



(RESEARCH ARTICLE)



Science, technology and society in chemistry education through the lens of topic modeling

Flávio Leite Rodrigues ^{1,*}, Galileu Batista de Sousa ² and Marcelo de Andrade Lima Maia ¹

¹ Universidade Federal do Rio Grande do Norte (UFRN), Brazil.

² Instituto Federal do Rio Grande do Norte (IFRN), Brazil.

International Journal of Science and Research Archive, 2025, 16(01), 2122-2129

Publication history: Received on 19 June 2025; revised on 27 July 2025; accepted on 30 July 2025

Article DOI: <https://doi.org/10.30574/ijrsra.2025.16.1.2248>

Abstract

This study investigates how the principles of Science, Technology, and Society (CTS) are reflected in chemistry education research. A corpus of 2,468 peer-reviewed abstracts from the Scopus database, retrieved in May 2025, was analyzed using Latent Dirichlet Allocation (LDA) to identify thematic patterns. Five main topics emerged: (1) Pedagogical Models and Conceptual Understanding, (2) Experimental Activities and Laboratory Learning, (3) Green Chemistry and Critical Thinking, (4) Educational Data and Learning Outcomes, and (5) Institutional Frameworks and Scientific Divisions. Temporal analysis revealed shifts from sustainability-focused discourse to data-driven evaluation. Gender-based analysis showed that female authors are more associated with laboratory-oriented research, while male authors are more prominent in data-focused studies. The results suggest that CTS-related themes are implicitly present in chemistry education literature and provide insights for future curricular development, policy, and teacher training. The study demonstrates the value of text mining for mapping academic discourse and offers methodological and theoretical contributions to science education research.

Keywords: Chemistry education; CTS; Topic modeling; LDA; Science teaching; Educational trends; Bibliometric analysis; Sustainability in education; Gender authorship

1. Introduction

The integration of Science, Technology, and Society (STS or CTS in Portuguese) into science education has gained increasing attention worldwide since the 1960s. This pedagogical orientation proposes that science education should transcend the mere transmission of conceptual knowledge, encouraging students to engage critically with the social dimensions of scientific development. Within this context, the CTS approach has been recognized as a powerful framework to promote scientific literacy and prepare individuals for active and responsible citizenship [1].

In chemistry education, CTS perspectives are essential to demystify the image of science as a neutral, purely objective endeavor. It helps students understand the epistemological, historical, and cultural underpinnings of chemical knowledge. As Santos and Mortimer argue, CTS-based curricula aim to contextualize scientific content within real-life sociotechnical problems, aligning science education with democratic values and civic responsibilities [2].

This approach is relevant across multiple educational levels. In secondary education, it aligns with curricular reforms that emphasize interdisciplinarity and the development of critical thinking. Pinheiro et al. highlight that CTS fosters a broader understanding of the interactions among science, technology, and society, particularly pertinent to the Brazilian Parâmetros Curriculares Nacionais (National Curriculum Parameters) [3].

* Corresponding author: Flávio Leite Rodrigues

At the university level, the implementation of CTS in teacher education programs has shown promise. Domiciano and Lorenzetti's case study at UFPR Litoral revealed that CTS contributes to the formation of socially engaged science teachers [4]. Similarly, Pelissari points to the convergence between CTS and polytechnic education as a strategy to promote ethical and political commitments in technological training [5].

Nevertheless, the operationalization of CTS in chemistry education remains challenging. Barriers include the persistence of traditional teaching models, limited teacher preparation, and the scarcity of pedagogical resources explicitly designed within a CTS framework. Given these challenges, this study aims to analyze how CTS has been incorporated into chemistry education across various educational contexts. By synthesizing key contributions from the literature, it also discusses future directions for research and practice in this field.

To better understand the presence and characteristics of CTS elements in chemistry education literature, this study conducted a systematic search in the Scopus database in May 2025. The search strategy included the keywords: "chemistry education" OR "chemistry teaching" OR "chemistry literacy". The resulting corpus of academic literature was analyzed through Latent Dirichlet Allocation (LDA), a probabilistic topic modeling technique that identifies latent themes within textual data [6].

The use of LDA for exploring open-ended textual corpora has proven effective in educational research, especially for capturing nuanced perceptions and emergent topics without imposing predefined categories [7]. In this study, topic modeling enabled the identification of prevailing discourses, thematic gaps, and the extent to which CTS-related concepts permeate chemistry education literature.

