ANNALS OF DATA ANALYSIS

Evidências sobre o uso de Processamento de Linguagem Natural e Machine Learning como ferramenta diagnóstica à transversalidade em Pernambuco

Ryan Pablo Alves de Almeida Pires [1] * - https://orcid.org/0009-0000-0215-2786
Mariana Crystinne Pereira Batista [2] - https://orcid.org/0009-0008-1710-4048
Bruno Costa de Abreu e Melo [2] - https://orcid.org/0009-0001-7664-5337
Larissa Karla Souza da Silva [2] - https://orcid.org/0009-0000-1011-9337
Cicera Vitória Amaro da Silva [2] - https://orcid.org/0009-0000-4454-9597
Ana Beatriz Castelo Branco de França Santos [2] - https://orcid.org/0009-0001-8577-9679
Maria do Carmo Soares de Lima [2] - https://orcid.org/0000-0002-5480-3103

[1] Departamento de Ciência Política, Universidade Federal de Pernambuco, Recife - PE, Brasil
Emails: ryan.almeida@ufpe.br, ryallmeida@gmail.com
* autor correspondente

[2] Departamento de Ciência Política, Universidade Federal de Pernambuco, Recife - PE, Brasil
Emails: mariana.crystinne@ufpe.br, bruno.cmelo@ufpe.br, larissa.karla@ufpe.br, cicera.vitoria@ufpe.br, ana.castelo@ufpe.br, maria@de.ufpe

RESEARCH ARTICLE

Annals of Data Analysis
https://periodicos.ufpe.br/revistas/ADA/

Submitted on: 12.12.2025.
Accepted on: 18.08.2026.
Published on: 18.08.2026.

Resumo

Este estudo concentra-se na identificação e análise de práticas de transversalidade presentes no estado de Pernambuco, com base na extração de padrões e evidências empíricas a partir de dados textuais oriundos de documentos públicos do Estado, como os Planos Plurianuais, Leis de Diretrizes Orçamentárias e Lei Orçamentária Anual. A investigação busca diagnosticar, em que medida há expressão de integração entre níveis de governo na formulação políticas no contexto pernambucano. Para tanto, adota-se uma abordagem empírico-analítica orientada pela abordagem quantitativa, visando identificar indícios da intensidade das práticas transversais no território analisado utilizando o Processamento de Linguagem Natural e Aprendizado de Máquina.

Palavras-chaves

Transversalidade, Política Pública, Processamento de Linguagem Natural, Naive Bayes, Linear SVC

1. Por que políticas transversais importam e o que o Índice de Efetividade da Gestão Municipal nos conta?

De acordo com Gallo, a transversalidade pode ser entendida como um processo de atravessamento recíproco entre diferentes campos de saber, os quais, mesmo partindo de suas especificidades, se interpenetram, se misturam e se mesclam, e se cruzam. Contudo, essa interação não anula suas particularidades; ao contrário, cada saber amplia-se justamente nesse encontro com a multiplicidade (Gallo, 2007, *apud* Avelino e Santos, 2014, p. 11).

Esse fundamento reconhece que os problemas sociais são complexos e interconectados. Logo, exigem abordagens integradas para serem enfrentadas. Assim, ele busca superar a fragmentação das políticas setoriais, vindo com uma visão mais holística e coordenada. Assim, observa-se que a transversalidade tem sido uma variável independente que historicamente foi negligenciada pela literatura como preditor em potencial para explicar o sucesso e efetividade de Políticas Públicas.

Há evidências de que a transversalidade nas Políticas Públicas começou a ser introduzida pela primeira vez no governo federal brasileiro em meados de 2003, representando uma inovação essencial na gestão de questões complexas. Nesse sentido, o Plano Plurianual 2004-2007, destacou a importância de temas transversais, como gênero, meio ambiente, desigualdades, ciência e emprego, que deveriam perpassar todas as ações governamentais.

Ainda nesse documento, o "*Mega-objetivo I: Inclusão Social e Redução das Desigualdades Sociais*" foi o precursor que operacionalizou, na prática, os objetivos previstos por meio do "*Programa Fome Zero*" que coordenou a ação de 19 ministérios para implementar políticas de combate à fome no Brasil. A articulação transversal entre diferentes áreas da administração pública demonstrou como políticas integradas podem gerar impactos significativos ao atacar a fome e a pobreza em várias frentes simultaneamente.

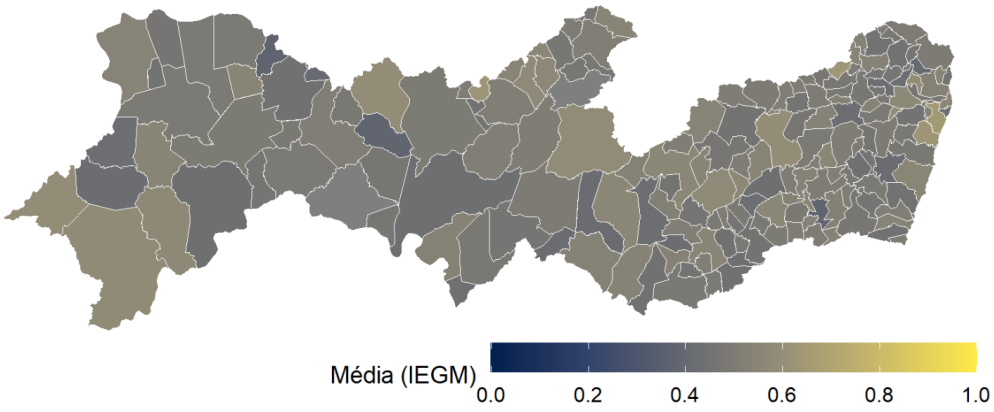
Assim como a experiência nacional demonstrou, a realidade em Pernambuco também pode ser investigada por meio dos seus principais instrumentos de planejamento: o Plano Plurianual (PPA), a Lei de Diretrizes Orçamentárias (LDO) e a Lei Orçamentária Anual (LOA). Esses documentos, revelam como o governo estadual organiza e executa suas políticas, estabelecendo prioridades, metas e que em seu âmago podem indicar o grau de articulação entre diferentes setores e níveis de governo no estado.

E se governar bem não é apenas executar políticas, mas também decidir com precisão onde, como e por que fazê-lo? Assim, os instrumentos de planejamento cristalizam essas perspectivas e dão o "norte" a ser seguido. Em contextos como o brasileiro, a governança multinível e a estrutura federativa, dentro de uma lógica funcionalista, parte da premissa de que diferentes esferas de governo são mais eficientes na oferta de determinados serviços públicos. Mas como avaliar se essa eficiência e efetividade, teorizada por esse modelo, se concretiza na prática?

Nesse cenário, o Índice de Efetividade da Gestão Municipal (IEGM) surge como um instrumento de análise, ele padroniza a mensuração da qualidade da gestão municipal. Sua finalidade vai além da simples avaliação: o IEGM foi concebido para subsidiar o controle externo, orientar decisões administrativas e oferecer transparência à sociedade quanto ao desempenho das administrações locais. Integrado ao Marco de Medição de Desempenho, o IEGM qualifica o debate sobre a efetividade das políticas ao atuar como ferramenta diagnóstica. Além disso, possibilita avaliar como os municípios organizam ações em áreas interdependentes, fornecendo evidências *in proxy* sobre o grau de transversalidade e articulação intersetorial, dado que é uma variável que versa sobre qualidade administrativa.

