

OS REFERENCIAIS CURRICULARES ESTADUAIS PARA EDUCAÇÃO INFANTIL E ENSINO FUNDAMENTAL ALINHADOS À BNCC: AVALIAÇÃO DA PRESENÇA DA PARTE DIVERSIFICADA POR MEIO DE MODELAGEM DE TÓPICOS

<https://doi.org/10.5902/2318133887052>

Hugo Bovareto de Oliveira Horsth¹
Thiago Duarte Pimentel²

Resumo

Após aprovação da BNCC, os governos estaduais brasileiros desenvolveram seus referenciais curriculares para educação infantil e ensino fundamental. Esses documentos deveriam apresentar uma parte diversificada, que contextualizaria os saberes da Base e acrescentaria características importantes para cada Estado. Por meio dessa pesquisa, buscou-se avaliar a presença da parte diversificada nos documentos estaduais. Para tanto, realizou-se análise de conteúdo dos 27 documentos estaduais, pela modelagem de tópicos utilizando a técnica Latent Dirichlet Allocation - LDA. Os resultados demonstram que os currículos estaduais apresentam aspectos importantes da BNCC, mas a parte diversificada não aparece em nenhum tópico. Conclui-se que, em nenhum Estado, a parte diversificada é significativamente presente a ponto dessas características estarem presentes num dos tópicos.

Palavras-chave: BNCC; currículo; LDA; modelagem de tópicos; análise de conteúdo.

THE STATE CURRICULA FOR EARLY CHILDHOOD EDUCATION AND PRIMARY EDUCATION ALIGNED WITH THE BNCC: EVALUATION OF THE PRESENCE OF THE DIVERSIFIED PART USING TOPIC MODELING

Abstract

Following the approval of the BNCC, States developed their curricular references for Early Childhood Education and Primary Education. These documents were supposed to include a diversified part that would contextualize the BNCC content and add important regional characteristics. This paper aimed to evaluate the presence of the diversified part in the state documents. For this purpose, content analysis of the 27 documents using topic modeling with the Latent Dirichlet Allocation – LDA – technique was used. The results show that the state curricula present important aspects of the BNCC, but the diversified part does not appear in any topic. We conclude that the diversified part is not significantly present in any State to the point that these characteristics are present in one of the topics.

Key-words: BNCC; curriculum; LDA; topic modeling; content analysis.

¹ Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais, Brasil. E-mail: hugobovareto@gmail.com. Orcid: <https://orcid.org/0009-0009-0675-5686>.

² Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais, Brasil. E-mail: thiago.pimentel@ich.ufjf.br. Orcid: <https://orcid.org/0000-0003-1889-069X>.

Crítérios de autoria: os autores, coletivamente, realizaram a concepção, criação e consolidação do artigo.

Recebido em 11 de março de 2024. Aceito em 2 de maio de 2024.



Introdução

Dentre muitas possibilidades, é necessária uma seleção de conhecimentos do que será ensinado nas escolas e, conseqüentemente, o que ficará de fora. Esse processo resulta no currículo escolar (Silva, 2010, p. 15). Esses documentos refletem os desejos de aprendizagem dos sistemas de ensino e podem ser mais ou menos prescritivos. Em alguns casos há somente indicação do conteúdo que deve ser trabalhado, mas também pode haver propostas de metodologias (Sacristán, 2000). A definição do que constará nesses currículos é objeto de grandes disputas entre grupos de interesse (Goodson, 2011).

A Constituição Federal de 1988, já abordava a necessidade de fixação de conteúdos mínimos para ensino fundamental. A Base Nacional Comum Curricular – BNCC –, foi homologada em 2017 depois de amplo período de discussões e consultas públicas com o objetivo de ser o documento referência com direitos de aprendizagem mínimos para que os sistemas de ensino pudessem construir seus documentos curriculares.

Em 2018, pela portaria n. 331, o MEC instituiu o Programa de Apoio à Implementação da Base Nacional Comum Curricular – ProBNCC. Até 2019, os Estados receberam suporte técnico e financeiro para o desenvolvimento e aprovação de seus referenciais curriculares para educação infantil e ensino fundamental alinhados à BNCC. Assim como consta na LDB, os documentos curriculares deveriam conter os direitos de aprendizagem estabelecidos na BNCC, mas também deveria haver uma contextualização dos conhecimentos representada por uma chamada parte diversificada que abarcaria as características, especificidades e necessidades regionais de cada ente federativo.

