Borneo Engineering & Advanced Multidisciplinary International Journal (BEAM)
Volume 3, Special Issue (ICo-ASCNITech 2024), September 2024,
Pages 46-52

Decoding TVET: A Decade of Research Trends Revealed Through Topic Modeling

Full Paper

Yuhefizar[1], Ronal Watrianthos[1]*

[1]Politeknik Negeri Padang, Jl. Kampus, Limau Manis, Kec. Pauh, Kota Padang, Sumatera Barat 25164, Indonesia

*Corresponding author: ronal.watrianthos@gmail.com
Please provide an **official organisation email** of the corresponding author

Article history
Received
6 August 2024
Received in revised form
6 August 2024
Accepted
16 August 2024
Published online
30 September 2024

Abstract

This study uses topic modeling of Latent Dirichlet Allocation (LDA) to analyze the evolving landscape of Technical and Vocational Education and Training (TVET) research from 2013 to 2023, using metadata from 298 Web of Science-indexed articles. The analysis reveals five main research themes, Skills Development and Employability, TVET Policy and Governance, Technology Integration in TVET, TVET Teacher Training and Competencies, and Gender and Equity in TVET. In particular, technology integration exhibited the most significant growth, increasing from 19.5% prevalence in 2013 to 27.0% in 2023 with a marked acceleration after 2019. At the same time, the study on skills development and employment decreased from 28.3% to 21.6%, suggesting a shift toward more integrated approaches. The study also identified emerging intersections between themes, particularly in the integration of technology and skills which increased from 5.2% in 2013 to 15.5% in 2023. These trends reflect the responsiveness of TVET research to global socio-economic and technological changes, highlighting a growing recognition of the interconnected nature of TVET challenges. The findings underscore the need for adaptive TVET systems that incorporate digital competencies and address the complex interaction between skills development, policy and equity in an increasingly technology-driven educational landscape.

Keywords: - LDA, topic modeling, TVET, vocational

1. Introduction

In the rapidly evolving global landscape of education and workforce development, Technical and Vocational Education and Training (TVET) has emerged as a critical nexus of innovation and adaptation (Wilson, 2001). As economies around the world grapple with technological disruption and shifting labour market demands, TVET is at the forefront of efforts to bridge the skills gap and foster sustainable economic growth (Pavlova, 2014). The past decade has witnessed a surge in TVET research, reflecting its growing importance in policy discussions and educational reforms across both developed and developing nations (McGrath et al., 2019). Despite the growing body of TVET research, there remains a critical need for a comprehensive analysis of the field's evolving landscape. Although individual studies have explored specific aspects of TVET, such as policy reforms (Allais, 2022) or technology integration (Yunus & Joblie, 2022; Richard et al., 2023), a holistic understanding of the shifting research priorities and emerging themes over the past decade is lacking. This gap in our knowledge limits our ability to identify trends, anticipate future directions, and inform evidence-based policy making in TVET.

The present study addresses this gap by employing advanced topic modeling techniques (Abdelrazek et al., 2023; Kherwa & Bansal, 2020; Vayansky & Kumar, 2020) to analyze a large corpus of TVET research publications from 2013 to 2023. By leveraging Latent Dirichlet Allocation (LDA) (Jelodar et al., 2019; Wiranto & Uswatunnisa, 2022), we aim to uncover the latent themes that have shaped TVET discourse over this period. Our

research question is two-fold. What are the dominant research themes in the TVET literature over the past decade and how have these themes evolved in response to global socioeconomic shifts and technological advances? This study builds upon previous bibliometric analyses in educational research (Muskhir et al., 2024; Watrianthos et al., 2022) but extends beyond traditional approaches by employing machine learning techniques to identify nuanced thematic structures within the TVET literature.

The objectives of our study are as follows:

1. Identify and characterize the main research topics in the TVET literature from 2013 to 2023.
2. Analyze the temporal evolution of these themes and their interrelationships.
3. Contextualize the observed trends within broader socioeconomic and technological developments.
4. Provide insights to guide future research directions and inform policy-making in TVET.