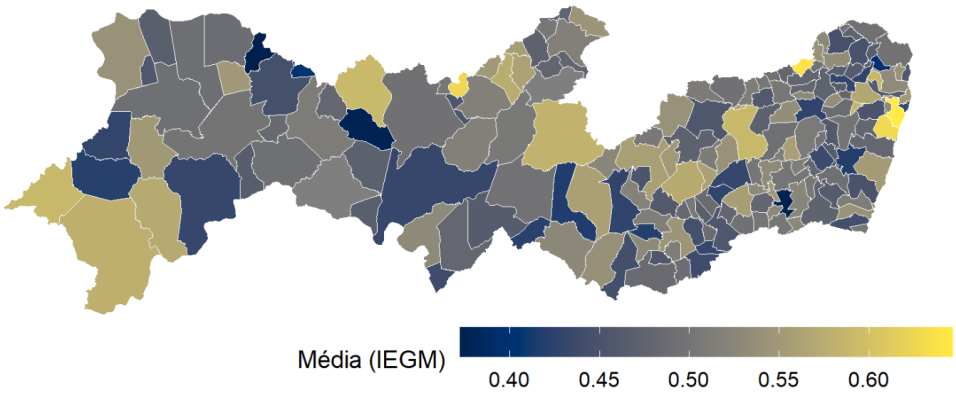
A partir da análise do índice em questão, foi possível desenvolver dois mapas temáticos e representativos do estado de Pernambuco. O primeiro mapa (*vide* Figura 1) demonstra a média estimada do IEGM para o estado oferecendo uma perspectiva projetada da efetividade administrativa municipal em escala de território nacional. Já o segundo mapa (*vide* Figura 2) retrata a realidade observada no território pensando apenas o estado de Pernambuco, revelando possíveis disparidades entre as expectativas institucionais e os desempenhos reais das gestões locais. Essa abordagem permite uma análise comparativa entre territórios do estado.

Figura 1. Efetividade Média da Gestão Municipal em Pernambuco considerando o território nacional (2017-2023)



Elaboração própria. Pires et al. (2026)
Fonte: Instituto Rui Barbosa/TCE-PE

Figura 2. Efetividade Média da Gestão Municipal em Pernambuco considerando apenas o território estadual (2017-2023)



Elaboração própria. Pires et al. (2026)
Fonte: Instituto Rui Barbosa/TCE-PE

Ora, se a efetividade média da gestão municipal em Pernambuco permanece abaixo do ideal, o que isso revela sobre a capacidade dos entes locais de produzir políticas com resultados concretos? Mais do que uma questão de desempenho administrativo, esse dado lança luz sobre a questão e convida a uma reflexão mais ampla: até que ponto a baixa efetividade é sintoma de uma estrutura federativa que, embora descentralizada, ainda opera de forma fragmentada e sem

cooperação em um determinado grau? Ao reconhecer que políticas públicas eficazes dependem de interações coordenadas entre diferentes esferas de governo, esse modelo explicita a necessidade de articulação e cooperação.

No entanto, a simples existência de múltiplos níveis de governança não garante, por si só, a superação dos gargalos na gestão municipal. É nesse cenário que a transversalidade pode ser mais do que um princípio abstrato, pode ser uma resposta concreta. Ao integrar políticas e setores historicamente compartimentalizados, a transversalidade pode ampliar a capacidade do município de responder a problemas complexos, reforçando sua efetividade e qualificando sua atuação no arranjo federativo

Dado o exposto, embora o uso de técnicas computacionais para análise textual tenha antecedentes na linguística computacional desde meados do século XX, sua incorporação sistemática às ciências sociais é relativamente recente. Estudos pioneiros na área de Ciência Política começaram a empregar métodos automatizados de análise de texto a partir dos anos 1990, sobretudo em pesquisas sobre discursos parlamentares, manifestos partidários e posicionamentos ideológicos (Laver; Benoit; Garry, 2003). Contudo, foi apenas a partir da década de 2010, com o avanço do *machine learning*, do aumento da capacidade computacional e da disponibilização massiva de dados públicos digitais, que essas técnicas passaram a ser amplamente utilizadas para a análise de políticas públicas.

A popularização do Processamento de Linguagem Natural (PLN, ou do inglês NPL) no campo da Administração Pública está diretamente associada à consolidação da chamada *data-driven public policy analysis*, que busca integrar grandes bases de dados administrativos a métodos quantitativos avançados para fins de diagnóstico, monitoramento e avaliação de políticas. Nesse movimento, técnicas como classificação supervisionada e aprendizado de máquina passaram a ser mobilizadas não apenas para prever resultados, mas para investigar dimensões tradicionalmente consideradas abstratas ou de difícil mensuração, como coordenação intersetorial, capacidade institucional, coerência programática e governança multinível (Peters, 1998; Pollitt, 2003; Peters, 2015).

Diferentemente de variáveis diretamente observáveis, como gastos e indicadores de desempenho, já conceitos centrais da literatura de políticas públicas, como o objeto de estudo deste *paper* manifestam-se predominantemente no plano discursivo e normativo. Eles aparecem incorporados à linguagem formal do Estado, expressos em diretrizes, objetivos, metas e justificativas presentes nos documentos de planejamento e gestão. Assim, o PLN torna-se particularmente adequado para operacionalizar empiricamente esses conceitos, ao permitir a identificação de padrões semânticos, frequências lexicais, associações contextuais e estruturas latentes que dificilmente seriam captadas por leituras manuais extensivas (Jurafsky; Martin, 2018).

Nesse sentido, a literatura recente tem demonstrado que documentos como Planos Plurianuais, Leis de Diretrizes Orçamentárias e Leis Orçamentárias Anuais não são apenas instrumentos técnicos de planejamento, mas registros privilegiados das prioridades políticas e dos modos de coordenação do Estado. A análise automatizada desses documentos permite examinar empiricamente se princípios amplamente defendidos no plano normativo como a transversalidade, efetivamente se consolidam na linguagem e na estrutura discursiva da ação governamental.

O uso combinado de PLN e algoritmos de aprendizado de máquina podem responder a um desafio metodológico recorrente na análise de políticas públicas: a ambiguidade semântica da linguagem institucional. Textos administrativos tendem a ser vagos, genéricos e marcados por fórmulas retóricas que dificultam classificações dicotômicas simples. Ao empregar modelos estatísticos e regras de decisão calibradas teoricamente, essas técnicas

permitem lidar com zonas cinzentas do discurso estatal, oferecendo diagnósticos mais robustos e sistemáticos (Izbicki; Santos, 2017).

Em suma, talvez o desafio não esteja apenas em mensurar a transversalidade, mas em reinterpretá-la à luz das lacunas estruturais da governança local, buscando estratégias que a elevem, assim como reestruturar as formas de governar para que a transversalidade se torne um vetor de transformação.

Nesse cenário, este trabalho parte do seguinte problema central: em que medida a transversalidade é evidenciada nos documentos oficiais de planejamento do Estado de Pernambuco e como técnicas de mineração de texto baseadas em Processamento de Linguagem Natural e algoritmos de *Machine Learning* podem colaborar para captá-la, especificamente quando se trata de um conceito abstrato em um contexto de escassez de consensos sobre seu estudo empírico?

2. Referencial Teórico

A literatura, acumulou diferentes perspectivas analíticas (Silva e Calmon, 2017; Lotta, 2019), mas ainda assim a transversalidade permanece um conceito pouco explorado em profundidade, apesar de seu uso recorrente. Uma política com uma implementação mal articulada, além de comprometer resultados, enfraquece a confiança pública e revela limites da coordenação estatal, da eficiência burocrática e do accountability (Pollitt, 2003). Nesse contexto, a literatura tem discorrido em três eixos centrais: governança multinível, teorias de implementação e capacidade institucional.

