2025 Journal of Engineering, Technology and Social Sciences
Jurnal Kejuruteraan, Teknologi dan Sains Sosial
Volume 11 Special Issue: ICETISM
International Conference on Emerging Technologies, Information Science and Mathematics
e-ISSN: 27166848

A REVIEW OF BIG DATA MANAGEMENT IN MALAYSIAN POLYTECHNICS: STATE-OF-THE-ART, APPLICATIONS, CHALLENGES AND FUTURE DIRECTIONS

Nor Azida Mansor[1*], Munirah Abdullah[2] and Sharifah Nur Syed Ismail[3]

[1,2,3] *Department of Information and Communications Technology, Ungku Omar Polytechnic*

[*]*azida@puo.edu.my*

ARTICLE INFO

Article history:
Received
14 July 2025
Received in revised form
18 Sept 2025
Accepted
3 Oct 2025
Published online
15 Oct 2025

***Keywords**:*
big data management, Malaysian Polytechnics, TVET, digitalisation

ABSTRACT

This review study looks at the existing condition, uses, problems, and future directions of big data management in Malaysian polytechnics. As part of Malaysia's Technical and Vocational Education and Training (TVET) ecosystem, polytechnics are increasingly expected to adopt data-driven approaches to enhance teaching, institutional efficiency, and graduate employability. Drawing on global research and guided by the Technology–Organization–Environment (TOE) and Diffusion of Innovation (DOI) frameworks, this study critically explores how big data tools such as learning analytics, predictive modelling, and dashboards are applied in polytechnic settings. Despite promising developments, polytechnics face persistent challenges: outdated infrastructure, fragmented systems, limited staff competence, and weak governance. National initiatives such as the TVET Digitalisation Strategic Plan and Dasar TVET Negara 2030 can help address these gaps by promoting centralized data governance and institutional readiness. The paper recommends strengthening infrastructure, staff capacity building, pilot analytics projects, and integrating graduate employment data. Future trends include AI-powered analytics, smart campus technologies, and inter-institutional data sharing. With strong leadership and coordinated governance, Malaysian polytechnics can evolve into agile, data-driven institutions aligned with national digital transformation goals.

1. Introduction

Big data is transforming higher education by enhancing teaching quality, administrative efficiency, and institutional planning. A bibliometric study by Samsul et al. (2023) shows rapid global growth in research on educational big data and learning analytics, underscoring their role as key drivers of innovation.

In Malaysia, the adoption of big data in higher education has been uneven. Public universities benefit from stronger infrastructure and policy support, while polytechnics face systemic

barriers such as limited digital infrastructure, undertrained staff, and fragmented systems (Jayashree et al., 2023; Zian et al., 2024). Given polytechnics' central role in the TVET ecosystem, a focused review is necessary to assess their readiness and chart a path forward.

This paper reviews the state-of-the-art in big data management in Malaysian polytechnics, explores applications, identifies challenges, and outlines strategic recommendations. The analysis is framed using the TOE and DOI models to highlight both institutional and systemic dimensions of adoption.

2. Methodology

2.1 Database Selection

The literature search was conducted across multidisciplinary and domain-specific databases widely recognized in education and information technology research: Scopus, Web of Science (WoS), IEEE Xplore, SpringerLink, ScienceDirect, and Google Scholar. These databases were selected to capture both global research trends and Malaysia-specific studies relevant to higher education and Technical and Vocational Education and Training (TVET).

2.2 Search Strategy

A Boolean search strategy was designed to combine terms related to big data, education/TVET, and Malaysia. The search was conducted in January 2025, covering publications between 2012 and 2025 to reflect the decade of accelerated growth in big data adoption within education. Studies were required to focus on big data management, analytics, or adoption in higher education, polytechnics, or TVET, and could be conceptual, empirical, or framework-based, provided they addressed applications, challenges, or governance. Excluded from the review were studies unrelated to education (such as big data in healthcare or finance), non-empirical commentaries, editorials, duplicate records, and papers without accessible full text.

2.3 Screening and Eligibility

The initial search yielded 310 records. After removing 50 duplicates, 260 titles and abstracts were screened for relevance. A total of 150 records were excluded at this stage, leaving 110 full-text articles for eligibility assessment. Following detailed evaluation, 75 studies met the inclusion criteria and were retained for synthesis.

2.4 Data Extraction and Synthesis

Data from each study were extracted based on thematic focus (applications, challenges, governance, employability), theoretical frameworks (e.g., TOE, DOI, UTAUT), geographical scope, and type of contribution. A thematic synthesis was then conducted to compare global best practices with the current state of Malaysian polytechnics, highlighting gaps, challenges, and opportunities for improvement.