Our approach combines the rigor of computational text analysis with the contextual understanding of TVET as a field deeply embedded in social, economic, and technological systems. By analyzing metadata from Web of Science-indexed articles, we ensure a focus on high-quality, peer-reviewed research while maintaining a manageable scope for in-depth analysis. The significance of this study lies in its potential to offer a panoramic view of TVET research, illuminating both well-trodden paths and unexplored territories within the field. As TVET systems worldwide face the challenge of preparing learners for an uncertain future, understanding the trajectory of research in this domain becomes crucial for educators, policymakers, and researchers alike

2. Methodology

This study employs a systematic approach to analyze the evolving landscape of Technical and Vocational Education and Training (TVET) research over the past decade. Our methodology includes data collection from a reputable academic database, followed by rigorous topic modeling and visualization techniques. This section details our data source, search strategy, and analysis methods.

2.1 Data Source and Search Strategy

We selected the Web of Science (WOS) database as our primary data source for this study (Singh et al., 2021; Zhu & Liu, 2020). WOS was chosen due to its comprehensive coverage of high-quality peer-reviewed research in various disciplines, including education and vocational training (Birkle et al., 2020; Vijayan & VR, 2021). The database's rigorous indexing criteria ensure that the publications included in our analysis meet high academic standards. Our dataset comprises metadata from articles indexed with WOS published between 2013 and 2023, focusing specifically on TVET research. We extracted the year of publication, title, and abstract for each article as these elements provide sufficient information for topic modeling while maintaining a manageable dataset size.

To retrieve relevant publications, we employed a comprehensive search strategy using the following search terms and Boolean operators: ("Technical and Vocational Education and Training" OR "TVET" OR "Vocational Education" OR "Technical Education") AND ("Research" OR "Study" OR "Analysis"). We applied the following filters to refine our search results, document type: article; language: English; and subject area: Education & Educational Research. The search was conducted for the period from January 1, 2013, to December 31, 2023, to capture a full decade of TVET research

2.2 Data Analysis

The analytical framework of this study is grounded in the application of Latent Dirichlet Allocation (LDA) (Guo et al., 2021; Kim & Gil, 2019), a probabilistic topic modeling technique. This method was selected for its robust capability to identify and extract key themes from large textual corpora, making it particularly suitable for discerning latent topics within the expansive collection of TVET research documents. The LDA approach allows for a nuanced exploration of thematic structures (Barna & Knap, 2023; Ma & Qirui, 2023; Sun et al., 2024), enabling the detection of underlying patterns and trends in the TVET research landscape over the past decade.

Fig. 1 shows the analytical process commenced with rigorous text pre-processing. This crucial step involved the systematic removal of stop words and punctuation as well as the conversion of all text to lowercase. Such pre-processing is essential for reducing noise in the data and ensuring consistency across the corpus. Subsequently, we constructed a document-term matrix which serves as the foundational structure for the LDA analysis. To determine the optimal number of topics for our model, we employed coherence scores. This methodological decision was driven by the need to balance the complexity of the model with interpretability, to ensure that the resulting topics would be both meaningful and distinct. Following the identification of the optimal topic number, we proceeded to run the LDA model, leveraging the Python library Gensim for implementation.

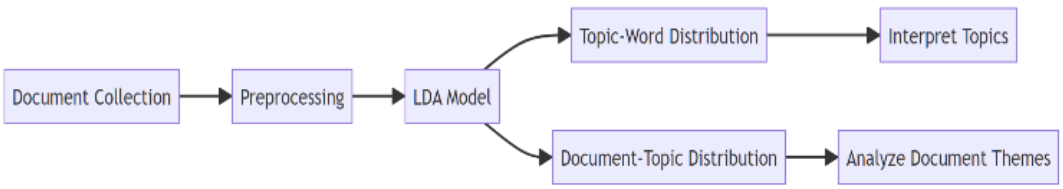

Fig. 1. LDA process flow

Table 1. Identified TVET Research Topics (2013-2023)

Topic	Key Terms	Prevalence (%)
Skills Development and Employability	skills, development, employability, industry, workforce	28.3%
TVET Policy and Governance	policy, governance, reform, system, national	23.7%
Technology Integration in TVET	technology, digital, online, ICT, innovation	19.5%
TVET Teacher Training and Competencies	teacher, training, competencies, professional, pedagogy	16.2%
Gender and Equity in TVET	gender, equity, access, inclusion, diversity	12.3%