A governança multinível, marcada pela dispersão de competências entre jurisdições, implica desafios de coordenação (Hooghe, Marks e Schakel, 2017; Peters, 1998; Peters, 2015). A literatura evidencia que problemas interconectados exigem arranjos mais integrados, mas disputas políticas e interesses setoriais dificultam a cooperação (Pollitt, 2003), portanto, é congruente advogar e alarmar acerca do princípio da transversalidade política. As teorias de implementação, por sua vez, oscilam entre visões *top-down*, que enfatizam o planejamento e a clareza dos objetivos (Pressman e Wildavsky, 1973; Arretche, 2001), e abordagens *bottom-up*, que analisam práticas concretas e dinâmicas locais (Elmore, 1979; Hjern; Porter, 1981; Barrett, 2004). Ambas, contudo, enfrentam limitações diante da complexidade das políticas transversais.

A capacidade institucional emerge, assim, como fator decisivo: sem burocracias qualificadas, incentivos institucionais e mecanismos de cooperação intersetorial, a descentralização se traduz em ambiguidade e sobreposição de competências (Arretche, 2002; 2011; Lotta, 2019). Nesse cenário, redes colaborativas podem mitigar a fragmentação (Klijn; Koppenjan, 2000), mas também reproduzem desigualdades de poder. A transversalidade, portanto, desafia tanto a teoria quanto a prática, exigindo assim modelos de governança mais policêntricos e mecanismos robustos de responsabilização (Ostrom, 2005; Goodin, 2006).

3. Metodologia

O Processamento de Linguagem Natural é uma das ramificações da inteligência artificial, onde concentra seu processamento na interação entre máquina e linguagem humana. Sua principal função é criar processos de análise de grandes quantidades de dados de texto, sendo capaz de analisar, entender, interpretar e gerar uma linguagem próxima ou exatamente correspondente com a linguagem humana (Bird & Klein, 2009 *apud* Lavezzo e Conceição,

2023, p. 67). Por meio do uso eficaz do Processamento de Linguagem Natural e da análise de tópicos, pode-se decifrar tendências e padrões ocultos que possam fornecer *insights* valiosos para diversas aplicações.

O processo de descoberta de conhecimento tem sido empregado para encontrar informações ou padrões de informações, utilizando técnicas de análise e extração de dados. A descoberta de conhecimento em documentos textuais está inserida no contexto da descoberta de conhecimento em dados não estruturados. Este estudo coletou dados de documentos textuais disponibilizados pela Secretaria de Planejamento, Gestão e Desenvolvimento Regional de Pernambuco[1] (SEPLAG - PE) e por utilizarmos dados em domínio público não houve necessidade de submissão do projeto ao Comitê de Ética da Universidade Federal de Pernambuco.

Assim, o presente estudo adota uma abordagem quantitativa, fundamentada no uso de algoritmos computacionais para a análise automatizada de documentos oficiais. O algoritmo desenvolvido realiza a extração, o processamento e a análise de conteúdo textual desses documentos, convertendo-os integralmente em uma *corpora* estruturada em formato *tidy (vide* Tabela 1), no qual cada observação representa um parágrafo dos textos oficiais de planejamento do Estado.

Tabela 1. Estruturação da Corpora em formato Tidy

Nome	Variável	Tipo	Descrição
Arquivo	arquivo	character	Essa variável nasce de quando o texto completo do arquivo é dividido em parágrafos usando quebras de linha duplas como delimitador. Assim, o nome do arquivo de origem de onde esse parágrafo foi minerado é conservado, servindo como variável de id.
Id do Paragrafo	paragrafo_id	integer	Variável de id que cada parágrafo recebe um identificador sequencial que varia de 1 ao enésimo parágrafo.
Documento	documento	factor	Versa sobre qual instrumento de planejamento do Estado o parágrafo corresponde. Divide-se em três categorias: “LDO”, “PPA” e “LOA”.
Ano	ano	factor	Respectivo ano a qual o documento/parágrafo pertence.
Texto	texto_limpo	character	Corresponde ao texto do corpus limpo e normalizado: aplica-se a remoção de *stopwords* em português e de palavras com menos de 4 caracteres, assim como, o texto é tokkenizado em palavras e as palavras restantes são recombinadas em uma única string com espaços normalizados.
Classificador Tipo 1 (TP1)	label_final	integer	Cada parágrafo recebe uma classificação trivalente que indica presença, ambiguidade/neutra, ou ausência/negação de transversalidade conforme dicionários semânticos pré-definidos.

[1] Cf. Repositório de dados e informações da Secretaria de Planejamento, Gestão e Desenvolvimento Regional de Pernambuco. Disponível em: https://www.seplag.pe.gov.br/orcamento

Soma das Intensidades	soma_intensidades	double	Dá-se a construção de uma medida quantitativa de sentimento a partir de um léxico ponderado manualmente (*open access vide* Anexos), aplicada aos textos previamente limpos. O léxico é consolidado por termo, somando-se as intensidades associadas a cada palavra, de modo a obter um único peso semântico agregado por entrada lexical. Na sequência, para cada texto identificado, as intensidades associadas às palavras reconhecidas no léxico são somadas, resultando em um escore total de sentimento por unidade textual. Por fim, textos que não contêm nenhum termo do léxico recebem valor zero na variável de *soma_intensidades.*
Classificador Tipo 2 (TP2)	espectro_tipo2	integer	Essa variável deriva-se da extração dos limites inferior e superior, os quartis e a mediana do boxplot da variável soma_intensidades. Com base nesses parâmetros, o intervalo negativo (entre o limite inferior e a mediana) e o intervalo positivo (entre a mediana e o terceiro quartil) são subdivididos em faixas internas de igual comprimento. Em seguida, cada texto é classificado em uma única categoria de sentimento conforme a posição do seu escore: valores extremos abaixo do limite inferior e acima do limite superior são definidos como hipernegativos [- 6] e hiperpositivos [+ 6]; valores nulos são classificados como neutros [±0]; e os demais são distribuídos em níveis graduais de negatividade [-5, -2] ou positividade [+5, +2].
TF-IDF	tfidf_medio	double	Inicialmente, destrincha-se a corpora em nível de palavra. Em sequência, gera-se a matriz de frequências de termos (DFM) agrupada pela variável temporal, consolidando os documentos por ano. Sobre essa matriz agrupada é aplicado o cálculo de TF-IDF, produzindo pesos que refletem a relevância relativa de cada termo no conjunto anual de documentos. Em seguida, os valores de TF-IDF são convertidos para formato tabular e somados por termo, gerando um peso global de TF-IDF para cada palavra na *corpora*. Esses pesos são associados às palavras presentes em na *corpora* e atribuindo valor zero aos termos sem correspondência. Por fim, para cada *paragrafo_id*, a variável *tfidf_medio* é calculada como a média aritmética dos valores de TF-IDF das palavras que compõem o parágrafo, resultando em uma medida sintética da relevância média dos termos daquele parágrafo no corpus anual.
	tfidf_intensidade	double	Multiplicação direta entre *tfidf_medio* e *soma_intensidades* em nível de parágrafo. Essa operação combina a relevância lexical média do texto com a intensidade de sentimento, gerando um indicador composto que pondera a carga semântica do parágrafo

			pela importância relativa dos termos no corpus.
Classificador Tipo 3 (TP3)	espectro_tipo3	integer	*Idem* ao *espectro_tipo2*, onde substitui-se *soma_intensidades* por *tfidf_intensidad*e.

Elaboração própria. Pires *et al.* (2026)

A etapa de preparação e análise dos dados textuais foi realizada no R, com o auxílio de um conjunto integrado de pacotes voltados à manipulação e organização. Os pacotes *pdftools*, *stringr* e *stopwords* foram utilizados para a leitura dos documentos, limpeza textual e remoção de elementos linguísticos não informativos. A estruturação da base de dados seguiu o paradigma *tidy*, por meio dos pacotes do *tidyverse* como *tibble* e *purrr,* e outros como *tidytext*, permitindo a organização dos textos em formato tabular e a aplicação sistemática de operações de transformação e agregação.