Table 1 : Comparison of global and Malaysian adoption of big data in higher education

Dimension	Global Practices	Malaysian Polytechnics
Adoption Level	Mature adoption in universities; integrated analytics platforms, data lakes	Early-stage adoption; mostly pilot projects, fragmented systems (e.g., CIDOS)
Theoretical Frameworks	TOE, DOI, UTAUT widely applied in adoption studies	Limited application; TOE–DOI considered in some studies but weak institutional use
Applications	Predictive analytics for dropout prevention, adaptive learning, smart campuses	Basic dashboards, student tracking, limited predictive models
Data Governance	Centralized strategies, strong institutional backing	Fragmented governance across ministries (TVET under multiple agencies)
Staff Competence	High training investment, established data literacy programs	Skills gap in data analytics; limited staff training opportunities
Industry Integration	Strong university–industry collaboration in labour market analytics	Limited collaboration; tracer studies exist but not fully integrated

Table 2 : Challenges and strategic solutions for Malaysian polytechnics

Challenge Category	Key Issues	Strategic Solutions
Technological	Poor infrastructure; fragmented systems (CIDOS, eKampus); low analytics maturity	Upgrade infrastructure; integrate systems; adopt cloud-based scalable platforms
Organizational	Staff lack data literacy; weak leadership; poor data-driven culture	Professional development; leadership training; create Data Governance Units
Environmental	Fragmented governance (multiple ministries); limited funding	Centralize under TVET 2030; secure dedicated funding streams; foster inter-polytechnic collaboration
Employability Data	Tracer studies underutilized; employability models not widely scaled	Integrate tracer data into real-time dashboards; deploy machine learning employability models
Research Integration	Limited analytics in teaching and research	Encourage industry partnerships; embed big data in polytechnic research and final-year projects

3. Results

3.1 State-of-the-Art in Big Data Management at Malaysian Polytechnics

Reviewing the current status of big data management techniques, this section highlights their theoretical underpinnings, national context, global development, and particular applications in Malaysian polytechnics.

a. Global Perspective

Big data adoption in education has expanded rapidly through learning analytics, educational data mining, and dashboards. These tools enable monitoring, personalized instruction, and planning. Platforms such as Hadoop and Spark, with AI/ML, support predictive models for dropout risk (Rani et al., 2024).

b. Theoretical Foundations

The TOE framework together with the Diffusion of Innovation (DOI) theory serve as common research tools to study institutional big data adoption and management practices. The TOE model examines three factors which include technological readiness and internal organizational elements and external pressures while the DOI model emphasizes innovation characteristics such as relative advantage and compatibility and observability (Jayashree et al., 2023). (Harun et al., 2022) showed that Malaysian educational institutions face two essential technological challenges regarding big data adoption which include compatibility and data security. These models provide a useful starting point for assessing the readiness of Malaysian polytechnics.

c. Big Data in Malaysian Higher Education Institutions

The Ministry of Higher Education (MoHE) in Malaysia has established digital transformation as a strategic priority through the Malaysia Education Blueprint 2015–2025. The implementation of institutional data lakes and academic analytics dashboards and AI-based assessment tools by several public universities has improved both administrative and academic performance according to (Zian et al., 2024). The implementation of big data in Malaysian polytechnics exists at an initial stage despite recent advancements. According to (Jayashree et al., 2023) universities gain better infrastructure and policy support but polytechnics encounter systemic barriers including insufficient digital infrastructure and untrained staff and disconnected systems.

d. State-of-the-Art in Malaysian Polytechnics

The Technical and Vocational Education and Training (TVET) sector's key institutions Malaysian polytechnics are adopting big data integration according to the TVET Digitalisation Strategic Plan 2021–2025. The sector shows some advancement in administrative analytics yet the complete implementation of big data remains restricted and scattered. The Learning Management Systems (LMS) at polytechnic institutions have started to incorporate analytics features through CIDOS. The features remain underutilized because staff members lack proper training and face organizational obstacles.

(Ridzuan & Alina Ismail, n.d.) performed a review to determine the readiness of TVET institutions for learning analytics which revealed a general deficiency in analytics adoption maturity. The institutions recognize the potential advantages of learning analytics yet encounter obstacles because of inadequate data systems and insufficient staff training. The ALERT model at Temasek Polytechnic in Singapore serves as a benchmark because it demonstrates how predictive tracking and real-time student engagement data enables academic intervention. (Azmi et al., 2024) demonstrated that institutional culture together with staff training and system integration form essential elements for successful model implementation. The current state-of-the-art in Malaysian polytechnics exists in an emergent state but remains underdeveloped. The majority of initiatives exist at the pilot level without achieving systemic integration. A national framework specifically designed for polytechnics needs to be developed to achieve full-scale transformation while following the TOE–DOI model and considering institutional realities throughout Malaysia's TVET ecosystem.