This article aims to present the findings of this analysis, providing insights into how the CTS perspective has been incorporated—explicitly or implicitly—into the academic conversation surrounding chemistry teaching and learning. The results contribute to the ongoing dialogue on curriculum innovation and teacher education, and they offer evidence-based recommendations for future integration of CTS in chemistry education at various levels.

2. Materials and Methods

2.1. Data Collection

A comprehensive search was conducted in the Scopus database in May 2025. The search terms used were: "chemistry education" OR "chemistry teaching" OR "chemistry literacy". This query returned a total of 2814 documents. Inclusion criteria comprised empirical studies and review articles containing a readable abstract. Exclusion criteria eliminated book chapters, conference proceedings, editorials, errata, letters to the editor, and any documents not meeting the defined standards of completeness or relevance.

After this screening process, 2468 articles were retained for analysis. For each document, the following metadata were extracted: title, list of authors, year of publication, and abstract text.

The corpus of abstracts was preprocessed using standard natural language processing (NLP) techniques. These included:

- Conversion to lowercase
- Tokenization
- Removal of stop words
- Lemmatization
- Elimination of punctuation and numeric characters

This preprocessing pipeline follows best practices in text mining and ensures data quality and consistency for topic modeling applications [8,9].

2.2. Topic Modeling with LDA

The preprocessed text corpus was analyzed using Latent Dirichlet Allocation (LDA), a generative probabilistic model designed to discover latent themes in large textual datasets [10]. LDA models each document as a mixture of topics, and each topic as a distribution over words.

To determine the optimal number of topics (K), multiple models were tested with K ranging from 5 to 30. The models were evaluated using:

- Topic coherence metrics: c_v and u_{mass} [11].
- Perplexity scores: for both training and held-out sets [11].

Figure 1 shows the coherence and perplexity trends, highlighting that the trade-off between interpretability and model performance was best balanced at $K = 5$.