Adicionalmente, os pacotes *quanteda* e *quanteda.textstats* foram empregados na construção do corpus e na extração de medidas descritivas, tais como estatísticas de frequência e outras métricas textuais relevantes para a caracterização do material analisado. Por fim, os pacotes *scales*, *viridis* e o *ggplot2* foram utilizados para o tratamento de escalas e apoio à visualização exploratória dos resultados, contribuindo para uma análise mais consistente e replicável dos dados.

Dada os processos de categorização supracitados (*vide* Tabela 1), a corpora foi balanceada por *random undersampling*, com amostragem aleatória sem reposição limitada ao tamanho da menor classe, com o objetivo de reduzir viés de classe e riscos de *overfitting*. Em seguida, os dados balanceados foram processados em ambiente Python, utilizando os pacotes *pandas*, *numpy* e *Scikit-Learn*, para a construção e avaliação de modelos supervisionados de classificação de sentimentos.

Os textos pré-processados foram definidos como variáveis explicativas, enquanto os rótulos de sentimento constituíram as variáveis dependentes. Inicialmente, estimou-se um modelo baseado em *Support Vector Machines* (SVM), no qual os textos foram transformados em uma representação numérica por meio da vetorização TF-IDF, que pondera a frequência dos termos pela sua raridade no corpus, resultando em uma matriz de características esparsa e de alta dimensionalidade. Essa matriz foi particionada em subconjuntos de treinamento e teste, destinando-se 75% das observações ao ajuste do modelo e 25% à avaliação fora da amostra, com estratificação das classes e semente fixa para garantir reprodutibilidade. O classificador adotado foi um SVM linear (*LinearSVC*), e seu desempenho foi avaliado a partir das métricas de precisão, revocação, F1-score e acurácia.

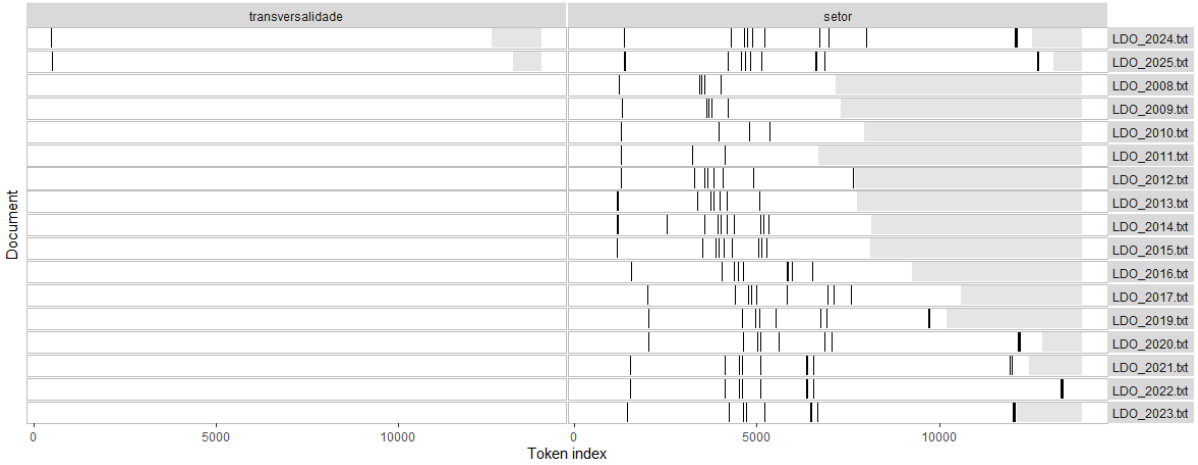
Complementarmente, estimou-se um segundo modelo supervisionado baseado no *Naive Bayes* Multinomial, com o objetivo de realizar uma análise comparativa entre abordagens fundamentadas em pressupostos distintos. Nesse caso, os textos foram vetorizados por meio do *CountVectorizer*, que converte o corpus em uma matriz documento-termo baseada na frequência absoluta das palavras. Para reduzir ruído lexical e evitar termos excessivamente dominantes, foram aplicados limiares mínimos e máximos de frequência dos termos. A esparsidade da matriz resultante foi calculada como medida descritiva do corpus vetorizado. Assim como no modelo SVM, os dados foram divididos em conjuntos de treinamento e teste mantendo a proporção de 75/25, com estratificação das classes. O modelo Naive Bayes Multinomial foi ajustado sobre o conjunto de treinamento e avaliado no conjunto de teste por meio de relatórios de classificação com métricas de precisão, revocação e F1-score e acurácia, considerando explicitamente a ordenação das classes em cenários multiclasse.

A análise automatizada de documentos administrativos, mediante técnicas de mineração de texto e algoritmos de classificação e predição, evidencia a qualidade metodológica deste estudo, assim como, caso bem avaliado ratifica a distância existente entre o plano formal das políticas e sua efetivação prática, revelando padrões recorrentes de desarticulação. O emprego do Processamento de Linguagem Natural e de *Machine Learning* possibilita identificar tendências e estruturas latentes, oferecendo subsídios relevantes para a compreensão de dinâmicas institucionais.

4. Resultados e discussão

Acerca dos resultados da análise exploratória inicial, fica evidente nos três gráficos de dispersão lexical que seguem que compara a frequência absoluta dos tokens “transversalidade”, “integrar” e de “setor” na corpora desses documentos. A palavra “transversalidade” e “integrar” aparece em menor volume e de forma descontínua, enquanto “setor” é mais constante, sugerindo que a agência institucional do Estado privilegia núcleos específicos sobre a coordenação intersetorial. (*vide* figura 3, 4 e 5)

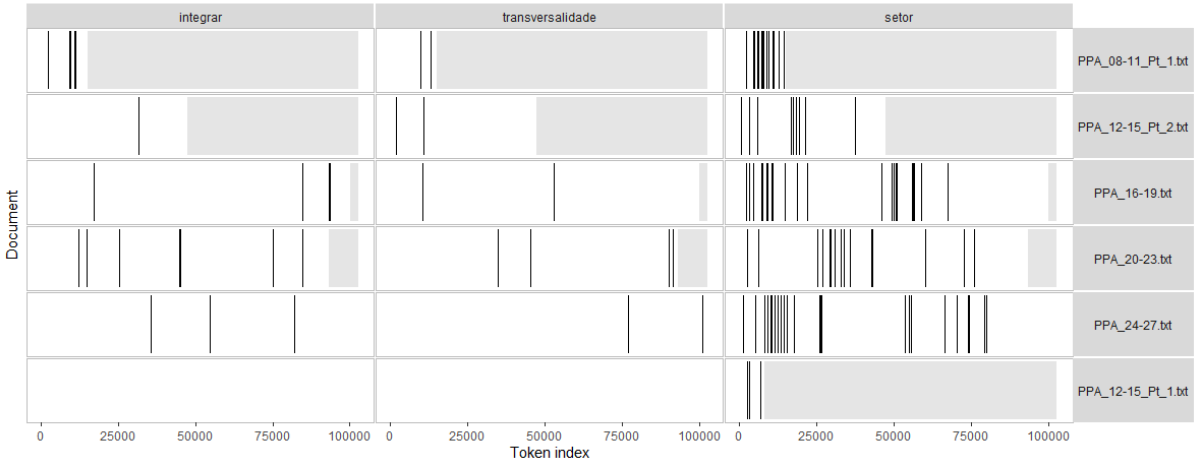
Figura 3. Dispersão lexical nas Leis de Diretrizes Orçamentárias[2]