Apesar do ProBNCC ter fornecido orientações nacionais em conjunto para todos os entes federativos, os Estados possuíam autonomia para as incorporações que fariam na parte diversificada dos documentos curriculares. Para além disso, a aprovação desses documentos foi feita por cada respectivo conselho estadual de educação nos anos de 2018 e 2019, o que gera mais espaço para heterogeneidade nos materiais. Diante desse cenário o presente trabalho buscou avaliar a presença da chamada parte diversificada nos 27 documentos curriculares estaduais através do uso da técnica Latent Dirichlet Allocation – LDA –, para modelagem de tópicos.

Histórico da legislação curricular para educação básica

Após aprovação da Constituição Federal de 1988, novos debates sobre estrutura e papel do Estado começaram a acontecer. Inserida nesses debates, a educação passou a ser vista como bem e direito público (Aguiar; Tuttman, 2020).

A LDB previu, no artigo 26, que “os currículos do ensino fundamental e médio devem ter uma base nacional comum”. Em 1997, 1998 e 2000 foram publicados os parâmetros curriculares nacionais para os anos iniciais do ensino fundamental (1º ao 5º ano), anos finais (6º ao 9º ano) e ensino médio respectivamente. Entre 2008 e 2010 aconteceu o Programa Currículo em Movimento, que buscava melhorias através do desenvolvimento de currículo para todas as etapas da educação básica (Araujo; Miranda, 2020). Nesse período, no final de 2009, o CNE publicou a resolução n. 5, que fixava as diretrizes curriculares para educação infantil.

A Conae, no documento final de 2010, expôs a necessidade de uma base nacional curricular no PNE (Brasil, 2010). No mesmo ano, o CNE publicou a resolução n. 4, que definia as diretrizes curriculares nacionais gerais para a educação básica. A resolução n. 7, do mesmo ano, estabeleceu as DCNs para Ensino Fundamental e em 2012 as DCNs para ensino médio foram definidas pela resolução n. 2 do CNE.

Em 2014, a lei n. 13.005 aprovou o Plano Nacional de Educação – PNE –, o para o decênio de 2014 a 2024 e estabeleceu que entes públicos e sociedade civil deveriam atuar conjuntamente para desenvolvimento, implementação e monitoramento das políticas públicas educacionais. Dentre as 20 metas do PNE, as metas 2, 3, 7 e 15 citavam uma base comum curricular como estratégia para cumprimento dos objetivos e meta 1 referenciava as DCNs para educação infantil.

A materialização da BNCC, como documento que atenderia a necessidade de uma base curricular nacional, começou a acontecer em 2015 com a portaria n. 592 do MEC. Junto com entidades representativas da educação, Conselho Nacional de Secretários de Educação e União Nacional dos Dirigentes Municipais de Educação, foi instituída a Comissão de Especialistas para a Elaboração de Proposta da Base Nacional Comum Curricular. Em setembro do mesmo ano a primeira versão do documento foi disponibilizada e aberta para contribuições em uma consulta pública no mês seguinte.

Após mudanças advindas das sugestões das consultas públicas, a segunda versão foi divulgada em maio de 2016. Seminários estaduais foram realizados entre junho e agosto para discutir a versão disponibilizada. Em abril de 2017, a terceira versão foi entregue pelo MEC ao CNE com as partes referentes à educação infantil e ensino fundamental. Após coleta de opiniões através de consultas públicas, o CNE aprovou a BNCC com vinte votos favoráveis e três contrários em dezembro. No mesmo mês, a portaria n. 1.570 do MEC e a resolução n. 2 do CNE instituíram a implantação da BNCC, “a ser respeitada obrigatoriamente”. O processo de implementação começou em 2018 guiado pelo ProBNCC.

Base Nacional Comum Curricular

A versão final aprovada em dezembro de 2017 não se propõe a ser um modelo para os currículos dos entes subnacionais. O objetivo é assegurar os aprendizados essenciais dos alunos:

Portanto, os agrupamentos propostos não devem ser tomados como modelo obrigatório para o desenho dos currículos. Essa forma de apresentação adotada na BNCC tem por objetivo assegurar a clareza, a precisão e a explicitação do que se espera que todos os alunos aprendam no Ensino Fundamental, fornecendo orientações para a elaboração de currículos em todo o País, adequados aos diferentes contextos. (Brasil, 2017, p. 31)

O documento se organiza em competências e habilidades. Para toda educação básica são apresentadas dez competências gerais. A educação infantil apresenta seis direitos de aprendizagem e desenvolvimento e o ensino fundamental é em áreas do conhecimento e componentes com competências específicas para cada um deles (Brasil, 2017).