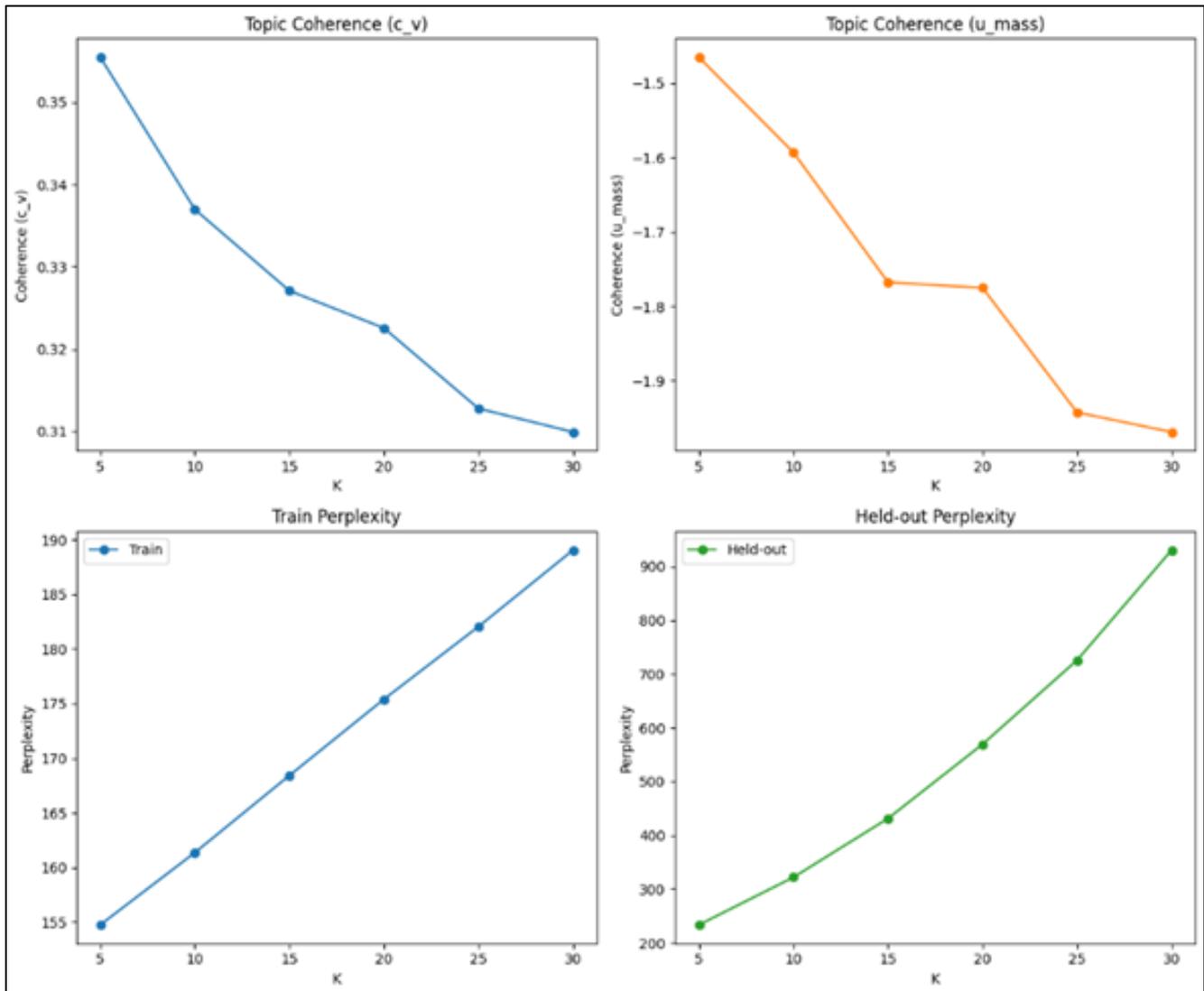
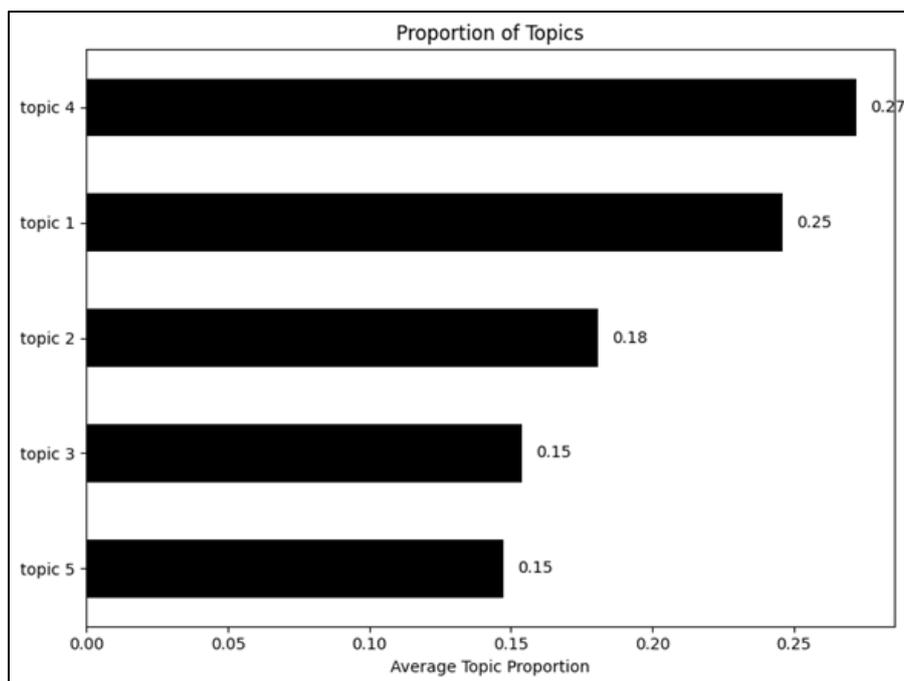


Figure 1 Evaluation metrics for LDA topic model selection across different numbers of topics (K). Source: Elaborated by the authors (2025)