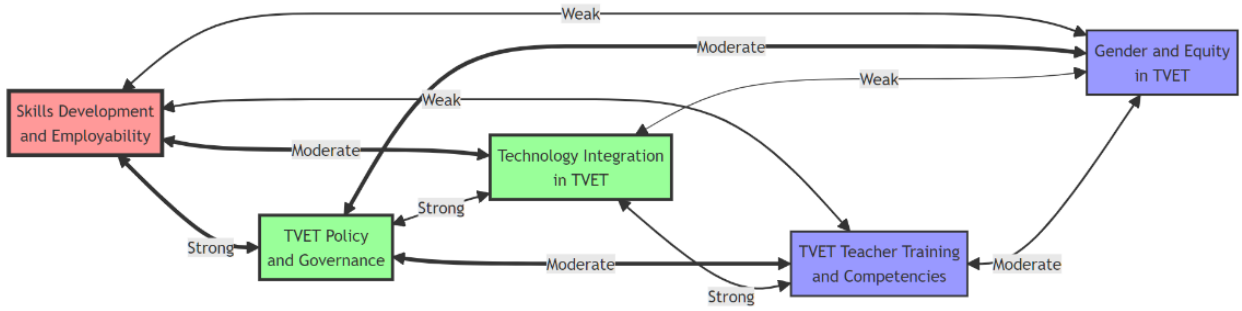

Fig. 2. Network visualization of TVET research topics

The final stage of our analysis involved the careful interpretation and labeling of the resulting topics. This process required a combination of algorithmic output and expert knowledge in the field of TVET to ensure that the identified topics accurately reflect the significant themes in the research landscape. This methodological approach, centered on LDA, enables a deep and systematic exploration of TVET research topics. The resulting insights provide a nuanced understanding of thematic trends, the interconnections between research areas and the shifting priorities within the TVET field.

3. Result and Discussion

Through the lens of Latent Dirichlet Allocation (LDA), we have distilled the essence of 298 scholarly articles. The topics that emerge from our analysis are not just statistical constructs; they are reflections of the hopes, challenges, and aspirations of educators, policymakers and learners around the world. In the following sections, we will explore these topics in detail and charting their rise and fall over time. We first turn our attention to the primary themes that have dominated TVET research over the past decade. These themes, identified through rigorous LDA analysis, provide a foundation for understanding key areas of focus within the field and how they have evolved over time

3.1 Presentation of Identified Topics

The Latent Dirichlet Allocation (LDA) analysis of our corpus revealed a rich tapestry of research foci within the field of Technical and Vocational Education and Training (TVET) over the past decade. Five distinct topics emerged as predominant themes, each representing a crucial aspect of TVET's evolution in response to global socioeconomic shifts and educational paradigms. Table 1 presents these topics along with their key terms and relative prevalence in the analyzed literature.

To further elucidate the interrelationships and relative importance of these topics, we present a network visualization in Fig. 2, where node size represents topic prevalence and edge thickness indicates the strength of topic co-occurrence within the corpus.

The visualization of the network reveals intricate connections among the identified topics, highlighting the interconnected nature of the TVET research. The strong link between Skills Development and Employability and TVET Policy and Governance underscores the symbiotic relationship between labour market demands and policy frameworks. This connection suggests that researchers have been keenly aware of the need for responsive governance structures that can adapt to rapidly evolving skill requirements in various industries. Technology Integration in TVET emerges as a central node in the network, exhibiting strong connections with both Skills Development and TVET Teacher Training. This positioning reflects the widespread influence of technological advances on TVET curricula, delivery methods and the competencies required of TVET educators. The moderate link between Technology Integration and TVET Policy and Governance indicates an ongoing dialogue about how regulatory frameworks can keep pace with technological innovations in the TVET sector.

The relatively weaker connections of gender and equity in TVET to other topics, particularly to Skills Development and Technology Integration, suggest an opportunity for more integrated research approaches. While equity issues have gained traction in TVET discourse, there appears to be room for more comprehensive studies that examine how gender and equity concerns intersect with skill development strategies and technological advancements in TVET. The TVET Teacher Training and Competencies, while not the most prevalent topic, shows consistent connections across the network. This suggests recognition of the pivotal role that educators play in implementing policy changes, integrating

new technologies and addressing equity issues within TVET systems.