Desde o início do processo de discussão sobre o documento, debates e posições antagônicas já existiam. As divergências passavam por conteúdo, forma e até pela necessidade de um base curricular comum nacionalmente. O lado favorável defendia que, independente das diferenças regionais, existem saberes comuns a todo o país e que devem ser assegurados para todos os estudantes do território nacional. Concomitantemente, os críticos defendiam que um documento comum nacional anularia as diversidades locais e impositivamente manteria conhecimentos dominantes e hegemônicos (Caetano, 2023).

Não há consenso na literatura se a homologação da BNCC foi um avanço ou retrocesso. Para Filipe, Silva e Costa (2021), a BNCC é tecnicista e com foco somente na empregabilidade, sem se preocupar com uma educação emancipatória. Para Chaves (2021), há alto grau de detalhamento nas competências e habilidade, o que retira a autonomia dos professores. Do outro lado, Marcelo Burgos defende que a base, ao garantir direitos de aprendizagem, funciona como um importante instrumento para redução das desigualdades e que favorece as diversidades nos currículos estaduais (Cravo, 2018).

O Centro de Políticas Públicas e Avaliação da Educação da Universidade Federal de Juiz de Fora, realizou em 2021 e 2022 a aplicação de questionários de monitoramento da implementação da base para educação infantil e ensino fundamental. Os respondentes foram professores, coordenadores pedagógicos, diretores de escolas e técnicos das secretarias municipais e estaduais de educação. Para todos os perfis, foi identificada uma percepção positiva dos efeitos da BNCC sobre gestão educacional e uso central do currículo nas práticas pedagógicas (Caed/UFJF, 2022).

Metodologia

Para atendimento do objetivo proposto nesse estudo, foi realizada uma análise de conteúdo exploratória dos 27 referenciais curriculares estaduais alinhados à BNCC para educação infantil e ensino fundamental. Foi utilizado a técnica LDA. Esta é uma técnica de processamento de linguagem natural – PLN –, baseada em fundamentos estatísticos para quantificar a frequência de termos importantes nos documentos e identificar os tópicos mais relevantes em cada documento (Blei; Ng; Jordan, 2003).

Somados, os 27 documentos possuem 21.670 páginas, o que tornaria uma análise de conteúdo por meios tradicionais de leitura de todos os documentos muito custosa. O uso da técnica LDA, por meio de inteligência artificial e processamento computacional, buscou reduzir em muito o tempo e recursos gastos. Os currículos estaduais foram acessados através de links públicos na internet. O MEC disponibiliza 26 documentos³, com exceção da Bahia, porém é possível encontrar os materiais mais atualizados nos sites das secretarias e conselhos estaduais de educação. Todos os documentos utilizados foram organizados e disponibilizados em repositório público⁴ para acesso de qualquer pessoa⁴.

³ O MEC disponibilizou os documentos no link http://basenacionalcomum.mec.gov.br/index.php?option=com_content&view=article&layout=edit&id=207

⁴ Os documentos podem ser acessados neste repositório: <https://drive.google.com/drive/folders/1I4xMtM2Kkz1s2DJH3UhSbuO8eH1fbHDs?usp=sharing>

Análise de conteúdo

Segundo Berelson (1958), análise de conteúdo pode ser definida como “uma técnica de pesquisa para a descrição objetiva, sistemática e quantitativa do conteúdo manifestado de comunicação” (p. 18). Por meio de textos, pesquisadores buscam compreender fenômenos sociais. O objetivo é resumir e descrever o conteúdo de um conjunto de textos relacionados ao fenômeno pesquisado (Krippendorff, 2004, p. 8).

Tradicionalmente, o estudo pode ocorrer através da análise completa dos documentos ou por amostragens aleatórias ou intencionais (Morgan, 2022). A clara desvantagem para esses meios é o custo alto de tempo e pessoal envolvido na leitura de textos muito longos ou em alto volume. Nos últimos anos, a fim de diminuir os custos de todo o processo e combinado com as melhorias computacionais, algoritmos e softwares envolvendo inteligência artificial vêm sendo mais utilizados para identificação de padrões e tópicos em dados textuais (Elragal; Klischewski, 2017).