Elaboração própria. Pires *et al.* (2026)
Fonte: SEPLAG/CEPO/PE

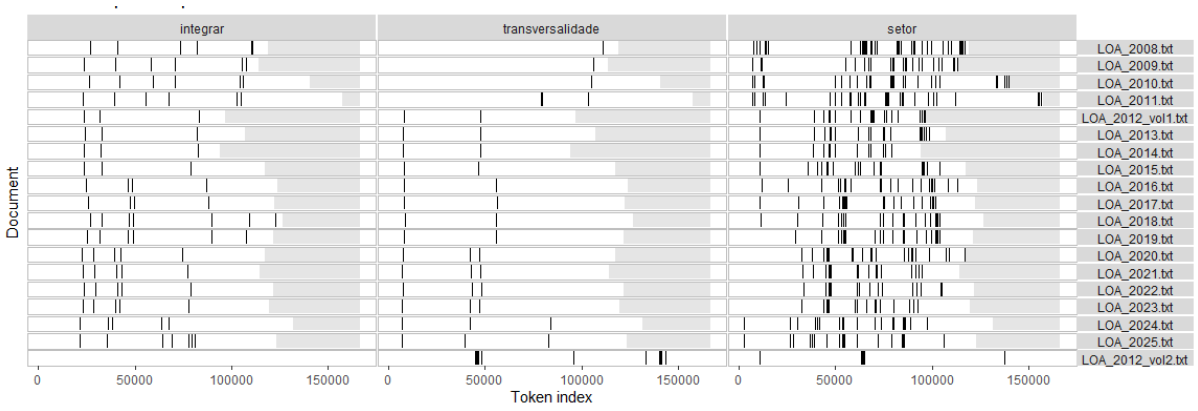
[2] Neste gráfico, não há a dispersão lexical da palavra “integrar", visto que o algoritmo não encontrou evidências desse token no documento oficial em questão.

Figura 4. Dispersão lexical nos Planos Plurianuais de Pernambuco



Elaboração própria. Pires *et al.* (2026)
Fonte: SEPLAG/CEPO/PE

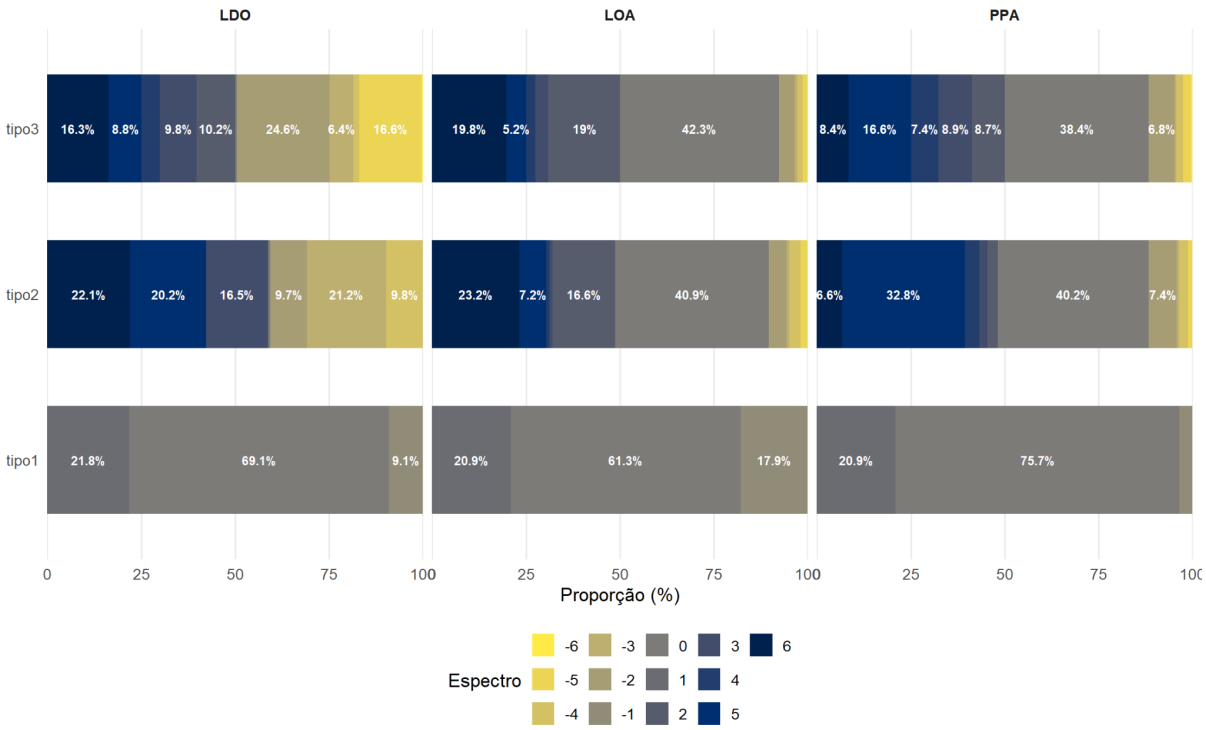
Figura 5. Dispersão lexical nas Leis Orçamentárias Anuais



Elaboração própria. Pires *et al.* (2026)
Fonte: SEPLAG/CEPO/PE

A distribuição descontínua da presença do termo sugere que sua menção é pontual e, em muitos casos, desvinculada de uma estratégia institucional sistemática. Esse uso impede que a transversalidade deixe de ser meramente retórica e se torne um princípio operacional recorrente. Se o uso do termo em documentos oficiais de planejamento é pouco, o que será que realmente acontece na prática efetiva da política? Perceba que dentre tantos tokens a frequência da palavra “transversalidade” é ínfima. Esses achados, ainda provisórios e passíveis de teste empírico adjacentes a esse estudo, apontam para uma fragmentação discursiva entre os níveis de planejamento, com baixa coesão semântica que pode dificultar a adoção efetiva da transversalidade como prática governamental. Contudo, se uma política for projetada de modo transversal, há indícios de que ela só tem a ganhar em eficiência e articulação. Ainda na análise exploratória de dados, a sistematização a seguir (*vide* Figura 6) evidencia a proporção de categorias dos parágrafos dos textos oficiais dado os três tipos de rotulação.

Figura 6. Proporção das categorias em relação à presença de evidências de transversalidade nos documentos oficiais de planejamento do Estado



Elaboração própria. Pires *et al*. (2026)
Fonte: Dados da pesquisa.