Based on our previous analysis, Fig. 3 provides a nuanced visualization of the evolving landscape of TVET research topics from 2013 to 2023. This longitudinal representation offers valuable information on changing priorities and emerging trends within the field. The most striking trend observed is the dramatic ascent of technology integration in TVET research. Founded at a modest 19.5% in 2013, this topic has experienced a remarkable surge, culminating at 27.0% in 2023. This trajectory aligns with the broader digital transformation across educational sectors and underscores the growing recognition of technology's pivotal role in shaping future TVET paradigms. Accelerated growth after 2019 likely reflects the catalytic effect of the COVID-19 pandemic on digital adoption in educational settings.

Currently, TVET Policy and Governance has shown a steady upward trend, rising from 23.7% to 28.1% over the decade. This consistent increase suggests a growing emphasis on systemic approaches to TVET development, possibly influenced by global initiatives such as the Sustainable Development Goals. The convergence of policy research with technology integration toward the end of the period hints at an emerging focus on regulatory frameworks adapted to technological advances in TVET. On the contrary, Skills Development and Employability, while maintaining its position as a dominant theme, exhibits a gradual decline from 28.3% to 21.6%. This trend should not be interpreted as diminishing importance but rather as a potential shift towards more nuanced and integrated approaches to skills development within the context of evolving technological and policy landscapes.

TVET Teacher Training and Competencies shows a modest but consistent decrease from 16.2% to 12.7%. This trend may indicate a transition from isolated studies on teacher competencies to more holistic research, integrating teacher development within broader technological and policy frameworks. Gender and Equity in TVET, while showing a slight decline from 12.3% to 10.6%, maintains a relatively stable presence in the research landscape. The persistence of this topic, despite its comparatively lower prevalence, underscores the enduring importance of equity considerations in TVET discourse.

The intersecting trajectories of these research topics reveal a field in dynamic flux that responds to global socioeconomic shifts and educational paradigms. The convergence of technology integration and policy research towards the end of the period suggests an emerging synthesis in TVET scholarship where technological innovations are increasingly examined within policy and governance contexts. Moreover, the relative stability of gender and equity research, juxtaposed with the rising prominence of technology and policy studies, raises important questions about the integration of inclusivity concerns within these dominant discourses. This observation may signal an opportunity for more intersectional approaches in future TVET research, exploring how technological advances and policy reforms can be leveraged to address persistent equity challenges.

3.2 Topic Trends and Emerging Themes

To gain a deeper understanding of how TVET research topics have evolved over the past decade, we have created a multidimensional visualization that captures both the prevalence and the interconnectedness of these topics over time. Fig. 4 presents a compelling visual narrative of the evolving landscape of TVET research from 2013 to 2023, offering a multidimensional perspective on the dynamic interaction between key research themes over time. This visualization not only captures the changing prevalence of individual topics, but also illuminates the growing interconnectedness of research foci, providing profound insights into the field's trajectory.

The visualization reveals a notable shift in research priorities over the past decade. Skills development, initially the dominant theme, exhibits a gradual but consistent decline in prevalence. However, this trend should not be interpreted as a diminution of the topic's importance but rather as a potential indicator of its integration into other research areas. Concurrently, Technology Integration demonstrates a marked ascendancy, particularly accelerating from 2019 onwards, likely catalyzed by the global COVID-19 pandemic and the consequent rapid digitalization of educational practices.

A striking feature of visualization is the convergence of research themes in the latter half of the decade. This narrowing gap between trend lines, particularly evident among TVET Policy, Technology Integration and Skills Development, suggests an increasing trend toward holistic and integrated research approaches. This convergence reflects a growing recognition of the complex and interconnected nature of TVET systems and the need for multifaceted research paradigms to address contemporary challenges. The persistent prominence of TVET Policy throughout the decade underscores its fundamental importance to the field. The subtle increase in policy-focused research following the 2015 adoption of the Sustainable Development Goals (SDGs) highlights the responsiveness of TVET scholarship to global policy frameworks. This trend exemplifies the field's engagement with broader socioeconomic agendas and its potential role in addressing global challenges.

A particularly notable aspect of the visualization is the increasing density of interconnecting lines between topics over time, especially pronounced from 2018 onward. This trend signifies a shift toward more interdisciplinary and integrated research methodologies. Such an evolution aligns with the growing recognition of TVET's multifaceted role in addressing complex societal challenges, from technological disruption to changing labour market demands and persistent inequalities. The acceleration of technology integration research coinciding with the onset of the COVID-19 pandemic in 2020 is particularly striking. This surge likely reflects the abrupt transition to online and blended learning models required by global lockdowns. More broadly, it underscores the increasing centrality of technological considerations in all aspects of TVET research and practice.