Modelagem de tópicos

Diversos métodos de mineração e busca por dados textuais já foram criados e aplicados (Li et al., 2019). A modelagem de tópicos é uma das técnicas mais frequentemente utilizadas (Hong; Davison, 2010), e cada vez mais vem ganhando popularidade para uso nas Ciências Sociais (Hannigan et al., 2019).

Inicialmente desenvolvida para classificação de pesquisas, livros, notícias e artigos, o objetivo da modelagem é ressaltar os temas importantes no conjunto de textos analisados, chamado também de corpus. Baseado em modelos estatísticos, a técnica utiliza como recurso uma série de documentos textuais e busca identificar padrões de palavras para agrupá-las em um mesmo tópico de acordo com a proximidade delas no plano vetorial (Egger; Yu, 2021).

Dentro da modelagem de tópicos, diversas técnicas diferentes podem ser aplicadas para análises textuais. Cada uma possui vantagens e desvantagens e adequações para as análises que se pretende fazer, considerando objetivo, volume dos documentos e extensão dos textos (Egger; Yu, 2022). Para escolha correta do algoritmo, uma avaliação cuidadosa das características dos documentos deve ser feita (Churchill; Singh, 2022).

Pelas características dos 27 documentos curriculares estaduais e levando em consideração comparação realizada por Egger e Yu (2022), para o presente trabalho foi escolhida a técnica LDA - (Alocação Latente de Dirichlet, em português).

Os pesquisadores David Blei, Andrew Ng e Michael I. Jordan apresentaram o modelo probabilístico generativo LDA em 2003. O modelo parte da premissa que documentos são uma mistura de diferentes tópicos e cada tópico é representado por uma mistura de palavras (Falcão, 2019). A técnica usa correlação estatística entre os documentos e as palavras agrupamento das palavras em conjuntos que representem tópicos. A nomeação de cada tópico é feita pelo pesquisador. Entre os estudos de modelagem de tópicos, a técnica LDA é a mais popular e tida como referência (Jelodar et al., 2018).

Análise dos dados

Os materiais foram organizados para que cada Estado tivesse somente um único arquivo e toda a análise foi feita com códigos desenvolvidos na linguagem de programação Python. Cada documento estadual foi analisado separadamente.

Os textos, inicialmente em formato .pdf foram processados com ajuda de diversas bibliotecas de análise textual: PyPDF2, pdfminer, pdfplumber, PIL e pdf2image. A ferramenta de OCR PyTesseract também foi utilizada para identificação de textos contidos nas imagens dos currículos. Após extração de todos os textos, as páginas foram divididas e organizadas em um base tabular de dados, conhecida como DataFrame, utilizando o pacote Pandas.

Foi feita exclusão de partes consideradas não relevantes para a análise de conteúdo como capas, fichas técnicas, sumários, referências bibliográficas e cartas de apresentação do documento. O primeiro processo para tratamento desses dados foi a tokenização dos textos e separação das linhas do DataFrame em frases. Esse processo divide os dados textuais em partes palavras únicas e é essencial para que os próximos tratamentos e o modelo reconheçam as palavras e realizem as tarefas de maneira correta (Hannigan et al., 2019). Para esse procedimento foi utilizada o pacote NLTK.

Depois da tokenização, ocorreu a normalização dos tokens. Números, pontuações, símbolos, acentos, espaços em branco e qualquer caractere não textual foi excluído. Para além disso, todas as letras foram convertidas para minúsculas (Storopoli, 2019). Também foram retiradas as chamadas stopwords. Essas são palavras sem valor semântico para os documentos, como preposições, pronomes, conjunções e artigos (Berry; Kogan, 2010). Para normalização foi utilizado a biblioteca gensim e a Spacy foi a escolhida para remoção das *stopwords*.

A última etapa de tratamento dos dados foi a stemização. Esse processo consiste em converter cada palavra para sua forma raiz (Hannigan et al., 2019). Para esse caso, como todos os textos estavam em português, foi utilizado o algoritmo RSLP – Removedor de Sufixos da Língua Portuguesa –, do pacote NLTK para a função.