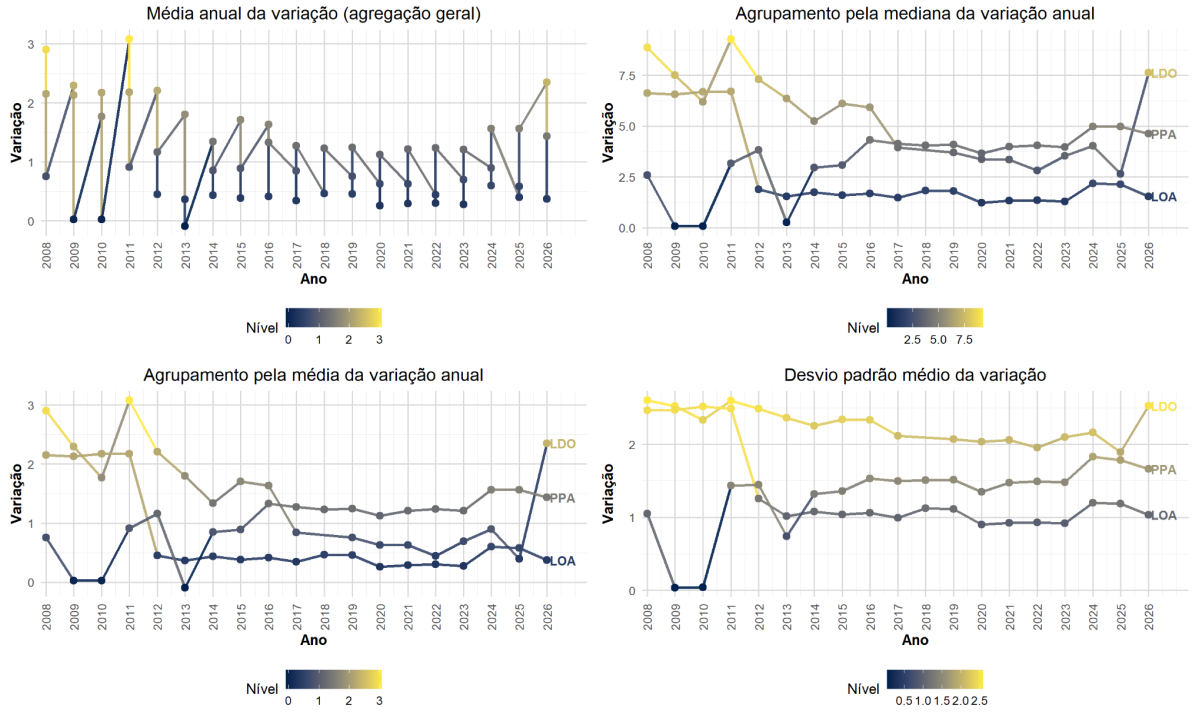
Com uma rotulação mais abrangente, pode-se ver como muitos dados ainda foram rotulados como zero, indicando que nem tudo versa sobre transversalidade ou a similitudes adjacentes à ela, o que não é um problema. A título de exemplo para ilustrar essa situação, os Planos Plurianuais são altamente utilizados para dar uma situação diagnóstica da infraestrutura física do Estado. Ainda que o mesmo possa também evidenciar problemáticas e questões a serem resolvidas em tempos futuros, ou em outros casos mostra o que e como o estado tem feito e/ou implementado determinadas políticas, muitas metas são traçadas nesse e em outros instrumentos de planejamento, o que pode justificar o porquê que o espectro positivo se sobressai nos resultados da Figura 6.

Em seguida, o interessante é que pode-se observar nitidamente como o espectro positivo se destaca ao comparar com o espectro negativo, sugerindo uma contraposição a análise da contagem absoluta de tokens na dispersão lexical pretérita ao contrastar a palavra "transversalidade" com outra que indica possibilidade semântica inversa, nas nuances os textos dos documentos oficiais mostram-se muito mais a favor da transversalidade do que na análise dos tokens em nível de palavra propriamente dita.

Entretanto, para além disso, como será que a presença média transversalidade variou em função do tempo nesses documentos oficiais de planejamento? Ora, se nossa unidade de análise é o parágrafo e detém-se de 3 tipos de rotulação, torna-se possível de minerar algumas estatísticas descritivas como média, mediana e desvio padrão e agrupar por ano e documento para observar a tendência média da série histórica do objeto de análise do presente estudo. Esse dado pode ser evidenciado na Figura 7.

Na agregação geral desconsidera-se a origem do parágrafo e para o cálculo usa-se da mediana para ponderar uma medida última entre os três tipos de rotulações, e tira-se a média por ano desse resultado. Subsequentemente, faz-se o mesmo procedimento, mas dessa vez considera-se a origem do dado. Em sequência, tira a média da soma das observações dos três tipos de rótulos por ano agrupado por documento e seu respectivo desvio padrão, a fim de estabelecer uma relação entre eles. De forma adjacente, é possível notar que os dados da Lei de Diretrizes Orçamentárias variam muito, ao inverso das Leis Orçamentárias Anuais que têm padrões mais consistente como evidencia o desvio padrão dessa medida, mas essa questão pode-se ser justificada dada a densidade de cada tipo de documento. Se fossemos colocar essa questão em um inequação encadeada pode-se ratificar que LOA ≻ PPA ≻ LDO e esse padrão pode tendenciar muitas das análises.

Figura 7. Variação da Transversalidade em função do tempo (*Index*) por documento de organização do Estado (2008-2026)



Elaboração própria. Pires *et al*. (2026)
Fonte: Dados da pesquisa.

Já sobre os resultados da predição de classes do machine learning, eles indicam um desempenho superior do modelo *Linear SVC* (*vide* Tebela 2) em comparação ao *Naive Bayes* (*vide* Tabela 3) em todas os três tipos analisadss, mesmo após o balanceamento das classes. No caso dos rótulos Tipo 1, o *Linear SVC* alcança acurácia de 95%, com valores elevados e estáveis de precisão, *recall* e *F1-score*. Esse padrão sugere que o modelo é particularmente eficaz na identificação de categorias semanticamente mais frequentes e estruturalmente mais homogêneas no *corpus*. Em contraste, o *Naive Bayes* apresenta desempenho inferior, com acurácia global de 79% no Tipo 1, refletindo limitações associadas à hipótese de independência condicional entre atributos, especialmente em contextos textuais complexos.

Nas subtarefas de rótulos Tipo 2 e Tipo 3, observa-se uma queda substancial de desempenho em ambos os modelos, evidenciada por baixos valores de *F1-score* macro e

ponderado. Ainda assim, o *Linear SVC* mantém vantagem relativa, superando sistematicamente o *Naive Bayes* em termos de *F1-score* médio. Esse resultado sugere que, à medida que a tarefa de classificação se torna mais fina e semanticamente exigente, a separabilidade linear entre as classes diminui, tornando o problema mais sensível a ruído, ambiguidade semântica e sobreposição entre categorias adjacentes do espectro.

As classes extremas (-6, -5, +5 e +6) que em tese indicam hiperpositividade e hipernegatividade apresentam, de modo geral, menores valores de recall, indicando maior dificuldade dos modelos em recuperar corretamente observações menos frequentes ou semanticamente mais dispersas, mesmo após procedimentos de balanceamento. De forma geral, os resultados reforçam a adequação do *Linear SVC* como modelo principal para o problema proposto, tanto por seu desempenho superior quanto por sua maior robustez frente à alta dimensionalidade típica de representações vetoriais de texto. O *Naive Bayes*, embora computacionalmente eficiente e conceitualmente simples, mostra-se menos capaz de capturar padrões discriminativos complexos presentes no corpus, sobretudo nas subtarefas de maior granularidade classificatória.

Tabela 2. Resultados do modelo Linear SVC com dados balanceados por categoria

	P (%)			Re (%)			F1-Score (%)			Supp		
Class	TP1	TP2	TP3	TP1	TP2	TP3	TP1	TP2	TP3	TP1	TP2	TP3
- 6	-	-	0.38	-	-	0.74	-	-	0.50	-	-	61
- 5	-	0.44	0.06	-	0.34	0.06	-	0.38	0.06	-	731	63
- 4	-	0.31	0.09	-	0.37	0.08	-	0.33	0.09	-	703	72
- 3	-	0.56	0.33	-	0.55	0.26	-	0.55	0.29	-	738	76
- 2	-	0.27	0.14	-	0.17	0.11	-	0.21	0.12	-	702	63
- 1	0.95	-	-	0.96	-	-	0.95	-	-	6446	-	-
0	0.94	0.20	0.16	0.92	0.10	0.15	0.93	0.14	0.16	6605	738	60
+ 1	0.97	-	-	0.98	-	-	0.98	-	-	6641	-	-
+ 2	-	0.32	0.09	-	0.26	0.13	-	0.29	0.11	-	729	52
+ 3	-	0.49	0.17	-	0.45	0.15	-	0.47	0.16	-	740	62
+ 4	-	0.31	0.06	-	0.46	0.05	-	0.37	0.05	-	731	65
+ 5	-	0.25	0.15	-	0.26	0.07	-	0.25	0.10	-	756	69
+ 6	-	0.42	0.19	-	0.68	0.21	-	0.51	0.20	-	747	61
ACC (%)							0.95	0.36	0.18	19692	7315	704
mAvg (%)							0.95	0.35	0.17			
wAvg (%)							0.95	0.35	0.17			

Elaboração própria. Pires *et al.* (2026)
Fonte: Dados da pesquisa.