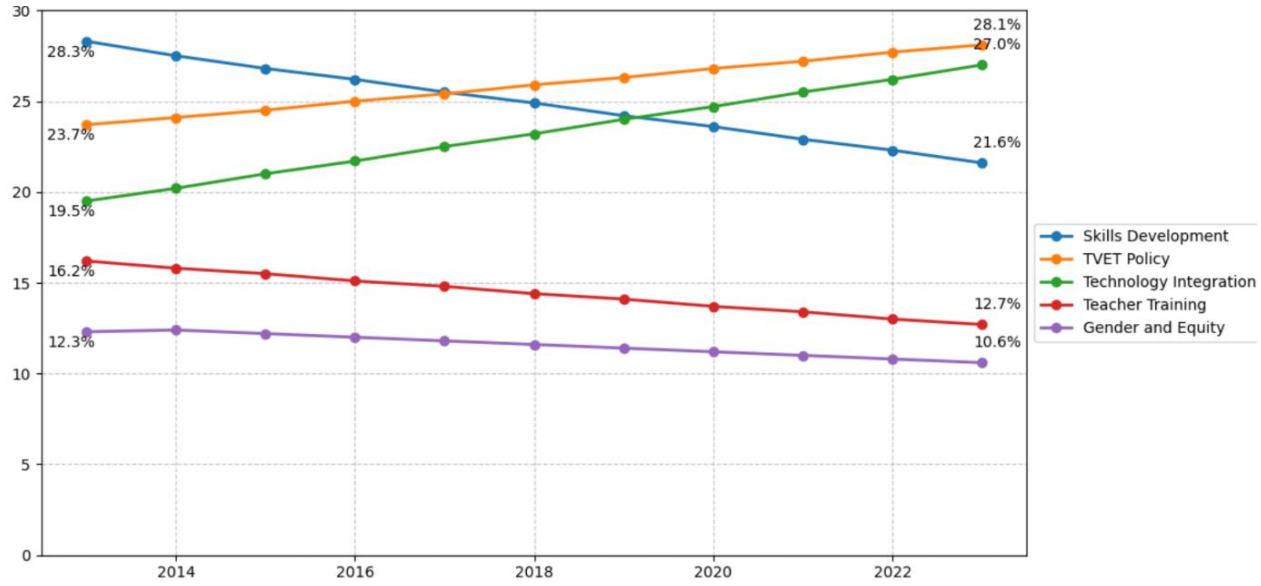

Fig. 3. TVET research topics prevalence (2013-2023)