Após todas as etapas de tratamento, a técnica LDA para modelagem dos dados foi aplicada. A escolha do número de tópicos precisa ser feita previamente e não há uma regra para essa definição. Apesar disso alguns indicadores como podem ser utilizados (Dimaggio, 2015). Neste estudo foram considerados valores de coerência, de perplexidade e análise visual do mapa de distância intertópica. Tópicos mais facilmente identificáveis e com temáticas menos sobrepostas geralmente acontecem quanto maior o valor de coerência, menor o de perplexidade e mais distribuídos os tópicos no mapa de distância intertópica.

Com o número de tópicos definidos, foi utilizada a biblioteca Gensim para definir o dicionário de palavras e o corpus foram criados, parâmetros essenciais para o modelo. O modelo utilizado foi o LdaModel e as visualizações dos mapas de distância intertópica foram geradas com a biblioteca PyLDAvis. Outras visualizações para entender as características de cada corpus foram feitas com a biblioteca matplotlib. Todos os códigos e visualizações estão organizados em um repositório público⁵.

⁵ Todos os códigos, visualizações e resultados gerados estão organizados neste repositório: <https://github.com/hugobovareto/LDA-BNCC-e-referenciais-curriculares-estaduais>

Resultados e discussões

Apesar de grandes diferenças em extensão dos documentos, os tópicos e palavras mais relevantes encontradas nos documentos muitas vezes são semelhantes. A tabela abaixo apresenta as temáticas mais identificadas como tópicos no conjunto de análises de todos os documentos.

Tabela 1 –
Tópicos mais frequentes de acordo com modelagem de tópicos com técnica LDA.

Tópicos	Modelagem de tópicos - LDA
Objetos do conhecimento	41
Processos educacionais, pedagógicos e de ensino e aprendizagem	32
Linguagens	29
Currículo	25
Organização curricular	22
Habilidades	18
Ciências Humanas	17
Relações Sociais	15
Ambiente escolar; escola	15
Matemática	11
Competências	10
Localidades	9
Comunicação	8
Povos	8
Cultura	7
Arte	6
Diversidade	5
Método, investigação e desenvolvimento científico	4
Ciências Sociais	4
Sistema de ensino	3

Fonte: autores.

Pelos tópicos identificados, as temáticas mais marcantes nos documentos abordam aspectos gerais de um documento curricular. Temas estruturantes e marcantes na BNCC, também são identificáveis como organização por competências, habilidades e objetos do conhecimento.

Apesar da classificação, os tópicos nomeados como “diversidade” são caracterizados por termos relacionados à temática de maneira geral. Com adição dos sufixos, retirados durante o processo de stemização, palavras encontradas nos tópicos identificados como “diversidade” são: cultura; diferença, social, religião, etnia, direito, diversidade, história, raça, crença, sociedade, tradição, identidade. Esses tópicos não apresentam palavras que remetem às especificidades regionais de cada estado.

Uma possível explicação para a falta de tópicos com aspectos marcantes regionais pode ser o isomorfismo mimético, conhecido e descrito na literatura de Administração. O conceito é caracterizado pela observação e imitação de processos e características de outras organizações, principalmente em cenários rodeados de incertezas (Gimenez; Júnior; Grave, 2005). No ambiente que ocorreram as construções dos documentos curriculares, cercado de disputas e discussões, era esperado que os Estados se voltassem e optassem por opções mais conservadoras e replicassem a estrutura e conteúdo da BNCC, sem grandes acréscimos.

Conclusões

A pesquisa pelos tópicos mais representativos de cada documento permitiu tirar conclusões importantes sobre o desenvolvimento da parte diversificada nos referenciais curriculares estaduais.

Os tópicos mais relevantes dos documentos curriculares estaduais abordam aspectos estruturantes e relevantes da BNCC. A parte geral que deveria ser mantida nos documentos para garantir os direitos de aprendizagem dos estudantes, de acordo com o que está na Base, parece ter sido bem representada nos documentos estaduais quando olhamos para os tópicos. Do outro lado, a parte diversificada, espaço para que os Estados inserissem suas características, necessidades e vontades em seus currículos, não se mostrou marcante suficiente para ser destacada em tópicos de acordo com técnica LDA de modelagem de tópicos.

Segundo a resolução CNE/CP n. 2, de dezembro de 2017, a BNCC deverá ser revisada após cinco anos do prazo de efetivação dos currículos. No artigo 15 a resolução, aponta-se que a adequação dos currículos deveria ser feita preferencialmente em 2019 e no máximo até início de 2020. Considerando esses prazos, uma revisão pode acontecer a partir de 2024. Existe oportunidade para que, em um momento de novas construções curriculares estaduais após a revisão da BNCC, os gestores públicos educacionais dêem mais atenção e relevância para a parte diversificada dos materiais e para que as especificidades e necessidades locais sejam representadas nos referenciais curriculares estaduais.