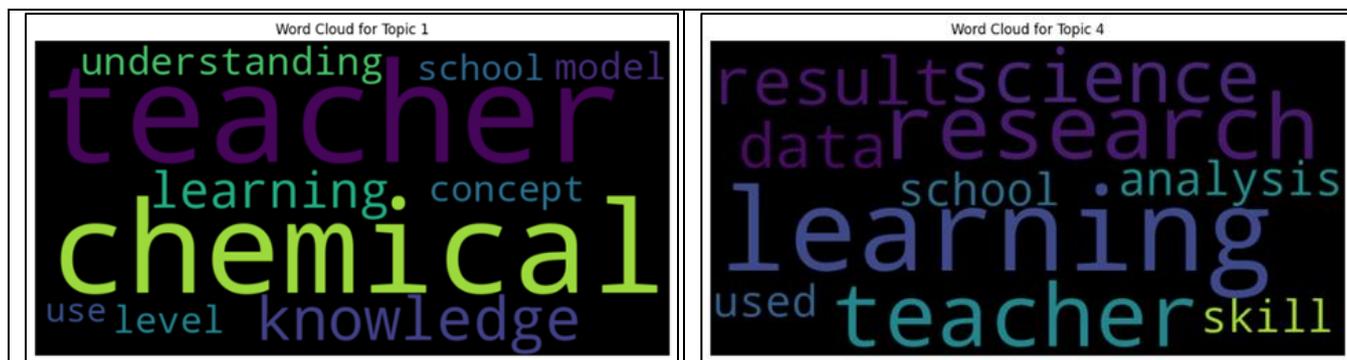
3. Results and Discussion

Once the optimal model was selected, each topic’s average proportion in the corpus was calculated (Figure 2). Word clouds were also generated for each topic to visually represent the most prominent terms (Figures 3 and 5).



Source: Elaborated by the authors (2025).

Figure 2 Average topic proportions in the LDA model (K = 5)



Source: Elaborated by the authors (2025).

Figure 3 and 4 Word cloud for Topic 1 – Pedagogical Models and Conceptual Understanding and Word cloud for Topic 4 – Educational Data and Learning Outcomes

This method has proven effective for exploratory analysis in educational research, as it allows for the inductive identification of thematic patterns and provides insight into discursive trends across large corpora without imposing a priori categories [12].

The Latent Dirichlet Allocation (LDA) model with K=5 topics revealed coherent and interpretable thematic clusters within the 2,468 abstracts. Topic labeling was performed using the FREX metric, which balances term frequency and exclusivity, making it particularly suitable for assigning descriptive and meaningful topic names [13].

The five topics identified are:

Topic 1 – Pedagogical Models and Conceptual Understanding

Characterized by terms such as teacher, learning, knowledge, and concept, this topic reflects studies focused on teaching approaches, cognitive development, and the role of teachers in chemistry education. FREX terms like societal and grading suggest concerns with assessment and broader educational goals.

Topic 2 – Experimental Activities and Laboratory Learning

Emphasizing group, experimental, laboratory, and method, this topic covers hands-on learning, control groups, and inquiry-based science education. FREX terms include equity and sensemaking, suggesting attention to inclusive and experiential learning environments.

Topic 3 – Green Chemistry and Critical Thinking

With high-probability words like green, thinking, sustainable, and curriculum, this topic aligns with sustainability, critical pedagogy, and environmentally oriented chemistry education.

Topic 4 – Educational Data and Learning Outcomes

Dominated by terms such as research, learning, data, and skills, this topic relates to empirical research on teaching effectiveness, assessment practices, and evidence-based instructional strategies.

Topic 5 – Institutional Frameworks and Scientific Divisions

Associated with course, chemical society, division, and instructor, this topic represents institutional and organizational aspects of chemistry education, including academic societies and program-level structures.

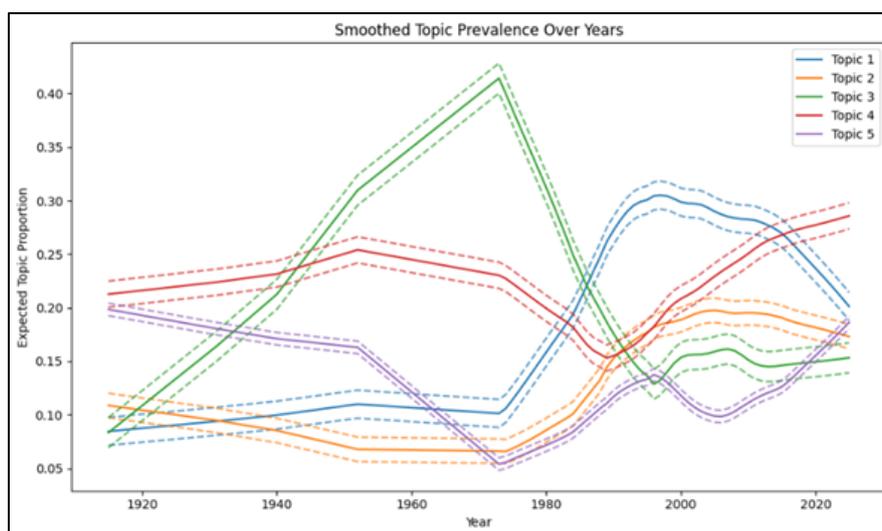
3.1. Temporal Evolution of Topics

Figure 5 displays the smoothed topic prevalence over time. The longitudinal trends reveal significant historical shifts:

Topic 3 – Green Chemistry and Critical Thinking peaked between the 1960s and 1980s, followed by a decline post-1990. This may reflect the saturation of curriculum development studies in that era and a later shift to more technical or data-driven approaches.