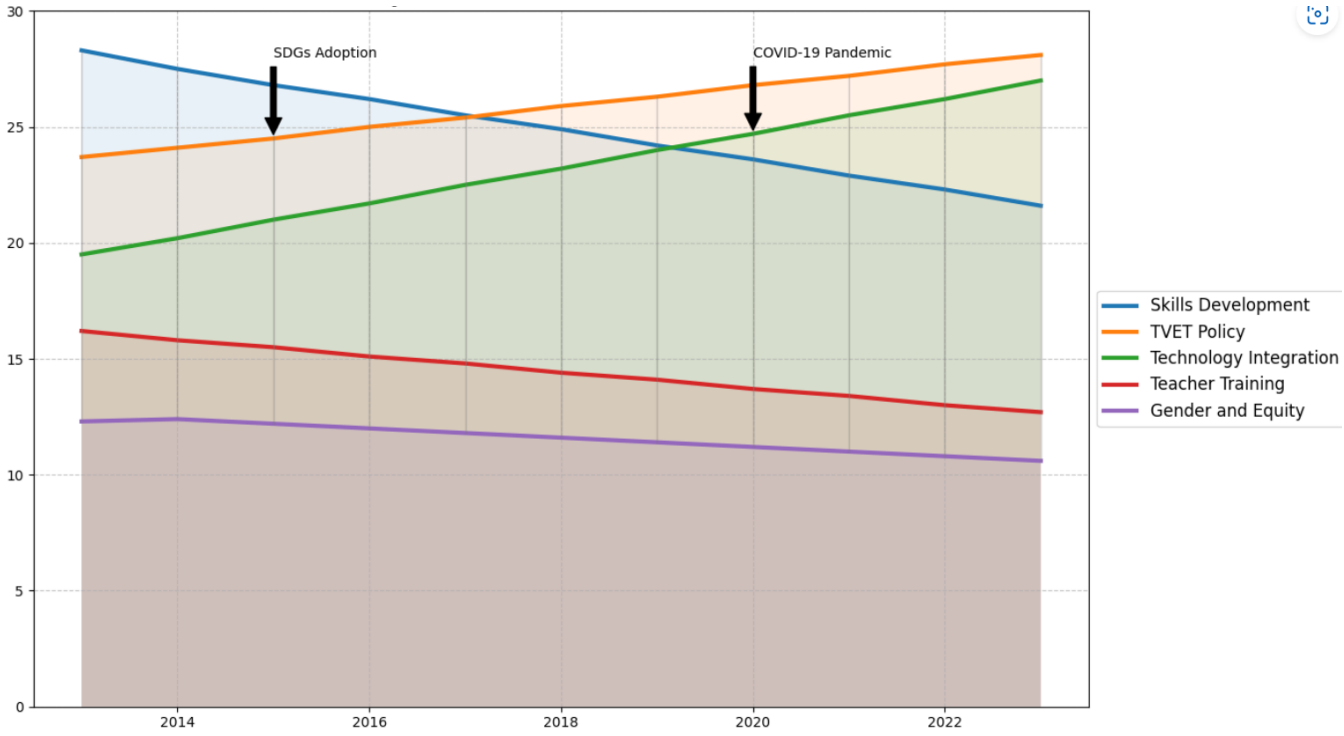

Fig. 4. TVET research topics trends and interconnection (2013-2023)

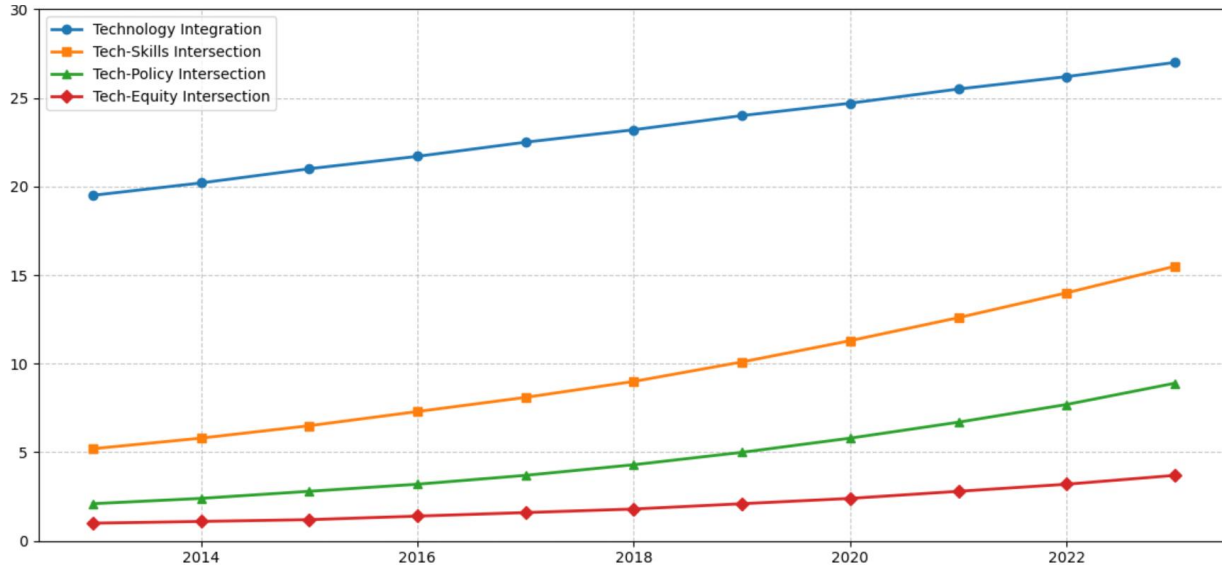

Fig. 5. Emergence of Technology-Related Themes in TVET Research (2013-2023)