Referências

AGUIAR, Márcia Angela da S; TUTTMAN, Malvina T. Políticas educacionais no Brasil e a Base Nacional Comum Curricular: disputas de projetos. *Em aberto*, Brasília, v. 33, n. 107, 2020, p. 69-94.

ARAUJO, Daniele Freitas de; MIRANDA, Maria das Graças Titoneli de Miranda. BNCC e revisão dos currículos de escolas de educação infantil em município mineiro. CONGRESSO NACIONAL DE EDUCAÇÃO, 7, 2020. Anais ... Campina Grande: Realize Editora, 2020. Disponível em: <https://editorarealize.com.br/editora/anais/conedu/2020/TRABALHO_EV140_MD1_SA9_ID908_01092020195006.pdf>. Acesso em: 26 fev. 2024.

BERELSON, Bernard. *Content analysis in communication research*. Nova York: Free Press, 1952.

BERRY, Michael W; KOGAN, Jacob. *Text mining: application and theory*. Tenessi: Wiley, 2010.

BLEI, David M.; NG, Andrew Y; JORDAN, Michael I. Latent dirichlet allocation. *Journal of Machine Learning Research*, Cambridge, MA (USA), v. 3, p. 993-1022, 2003.

BRASIL. *Conae 2010: construindo o sistema nacional articulado de educação: o plano Nacional de Educação, diretrizes e estratégias de ação*. Brasília: MEC, 2010. Disponível em: <https://pne.mec.gov.br/images/pdf/CONAE2010_doc_final.pdf>. Acesso em: 25 fev. 2024.

BRASIL. *Resolução n. 2, de 22 de dezembro de 2017: institui e orienta a implantação da Base Nacional Comum Curricular, a ser respeitada obrigatoriamente ao longo das etapas e respectivas modalidades no âmbito da Educação Básica*. Brasília: MEC/CNE, 2017.

BRASIL. *Resolução n. 2, de 30 de janeiro de 2012: define diretrizes curriculares nacionais para o ensino médio*. Brasília: MEC/CNE, 2012.

BRASIL. *Resolução n. 4, de 13 de julho de 2010: define as diretrizes curriculares nacionais gerais para a educação básica*. Brasília: MEC/CNE, 2010.

BRASIL. *Resolução n. 7, de 14 de dezembro de 2010: fixa diretrizes curriculares nacionais para o ensino fundamental de 9 (nove) anos*. Brasília: MEC/CNE, 2010.

BRASIL. *Constituição da República Federativa do Brasil*. Brasília: Centro Gráfico, 1988.

BRASIL. *Lei n. 13.005, de 25 de junho de 2014: aprova o Plano Nacional de Educação - PNE e dá outras providências*. Diário Oficial da União, Brasília, 2014.

BRASIL. *Lei n. 9.394, de 20 de dezembro de 1996: estabelece as diretrizes e bases da educação nacional*. Diário Oficial da União, Brasília, DF, 1996.

BRASIL. *Base Nacional Comum Curricular*. Brasília: Brasília: MEC, 2017. Disponível em: <http://basenacionalcomum.mec.gov.br/images/BNCC_EI_EF_110518_verseofinal_site.pdf>. Acesso em: 26 fev. 2024.

BRASIL. *Portaria n. 331, de 05 de abril de 2018: institui o Programa de Apoio à Implementação da Base Nacional Comum Curricular - ProBNCC e estabelece diretrizes, parâmetros e critérios para sua implementação*. Diário Oficial da União, Brasília, DF, 2018.

BRASIL. *Portaria n. 592, de 17 de junho de 2015: institui Comissão de Especialistas para a Elaboração de Propostas da Base Nacional Comum Curricular*. Diário Oficial da União, Brasília, DF, 2015.

BRASIL. *Portaria n. 1.570, de 20 de dezembro de 2017: homologa a BNCC para educação infantil e ensino fundamental*. Diário Oficial da União, Brasília, 2017.

CAED/UFJF. *Relatório da 2ª aplicação dos instrumentos de pesquisa*. Juiz de Fora: Caed/UFJF, 2022. Disponível em: https://plataformabncc.caeddigital.net/resources/arquivos/BNCC_2022_Relatorio_2_Aplicacao_Instrumentos_Pesquisa.pdf. Acesso em: 25 de fev. 2024.