Tabela 3. Resultados do modelo Naive Bayes com dados balanceados por categoria

	P (%)			Re (%)			F1-Score (%)			Supp		
Class	TP1	TP2	TP3	TP1	TP2	TP3	TP1	TP2	TP3	TP1	TP2	TP3
- 6	-	-	0.31	-	-	0.72	-	-	0.43	-	-	64
- 5	-	0.38	0.11	-	0.22	0.06	-	0.28	0.08	-	731	64
- 4	-	0.22	0.12	-	0.38	0.19	-	0.28	0.14	-	731	64
- 3	-	0.54	0.24	-	0.53	0.33	-	0.53	0.28	-	732	64
- 2	-	0.18	0.28	-	0.04	0.12	-	0.07	0.17	-	732	64
- 1	0.81	-	-	0.84	-	-	0.82	-	-	6564	-	-
0	0.70	0.16	0.17	0.66	0.15	0.19	0.68	0.15	0.18	6564	732	64
+ 1	0.84	-	-	0.87	-	-	0.85	-	-	6564	-	-
+ 2	-	0.32	0.15	-	0.17	0.11	-	0.22	0.13	-	731	64
+ 3	-	0.46	0.05	-	0.38	0.05	-	0.41	0.05	-	732	64
+ 4	-	0.26	0.17	-	0.41	0.08	-	0.32	0.11	-	731	64
+ 5	-	0.20	0.14	-	0.26	0.08	-	0.22	0.10	-	731	64
+ 6	-	0.38	0.31	-	0.53	0.23	-	0.44	0.27	-	732	64
ACC (%)							0.79	0.31	0.20	19692	7315	704
mAvg (%)							0.78	0.29	0.17			
wAvg (%)							0.78	0.29	0.17			

Elaboração própria. Pires *et al.* (2026)
Fonte: Dados da pesquisa.

5. Considerações finais e limitações

À luz dos resultados apresentados, este trabalho reconhece que a estratégia metodológica adotada, embora adequada ao seu propósito exploratório, revela tanto avanços quanto limites substantivos ao objeto de estudo. Ao mobilizar técnicas clássicas de Processamento de Linguagem Natural e algoritmos de *machine learning* supervisionado, foi possível captar sinais empíricos da presença (ainda que limitada, irregular e por vezes difusa)

da transversalidade nos documentos oficiais de planejamento de Pernambuco. Esses achados dialogam diretamente com a literatura que aponta os entraves estruturais à coordenação intersetorial no âmbito da administração pública, ao mesmo tempo em que sugerem a existência de nuances positivas, ainda incipientes, na linguagem administrativa estadual.

Entretanto, assumir esses resultados de forma acrítica seria negligenciar as próprias condições de produção da evidência empírica. A impossibilidade de processar imagens, tabelas e estruturas gráficas, elementos centrais em documentos orçamentários e de planejamento impôs um recorte analítico restrito ao texto linear. Ela carrega implicações substantivas, pois parte relevante da racionalidade governamental pode se materializar justamente em quadros, metas, indicadores e arranjos visuais que escapam à análise textual tradicional.

No mesmo sentido, este estudo assume que a utilização dos algoritmos *Naive Bayes* e *LinearSVC*, embora consolidados na literatura e computacionalmente eficientes, pode ter se mostrado insuficiente (ou mesmo obsoleta) diante da complexidade semântica da subtarefa de rotulação envolvida. A transversalidade, enquanto conceito normativo, relacional e abstrato, resiste a fronteiras rígidas de classificação. Modelos lineares, baseados em pressupostos de independência condicional ou separabilidade linear, tendem a capturar padrões superficiais da linguagem, mas apresentam dificuldades para apreender camadas mais profundas de significado, ambiguidade e contexto.

Diante disso, os próprios resultados do trabalho funcionam como mote metodológico a pesquisas futuras. Sugere-se que nelas avancem para o uso de *Large Language Models* (LLMs), munidos de estratégias contemporâneas como *fine tuning* e *Retrieval Augmented Generation* (RAG). Essas abordagens representam o estado da arte em NLP ao combinarem compreensão semântica mais sofisticada, adaptação a domínios específicos e incorporação dinâmica de bases externas de conhecimento. Diferentemente dos classificadores tradicionais, LLMs ajustados podem operar não apenas como rotuladores, mas como intérpretes contextuais, capazes de dialogar com a polissemia e a natureza relacional dos conceitos analisados. O *fine tuning*, em particular, desponta como estratégia promissora para refinar modelos genéricos a partir de exemplos cuidadosamente rotulados, reduzindo o ruído classificatório e aumentando a sensibilidade conceitual. Já o uso de RAG permitiria integrar os próprios documentos normativos, legislações e marcos institucionais como fontes externas de recuperação, diminuindo riscos de alucinação e ampliando a aderência empírica das análises.

Do ponto de vista normativo, permanece o pressuposto central que orienta este trabalho: realidades sociais complexas exigem respostas integradas. A transversalidade não pode ser reduzida a um jargão administrativo ou a um marcador retórico ocasional; ela constitui uma condição política e institucional para o enfrentamento das desigualdades. Políticas que negligenciam essa dimensão tendem a reproduzir, sob novas roupagens, os mesmos mecanismos de fragmentação e exclusão que afirmam combater.

Assim, ao mesmo tempo em que este estudo entrega evidências empíricas e metodológicas, ele também deixa uma agenda aberta. Avançar nessa agenda implica não apenas reestruturar dicionários temáticos ou adotar modelos mais recentes, mas fomentar uma cultura organizacional orientada à cooperação intersetorial, apoiada por formação técnica robusta e por sistemas contínuos de monitoramento e avaliação. Em última instância, trata-se de alinhar método, tecnologia e compromisso normativo para que a análise da linguagem do Estado contribua, de fato, para compreender e tensionar as possibilidades de uma gestão pública comprometida com a justiça social e com a tessitura do social em sua complexidade.

Agradecimentos

Agradecemos a todos do Workshop de Análise de Dados pelos comentários tecidos a esse trabalho independente. Suas considerações críticas condicionaram esse trabalho a melhorias não captadas pela óptica dos autores, e permitiu ele ser lapidado na forma que está hoje.

Declaração de Conflito de Interesses

Os autores declaram não haver conflitos de interesse. Todos os coautores leram e concordam com o conteúdo do manuscrito, e não há qualquer interesse financeiro a ser declarado.

Anexos

Os scripts utilizados no presente estudo estão publicamente disponíveis no repositório do autor correspondente no GitHub (Cf. Referências). Em razão de limitações de armazenamento e do grande volume dos arquivos, os corpora completos analisados neste artigo não puderam ser hospedados em repositórios públicos. Pesquisadores interessados em acessar integralmente os dados para fins de replicação ou análises adicionais podem entrar em contato com o autor correspondente, mediante condições razoáveis de uso e compartilhamento.

