em: 15 mai. 2026.

MINAS GERAIS. Governo do Estado. **Agência Minas – Portal Oficial de Notícias do Governo de Minas Gerais**. Disponível em: https://www.agenciaminas.mg.gov.br/. Acesso em: 15 mai. 2026.

MINAS GERAIS. Secretaria de Estado da Fazenda. **História institucional da Secretaria de Estado da Fazenda de Minas Gerais**. Disponível em: https://www.fazenda.mg.gov.br/secretaria/historia/. Acesso em: 15 mai. 2026.

MINAS GERAIS, Secretaria de Estado da Fazenda. **Secretaria de Estado da Fazenda**: 100 anos de história, 1891-1991. Dossiê: Agência de Investigação Histórica, Belo Horizonte,1991. Acesso em: 14 ago. 2026.

MIOLA, Edna; MARQUES, Francisco Paulo Jamil. Abordagens teóricas e definições mínimas na pesquisa em Comunicação Pública: construindo o conceito de Comunicação Pública de Estado. **Organicom**, v. 21, n. 45, p. 19-34, 2024.

NEUHOLD, Roberta dos Reis; POZZER, Márcio Rogério Olivato; PELLEJERO, David Cristopher. A desvalorização das licenciaturas na educação superior brasileira: análise do planejamento estratégico dos Institutos Federais. **Revista Electrónica "Actualidades Investigativas en Educación"**, [S. l.], v. 25, n. 1, p. 1-33, 2025. Disponível em: https://www.redalyc.org/journal/447/44780305027/html/. DOI: https://doi.org/10.15517/aie.v25i1.60755. Acesso em: 16 abr. 2026.

MINAS GERAIS. Secretaria do Estado de Fazenda. **RESOLUÇÃO CONJUNTA SEF/AGE Nº 5.625 DE 11 DE NOVEMBRO DE 2022 - SEF/MG**. Disponível em: https://www.fazenda.mg.gov.br/empresas/legislacao_tributaria/resolucoes/2022/rr5625_2022.html. Acesso em: 16 de jul. 2026.

ROCHA, Saul Sousa da. Mineração de padrões textuais sobre segurança pública: modelos e aplicações em plataformas de mídias sociais. 2026. **Repositório Institucional da UFPI (RI/UFPI)**. Disponível em: https://repositório.ufpi.br/handle/123456789/4212/. Acesso em: 05 jun. 2026.

SALAS-Canales, H. J. (2021). Endomarketing: Una herramienta para la gestión efectiva del capital humano. **Revista Internacionalde Investigación en Ciencias Sociales**, 17(1), 126–142. https://doi.org/10.18004/riics.2021.junio.126. Acesso em: 21 jun. 2026.

SCHUMANN, Lívia Rejane Miguel Amaral; MOURA, Leides Baroso Azevedo. Índices sintéticos de vulnerabilidade: uma revisão integrativa de literatura. **Ciência e Saúde Coletiva**, v. 20, n. 7, p. 2105-2120, 2015.

SHABBIR, Muhammad Khurram. Tax Salience: A review of the literature. **JISR management and social sciences e economics**, v. 21, n. 4, p. 55-72, 2023.

SILVA, M. R; HAYASHI, C. R.; HAYASHI, M. C. P. I. Análise bibliométrica e cientométrica: desafios para especialistas que atuam no campo. **InCID: Revista de Ciência da Informação e Documentação**, Ribeirão Preto, v. 2, n. 1, p. 110-129, jan./jun. 2011.

SOBRINHO, Samara Lima; CALLADO, Aldo Leonardo Cunha. Análise das relações entre os fatores contingenciais e as práticas de contabilidade gerencial adotadas à luz da teoria contingencial: um estudo nos hotéis dos municípios de Recife e Ipojuca do estado de Pernambuco. **REVISTA AMBIENTE CONTÁBIL**-Universidade Federal do Rio Grande do Norte-ISSN 2176-9036, v. 17, n. 1, p. 444-464, 2025.

SOUSA, Y. S. O. (2021). O Uso do Software Iramuteq: Fundamentos de Lexicometria para Pesquisas Qualitativas. **Estud. pesqui. psicol**. (Impr.), 21(4), 1541-1560. https://doi.org/10.12957/epp.2021.64034. Acesso em: 28 ago. 2026.

SPIAZZI, S. C., e Pozobon, R. de O. (2024). Estratégias de comunicação Estratégias de comunicação pública de uma universidade na sociedade de plataforma pública de uma universidade na sociedade de

plataformas. **Cuadernos.info**, 58, Article 58. https://doi.org/10.7764/cdi.58.68557. Acesso em: 30 ago. 2026.

TOZZI, Tatiana Es. Um método para o apoio e análise de documentos textuais empregando mineração de texto e taxonomia. 2024. **Repositório Institucional UFSC**. Disponível em: repositorio.ufsc.br/handle/123456789/260682. Acesso em: 25 jul. 2026.

VIOLA, Carmine e outros. Quanto mais você sabe, melhor você age? Comunicação institucional na gestão de crises da Covid-19. **Previsão Tecnológica e Mudança Social**, v. 170, p. 120929, 2021.