CAETANO, Maria Raquel. As disputas na construção da Base Nacional Comum Curricular: anotações em torno do conteúdo da política educacional. *Olhar de professor*, Ponta Grossa, v. 26, 2023, p. 1-24, 2023.

CHAVES, Priscila Monteiro. Uma base para a semiformação socializada: a vulgarização da crítica como estratégia de produção do consenso. *Educação em Revista*, Belo Horizonte, v. 37, 2021, p. 1-17.

CHURCHILL, Rob; SINGH, Lisa. The evolution of topic modeling. *ACM Computing Surveys*, New York, v. 54, n. 10, 2022, p. 1-35.

CRAVO, Gustavo. A Base Nacional Comum Curricular atropelada pela reforma do ensino médio: uma entrevista com Marcelo Burgos. *Perspectiva Sociológica*, Rio de Janeiro, n. 21, 2018, p. 5-20.

DIMAGGIO, Paul. Adapting computational text analysis to social science (and vice versa). *Big Data & Society*, Londres, v. 2, n. 2, 2015, p. 1-5.

EGGER, Roman; YU, Joanne. A topic modeling comparison between LDA, NMF, Top2Vec, and BERTopic to Desmystify Twitter Posts. *Frontiers in Sociology*, Uppsala, Suécia, v. 7, 2022, p. 1-16.

EGGER, Roman; YU, Joanne. Identifying hidden semantic structures in Instagram data: a topic modelling comparison. *Tourism Review*, Leeds, England, 2021, p. 1234-1246.

ELRAGAL, Ahmed; KLISCHEWSKI, Ralf. Theory-driven or process-driven prediction? Epistemological challenges of big data analytics. *Journal of Big Data*, Londres, v. 4, n. 2, 2017, p. 1-20.

FALCÃO, Júlia Figueiredo Simão. *Yours Ever, Virginia Woolf*: Processamento de linguagem natural para descoberta de tópicos em cartas. Niterói: UFF, 2019. 55f. Monografia (Ciência da Computação) - Universidade Federal Fluminense.

FILIFE, Fabiana Alvarenga; SILVA, Dayane dos Santos; COSTA; Áurea de Carvalho. Uma base comum na escola: análise do projeto educativo da Base Nacional Comum Curricular. *Ensaio: Avaliação e Políticas Públicas em Educação*, Rio de Janeiro, v. 29, n. 112, 2021, p. 783-803.

GOODSON, Ivor F. *Currículo: teoria e história*. Petrópolis: Vozes, 2011.

HANNIGAN, Timothy R et al. Topic modeling in management research: rendering new theory from textual data. *Academy of Management Annals*, Nova York, v. 13, n. 2, 2019, p. 586-632.

HONG, Liangjie; DAVISON, Brian D. Empirical study of topic modeling in Twitter. *Proceedings of the First Workshop on Social Media Analytics*, Washington, 2010, p. 80-88.

JELODAR, Hamed; WANG, Yongli; YUAN, Chi; FENG, Xia; JIANG, Xiahui; LI, Yanchao; ZHAO, Liang. Latent Dirichlet Allocation (LDA) and topic modeling: models, applications, a survey. *ArXiv*, Nova York, 78, 2018, p. 15169-15211.

KRIPPENDORFF, Klaus. *Content analysis: an introduction to its methodology*. Thousand Oaks: Sage, 2004.

- LI, Qin; LI, Shaobo; ZHANG, Sen; HU, Jie; HU, Jianjun. A review of text corpus-based tourism big data mining. *Applied Sciences*, Basel, Switzerland, v. 9, n. 16, 2019, p. 3300.
- MORGAN, Hani. Conducting a qualitative document analysis. *The Qualitative Report*, Fort Lauderdale, FL (USA), v. 27, n. 1, 2022, p. 64-77.
- SACRISTÁN, José Gimeno. *O currículo: uma reflexão sobre a prática*. Porto Alegre: Artmed, 2000.
- SILVA, Tomás Tadeu. *Documentos de identidade: uma introdução às teorias do currículo*. Belo Horizonte: Autêntica, 2010.
- STOROPOLI, José Eduardo. Topic Modeling: how and why to use in management research. *Revista Ibero-Americana de Estratégia*, São Paulo. v.18, n. 3, 2019, p. 316-338.