Referências

1. Avelino, D. P., & Santos, J. C. (2014, março). O desafio do Fórum Interconselhos na consolidação das estruturas participativas de segundo nível. In *Anais do 7º Congresso Consad de Gestão Pública*. Brasília, DF.
2. Benoit K, Watanabe K, Wang H, Nulty P, Obeng A, Müller S, Matsuo A (2018). "quanteda: Um pacote R para a análise quantitativa de dados textuais." *Journal of Open Source Software* , 3 (30), 774. doi:10.21105/joss.00774 , https://quanteda.io
3. Bird, Steven; Klein, Ewan. *Natural Language Processing with Python: analyzing text*
4. Brasil. Ministério do Planejamento, Orçamento e Gestão. (2003). *Plano Plurianual 2004–2007: Desenvolvimento com inclusão social*. Brasília, DF: Ministério do Planejamento, Orçamento e Gestão. Disponível em https://www.gov.br/planejamento/pt-br/assuntos/planejamento/plano-plurianual/paginas/ppa-2004-2007
5. Dryzek, J. S. (1990). Complexity. In J. S. Dryzek, *Discursive democracy: Politics, policy, and political science* (pp. 59–74). Cambridge: Cambridge University Press.
6. Géron, A. (2019). *Hands-On machine learning with Scikit-Learn and TensorFlow* (R. Contatori, Trad.). Rio de Janeiro: Alta Books.
7. Goodin, R. E. (2006). Governança em rede. In R. E. Goodin, *The Oxford handbook of public policy* (pp. 204–222). Oxford: Oxford University Press.
8. Goodin, R. E. (2006). Talking politics: Perils and promises. *European Journal of Political Research, 45*(2), 235–261. https://doi.org/10.1111/j.1475-6765.2006.00298.x
9. Grimmer, J., & Stewart, B. M. (2013). *Text as data: The promise and pitfalls of automatic content analysis methods for political texts.* Political Analysis, 21(3), 267–297. https://doi.org/10.1093/pan/mps028
10. Harris, C.R., Millman, K.J., van der Walt, S.J. et al. *Array programming with NumPy*. Nature 585, 357–362 (2020). DOI: 10.1038/s41586-020-2649-2.

11. Izbicki, R., & Santos, T. M. dos. (2017). *Machine learning sob a ótica estatística: Uma abordagem preditivista para a estatística com exemplos em R* [Versão preliminar em preparação]. São Carlos: UFSCar; São Paulo: Insper. Disponível em https://www.est.ufmg.br/~marcosop/est171-ML/MachineLearning_Izbicki.pdf

12. Jurafsky, D., & Martin, J. H. (2018, setembro 23). *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition* (Draft of 3rd ed., Cap. 4). Disponível em https://web.stanford.edu/~jurafsky/slp3/4.pdf

13. Janssen, M., Kuk, G., & Wagenaar, R. W. (2020). *Digital transformation and public policy: The role of information, data and algorithms.* Government Information Quarterly, 37(4), 101–117. https://doi.org/10.1016/j.giq.2020.101493

14. Klijn, E. H., & Koppenjan, J. F. M. (2000). Public management and policy networks: Foundations of a network approach to governance. *Public Management Review, 2*(2), 135–158. https://doi.org/10.1080/14719030000000009

15. Laver, M., Benoit, K., & Garry, J. (2003). *Extracting policy positions from political texts using words as data.* American Political Science Review, 97(2), 311–331. https://doi.org/10.1017/S0003055403000698

16. Lavezzo, G. H., & da Conceição, G. C. . (2023). *Análise de Tópicos em Redes Sociais Utilizando Processamento de Linguagem Natural (NLP).* Revista Interface Tecnológica, 20(2), 65-76. https://doi.org/10.31510/infa.v20i2.1750

17. Lotta, G. (2019a). A política pública como ela é: Contribuições dos estudos sobre implementação para a análise de políticas públicas. In *Teorias e análises sobre implementação de políticas públicas no Brasil* (pp. 11–38). Brasília: Enap.

18. Lotta, G. (2019b). Capacidades institucionais e burocracia. In G. Lotta, *Teorias e análises sobre implementação de políticas públicas no Brasil* (pp. 98–123). São Paulo: Letra e Voz.

19. Nogueira, C. A. G., & Forte, S. H. A. C. (2019). Efeitos intersetoriais e transversais e seus impactos sobre a efetividade das políticas públicas nos municípios do Ceará. *Revista de Administração Pública, 53*(6), 1189–1208. https://doi.org/10.1590/0034-761220180167

20. Ooms J (2025). *pdftools: Extração de texto, renderização e conversão de documentos PDF* . Pacote R versão 3.7.0, https://ropensci.r-universe.dev/pdftools

21. Ostrom, E. (2005). *Understanding institutional diversity* (p. 283). Princeton: Princeton University Press.

22. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). *Scikit-learn: Machine learning in Python*. *Journal of Machine Learning Research, 12*, 2825–2830.

23. Peters, B. G. (1998). Managing horizontal government: The politics of coordination. *Public Administration, 76*(2), 295–311. https://doi.org/10.1111/1467-9299.00102

24. Peters, B. G. (2015). Governança multinível. In B. G. Peters, *Advanced introduction to public policy* (pp. 87–103). Cheltenham: Edward Elgar Publishing.

25. Peters, B. G., & Pierre, J. (2004). Politicization of the civil service: Concepts, causes, consequences. In B. G. Peters & J. Pierre, *Politicization of the civil service in comparative perspective: The quest for control* (pp. 1–11). London: Routledge.

26. Pires, R. (2026). *Scripts de replicação para Evidências sobre o uso de Processamento de Linguagem Natural e Machine Learning como ferramenta diagnóstica à transversalidade em Pernambuco.* [Repositório código-fonte]. GitHub. https://github.com/ryallmeida/transversalidade

27. Pollitt, C. (2003). Joined-up government: A survey. *Political Studies Review, 1*(1), 34–49. https://doi.org/10.1111/1478-9299.00003

28. Rua, M. G. (2009). Coordenação intersetorial. In M. G. Rua, *Análise de políticas públicas: Conceitos básicos* (pp. 7–20). Rio de Janeiro: Editora FGV.

29. Sabatier, P. A., & Mazmanian, D. (1980). The implementation of public policy: A framework of analysis. *Policy Studies Journal, 8*(4), 538–560. https://doi.org/10.1111/j.1541-0072.1980.tb01266.x

30. Silge J, Robinson D (2016). *"tidytext: Mineração e análise de texto usando princípios de dados organizados em R." JOSS* , 1 (3). doi:10.21105/joss.00037 , http://dx.doi.org/10.21105/joss.00037 .

31. Silva, T. D., & Calmon, P. C. P. (2017, novembro 14–17). Transversalidade e políticas públicas. In *Anais do XXII Congreso Internacional del CLAD sobre la Reforma del Estado y de la Administración Pública*. Madrid, Espanha: CLAD.

32. Simon Garnier, Noam Ross, Robert Rudis, Antônio P. Camargo, Marco Sciaini e Cédric Scherer (2024). viridis (Lite) - Mapas de cores compatíveis com daltônicos para o pacote R. viridis versão 0.6.5.

33. The pandas development team. (2020). *pandas-dev/pandas: Pandas* (Version 2.3.3) [Software]. Zenodo. https://doi.org/10.5281/zenodo.3509134

34. Wickham, H., Pedersen, T. L., Seidel, D., & Posit Software, PBC. (2025). *scales: Scale functions for visualization* (Version 1.4.0) [Computer software]. https://doi.org/10.32614/CRAN.package.scales

35. Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D. A., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K., & Yutani, H. (2019). *Welcome to the tidyverse*. Journal of Open Source Software, 4(43), 1686. https://doi.org/10.21105/joss.01686