Topic 1 – Pedagogical Models and Conceptual Understanding gained prominence from the mid-1990s onward, likely driven by global reforms emphasizing student-centered learning.

Topic 4 – Educational Data and Learning Outcomes has grown steadily since the 2000s, mirroring the rise of data-intensive educational research and performance-based evaluation systems.



Source: Elaborated by the authors (2025).

Figure 5 Smoothed topic prevalence over time (1920–2025)

3.2. Gender-based Topic Distribution

Figure 6 shows the influence of author gender on topic prevalence. The structural topic model included gender covariates to estimate differential topic proportions:

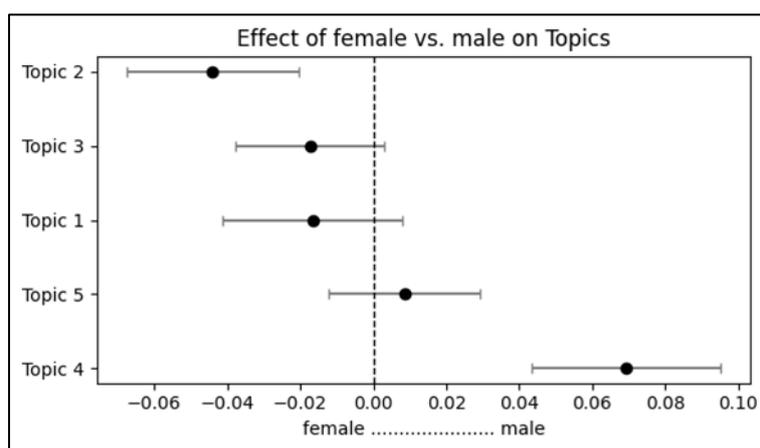
Topic 2 – Experimental Activities and Laboratory Learning was more frequently associated with male authors. This aligns with previous findings suggesting that women tend to focus on interactive and student-centered pedagogies.

Topic 4 – Educational Data and Learning Outcomes appeared more frequently in publications by female authors, suggesting a gendered difference in methodological preferences or research focus.

The other topics showed marginal or overlapping confidence intervals, indicating no strong gender effects in their distribution.

These findings highlight the diversity and dynamism of discourse in chemistry education. The presence of CTS-related themes (e.g., sustainability in Topic 3, equity in Topic 2, societal impact in Topic 1) underscores the indirect but persistent influence of the Science-Technology-Society approach across decades and subdomains of research.

Moreover, the prominence of institutional and data-driven narratives (Topics 4 and 5) suggests growing alignment between education research and accountability-oriented systems, while more humanistic and conceptual approaches (Topics 1 and 3) continue to shape pedagogical innovation.



Source: Elaborated by the authors (2025).

Figure 6 Gender-based differences in topic prevalence (female vs. male authors)

4. Conclusion

This study applied Latent Dirichlet Allocation (LDA) to a corpus of 2,468 abstracts from the Scopus database, aiming to explore how chemistry education research reflects and engages with the Science-Technology-Society (CTS) framework across time and authorship. Five coherent topics were identified and labeled based on their content and FREX-based keywords:

- Pedagogical Models and Conceptual Understanding,
- Experimental Activities and Laboratory Learning,
- Green Chemistry and Critical Thinking,
- Educational Data and Learning Outcomes, and
- Institutional Frameworks and Scientific Divisions.

The analysis revealed clear temporal dynamics. While sustainability-oriented discourse (Topic 3) peaked in earlier decades, recent years have seen a shift toward pedagogical innovation (Topic 1) and data-driven evaluation strategies (Topic 4). Additionally, gender-based analysis showed that male authors are more associated with experimental and hands-on learning research, while female authors more frequently contribute to topics involving data analytics and educational measurement.

These findings suggest that CTS-related themes are present in chemistry education literature, even when not explicitly labeled as such. Concepts like critical thinking, equity, sustainability, and the societal role of science permeate several of the dominant topics identified. This indicates that the CTS framework continues to inform educational research implicitly, guiding curricular and methodological developments.