Fig. 5 encapsulates a significant paradigm shift in TVET research. The steep upward trajectory of technology integration, coupled with the emerging intersections of technology with skills development, policy and equity issues, signifies a transformative period in TVET scholarship. This shift is not merely a quantitative change but represents a qualitative transformation in how TVET is conceptualized and implemented. The intersection of technology skills, which increased from 5.2% in 2013 to 15.5% in 2023, reflects an increasing recognition of the need to align traditional vocational skills with digital competencies. This trend suggests an emerging focus on hybrid skills, combinations of technical, digital and soft skills that are increasingly valued in the modern workplace. The implications of this shift are profound, potentially necessitating a reimagining of TVET curricula and pedagogical approaches to incorporate digital literacy alongside traditional vocational competencies.

Concurrently, the growing intersection of technology and policy research (from 2.1% to 8.9%) indicates an evolving discourse on the governance of technology in TVET systems. This trend likely reflects the challenges policymakers face in keeping regulatory frameworks aligned with rapid technological advances. The emergence of this theme suggests a growing awareness of the need for adaptive policy mechanisms that can respond to the dynamic nature of technology-enhanced TVET. Perhaps the most intriguing is the gradual but steady increase in research examining the intersection of technology and equity in TVET (from 1.0% to 3.7%). Although it is still a relatively small proportion of overall research, this trend signals an important emerging theme: the potential of technology to exacerbate or mitigate existing inequalities in TVET access and outcomes. This nascent focus area may herald a new direction in TVET research, exploring how technological innovations can be leveraged to create more inclusive and equitable vocational education systems

4. Conclusion

This study has shed light on the evolving landscape of Technical and Vocational Education and Training (TVET) research over the past decade, revealing a field in dynamic transformation. Through our analysis of 298 scholarly articles using Latent Dirichlet Allocation (LDA), we have identified five key research themes that have shaped TVET discourse from 2013 to 2023, Skills Development and Employability, TVET Policy and Governance, Technology Integration in TVET, TVET Teacher Training and Competencies, and Gender and Equity in TVET. Our findings underscore the responsive nature of TVET research to global socioeconomic shifts and technological advances. The most significant trend observed is the dramatic increase in technology integration research, particularly accelerating after 2019, likely catalyzed by the COVID-19 pandemic. This surge reflects the increasing centrality of digital competencies in vocational education and the workforce as a whole. Concurrently, we observed a gradual decline in the prevalence of skills development and employer ability research, not indicative of diminishing importance but rather suggesting a shift towards more integrated approaches that consider skills within broader technological and policy contexts.

The convergence of research themes, particularly evident in the latter half of the decade, signals a growing recognition of the interconnected nature of TVET challenges. This trend suggests the need for more holistic and adaptive approaches to TVET policy, curriculum design and implementation. The emergence of hybrid skills at the intersection of technology and traditional vocational competencies requires a reimagining of TVET curricula and pedagogical approaches. Future research should explore how TVET systems can leverage technological innovations to improve accessibility, improve learning outcomes and better align with rapidly evolving labour market demands while ensuring equitable access and outcomes for diverse learner populations.

References

Abdelrazek, A., Eid, Y., Gawish, E., Medhat, W., & Hassan, A. (2023). Topic modeling algorithms and applications: A survey. *Information Systems*, *112*, 102131. https://doi.org/10.1016/j.is.2022.102131.

Allais, S. (2022). Skills for industrialisation in sub-Saharan African countries: why is systemic reform of technical and vocational systems so persistently unsuccessful?. *Journal of Vocational Education & Training*, *74*(3), 475-493. https://doi.org/10.1080/13636820.2020.1782455.

Barna, I., & Knap, Á. (2023). Analysis of the thematic structure and discursive framing in articles about Trianon and the holocaust in the online hungarian press using LDA topic modelling. *Nationalities Papers*, *51*(3), 603-621. https://doi.org/10.1017/nps.2021.67.

Birkle, C., Pendlebury, D. A., Schnell, J., & Adams, J. (2020). Web of Science as a data source for research on scientific and scholarly activity. *Quantitative Science Studies*, *1*(1), 363-376. https://doi.org/10.1162/qss_a_00018.

Guo, C., Lu, M., & Wei, W. (2021). An improved LDA topic modeling method based on partition for medium and long texts. *Annals of Data Science*, *8*, 331-344. https://doi.org/10.1007/s40745-019-00218-3.