Future research may build on these findings by conducting full-text analyses and extending the topic modeling to include multimodal or non-English literature. Moreover, qualitative studies could deepen our understanding of how CTS principles are operationalized in classrooms and teacher training programs. Strengthening the connection between CTS theory and educational practice remains a vital goal for fostering scientifically literate and socially responsible citizens.

Contributions and Implications

This study contributes to chemistry education by revealing, through topic modeling, how core Science-Technology-Society (CTS) principles—such as sustainability, equity, critical thinking, and institutional context—are reflected across decades of research. Methodologically, it showcases the utility of LDA and bibliometric analysis for mapping thematic evolution and author demographics. The findings inform educators and policy-makers by highlighting implicit CTS influences in current discourse and supporting the design of curricula and training programs aligned with social and pedagogical demands. Overall, the study bridges educational theory and computational analysis to illuminate trends and guide future inquiry.

Limitations and Future Work

This study is limited by its reliance on abstracts rather than full-text articles, which may constrain the depth and nuance of topic identification. Additionally, the analysis was restricted to documents indexed in the Scopus database, potentially excluding relevant literature from other sources or in non-English languages. The gender analysis was based on inferred author names, which may introduce classification bias. Future research could address these limitations by incorporating full-text analysis, expanding the corpus across multiple databases, and integrating qualitative methods to explore how CTS principles are explicitly or implicitly embedded in instructional practices, curricular frameworks, and teacher education programs.

Compliance with ethical standards

Acknowledgments

We thank CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brazil) for maintaining open access to two essential platforms that support scientific research in Brazil: The Thesis and Dissertation Catalog, which provides visibility to national academic work, and the CAPES Journal Portal (www.periodicos.capes.gov.br), which grants institutional access to international scientific databases such as Scopus. These resources were fundamental for data collection and literature analysis in this study.

Disclosure of conflict of interest

The authors declare no conflict of interest.

References

- [1] Angotti JAP, Auth MA. Ciência e tecnologia: implicações sociais e o papel da educação. *Ciênc. educ.* (Bauru). 2001; 7:15-27.
- [2] Santos WLP, Mortimer EF. Uma análise de pressupostos teóricos da abordagem C-T-S (Ciência - Tecnologia - Sociedade) no contexto da educação brasileira. *Ens. Pesqui. Educ. Ciênc.* (Belo Horizonte). 2000; 2:110-132.
- [3] Pinheiro NAM, Silveira RMCF, Bazzo WA. Ciência, Tecnologia e Sociedade: a relevância do enfoque CTS para o contexto do Ensino Médio. *Ciênc. educ.* (Bauru). 2007; 13:71-84.
- [4] Domiciano TD, Lorenzetti L. A educação ciência, tecnologia e sociedade no curso de licenciatura em ciências da UFPR Litoral. *Ens. Pesqui. Educ. Ciênc.* (Belo Horizonte). 2020; 22:e14848.
- [5] Pelissari LB. A perspectiva Ciência, Tecnologia e Sociedade no contexto da educação tecnológica e politécnica: primeiras aproximações teóricas. *Rev. Bras. Educ.* 2023; 28:e280071.
- [6] Blei DM, Ng AY, Jordan MI. Latent Dirichlet Allocation. *J. Mach. Learn. Res.* 2003; 3:993–1022.
- [7] Wang M, Gao S, Gui W, Ye J, Mi S. Investigation of Pre-service Teachers' Conceptions of the Nature of Science Based on the LDA Model. *Sci Educ.* 2023;32(3):589–615.

- [8] Grimmer J, Stewart BM. Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Political Analysis*. 2013;21(3):267-297.
- [9] Roberts ME, Stewart BM, Tingley D, Lucas C, Leder-Luis J, Gadarian SK, et al. Structural topic models for open-ended survey responses. *Am J Polit Sci*. 2014;58(4):1064–82.
- [10] Blei DM, Ng AY, Jordan MI. Latent Dirichlet Allocation. *J. Mach. Learn. Res.* 2003; 3:993–1022.
- [11] Weston SJ, Shryock I, Light R, Fisher PA. Selecting the number and labels of topics in topic modeling: A tutorial. *Adv Methods Pract Psychol Sci*. 2023;6(2):25152459231160105.
- [12] Cifuentes J, Olarte F. A macro perspective of the perceptions of the education system via topic modelling analysis. *Multimed Tools Appl*. 2023;82(2):1783–1820.
- [13] Mi S, Zong T, Yang X, Gui W. Physics pre-service teachers' conceptual understanding of scientific literacy. *Sci Educ*. 2024; 32:589–615.