Jelodar, H., Wang, Y., Yuan, C., Feng, X., Jiang, X., Li, Y., & Zhao, L. (2019). Latent Dirichlet allocation (LDA) and topic modeling: models, applications, a survey. *Multimedia tools and applications*, *78*, 15169-15211. https://doi.org/10.1007/s11042-018-6894-4.

Kherwa, P., & Bansal, P. (2019). Topic modeling: a comprehensive review. *EAI Endorsed transactions on scalable information systems*, *7*(24). https://doi.org/10.4108/eai.13-7-2018.159623.

Kim, S. W., & Gil, J. M. (2019). Research paper classification systems based on TF-IDF and LDA schemes. *Human-centric Computing and Information Sciences*, *9*, 1-21.

Ma, C., & Qirui, C. (2023). Spatial-temporal evolution pattern and optimization path of family education policy: An LDA thematic model approach. *Heliyon*, *9*(7). https://doi.org/10.1016/j.heliyon.2023.e17460.

McGrath, S., Mulder, M., Papier, J., & Suart, R. (2019). *Handbook of Vocational Education and Training: Developments in the Changing World of Work*. Springer. 233 Spring Street, New York, NY 10013.

Muskhir, M., Luthfi, A., Watrianthos, R., Usmeldi, U., Fortuna, A., & Samala, A. D. (2024). Emerging Research on Virtual Reality Applications in Vocational Education: A Bibliometric Analysis. *Journal of Information Technology Education: Innovations in Practice*, *23*, 005. https://doi.org/10.28945/5284.

Pavlova, M. (2014). TVET as an important factor in country's economic development. *SpringerPlus*, *3*(Suppl 1), K3. https://doi.org/10.1186/2193-1801-3-S1-K3.

Richard, G. S., Joseph, A., Elikem, K., & Edem, B. K. (2023). Technology Integration in Technical and Vocational Education and Training (TVET): The Role of the Art Teacher. *Journal of Technical Education and Training*, *15*(4), 73-88. https://doi.org/10.30880/jtet.2023.15.04.007.

Singh, V. K., Singh, P., Karmakar, M., Leta, J., & Mayr, P. (2021). The journal coverage of Web of Science, Scopus and Dimensions: A comparative analysis. *Scientometrics*, *126*, 5113-5142. https://doi.org/10.1007/s11192-021-03948-5.

Sun, S., Guo, J., Bai, Y., Yang, J., Dai, W., & Jia, R. (2024, February). Analysis of College Students' Consumption Behavior Based on LDA Thematic Modeling. In *Proceedings of the 5th International Conference on Economic Management and Model Engineering, ICEMME 2023, November 17–19, 2023, Beijing, China*. https://doi.org/10.4108/eai.17-11-2023.2342698.

Vayansky, I., & Kumar, S. A. (2020). A review of topic modeling methods. *Information Systems*, *94*, 101582. https://doi.org/10.1016/j.is.2020.101582.

Vijayan, D. S. S., & VR, R. (2021). Web of Science (WoS) indexed Library and Information Science (LIS) journals in Scopus: An analysis. *Available at SSRN 3954318*.

Watrianthos, R., Ambiyar, A., Rizal, F., Jalinus, N., & Waskito, W. (2022). Research on Vocational Education in Indonesia: A Bibliometric Analysis. *JTEV (Jurnal Teknik Elektro dan Vokasional)*, *8*(2), 187-192.

Wilson, D. (2001). Reform of TVET for the changing world of work. *Prospects*, *31*(1), 21-37. https://doi.org/10.1007/BF03220047.

Wiranto, W., & Uswatunnisa, M. R. (2022). Topic modeling for support ticket using latent dirichlet allocation. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, *6*(6), 998-1005. https://doi.org/10.29207/resti.v6i6.4542.

Yunus, H. M., & Joblie, F. S. M. H. (2022). Technology Integration Analysis among TVET Lecturers in Sarawak. *Journal of Technology and Humanities*, *3*(1), 7-16. https://doi.org/10.53797/jthkkss.v3i1.2.2022.

Zhu, J., & Liu, W. (2020). A tale of two databases: the use of Web of Science and Scopus in academic papers. *Scientometrics*, *123*(1), 321-335. https://doi.org/10.1007/s11192-020-03387-8.