3.2 Applications of Big Data in Malaysian Polytechnics

Malaysian polytechnics are increasingly using big data to improve educational quality and operational efficiency and graduate employability in line with the national IR4.0 agenda. Although the adoption is still evolving, several key application areas can be identified based on current practices, pilot studies, and institutional strategies.

a. Academic Analytics and Student Monitoring

Academic analytics stands as a promising big data application in polytechnics because it uses student performance data to develop teaching and learning interventions. The national LMS CIDOS (Curriculum Information Document Online System) provides basic analytics tools to track student login activity and assignment submissions and quiz performance. The analytics tools remain underutilized because staff members lack awareness about them and have insufficient capabilities (Ridzuan & Alina Ismail, n.d.). Several educational institutions have started testing predictive models which use attendance records and assessment patterns and learning activity participation to forecast student failure and dropout risks. The experimental nature of these initiatives shows how real-time intervention can enhance student retention and academic performance (Md Hani et al., 2024).

b. Administrative and Institutional Planning

The implementation of big data enables institutions to make better decisions. Polytechnics gather extensive data regarding student enrollment and staff numbers and budgetary decisions and resource distribution. The effective use of this data enables organizations to create better forecasts and plans. The research by (Azmi et al., 2024) indicates that polytechnic institutions primarily use analytics for administrative purposes through student headcount dashboards and course evaluation summaries and scheduling analysis. The tools enable administrators to base their decisions on evidence yet their effectiveness remains restricted because of separate data systems and non-existent centralized data governance framework.

c. Curriculum and Programme Review

The analysis of big data enables institutions to develop curricula and evaluate their programs. Polytechnics can match their courses to labour market requirements through analysis of graduate employability rates and industry feedback and work-based learning (WBL) performance. The Department of Polytechnic Education (JPPKK) uses graduate tracer studies and industry satisfaction surveys to direct improvements in curriculum design. The TVET Transformation Plan supports data-driven curriculum reform to achieve industrial competency standards according to the (Da, 2025).

d. Research and Innovation

Polytechnics conduct research at lower levels than research universities yet they apply big data analytics extensively in engineering and smart manufacturing and IoT applications. The use of big data in predictive maintenance of smart labs and energy optimization through machine learning is being investigated through collaborative research between polytechnics and universities and industry partners. The initiatives develop institutional capacity while offering students hands-on learning experiences in data-intensive settings (Azmi et al., 2024).

e. Graduate Employability and Career Services

Graduate outcome analytics represents a developing application field. Polytechnics can evaluate their graduate employability performance through linking institutional data with job market intelligence to make strategic adjustments. The Graduate Employability Management System (GEMS) and social media analytics serve as tested initiatives to track alumni progress while detecting skill gaps. The data insights enable institutions to customize their career support services and build stronger relationships with industry partners (Zian et al., 2024). (*Azida - Malaysian Community College Graduates Employability Prediction Model Using Machine Learning Approach*, n.d.) created an employability prediction model for Malaysian TVET graduates through machine learning methods including logistic regression and neural networks and random forest. The model based on 10,000 tracer study records reached 84.8% prediction accuracy through random forest algorithm implementation. Predictive analytics proves useful in practice for detecting employment patterns which guide institutional strategic decisions. Polytechnics along with similar institutions can use their historical graduate data to create better curriculum alignment and industry partnerships which enhance graduate outcomes according to their research.

3.3 Challenges in Big Data Management at Malaysian Polytechnics

While Malaysian polytechnics are initiating big data practices, they face significant **technology**, **organization**, and **environmental** challenges. These collectively slow down adoption and limit impact.

a. Infrastructure Limitations

The research conducted by (Zian et al., 2024) shows that technological factors, particularly inadequate infrastructure such as low processing power and poor data integration capabilities, are the main barriers to big data adoption in Malaysia.

b. Data Quality and Integration Issues

Use of various LMS, administration, and finance platforms poses challenges for consolidated utilization of data analytics. For Malaysian TVET institutions, there are challenges for consolidated use of data because CIDOS and eKampus are separate systems and inhibit efficiency for analytics (Ridzuan & Alina Ismail, n.d.) (Jayashree et al., 2023).

c. System Maturity & Analytics Quality

The 2023 systematic review "The Role of Big Data Analytics on Improving TVET Outcomes" reports that most intervention programs are small and have limited strong analytics maturity. The report found organisational and technological gaps and determined that literacy in using data is dependent on strong data-use culture.

d. Lack of Skilled Personnel

The research conducted by (Md Hani et al., 2024) demonstrated that TVET educators show significant digital competence shortcomings. The lack of trained staff who understand data literacy prevents analytics tools from reaching their full potential.

e. Limited Leadership and Data Culture

Staff members avoid adopting analytics because the organization lacks a data-driven culture and receives insufficient leadership backing. According to (Harun et al., 2022) educational institutions need strong organizational readiness and cultural support to successfully integrate big data systems.

f. Digital Competence Among Staff

The 2023 study of polytechnic students who used institutional repositories showed that students actively used digital tools but staff members lacked both training and confidence in advanced analytics which limited technology adoption (Sani et al., 2024).

g. Policy and Governance Fragmentation

Malaysian TVET is directed by three main ministries and agencies namely Ministry of Higher Education (MoHE), Ministry of Education (MoE), and Ministry of Human Resources (MoHR). Decentralized governance exemplified by (Kamal et al., n.d.) causes duplications of programs and non-uniform accreditation standards and blurred institution functions that present great challenges for harmonized governance of data.
The Dasar TVET Negara 2030 recognizes these structural gaps and calls for the creation of a central accreditation body, a unified data platform, and a national data governance policy to support systematic digital transformation (Da, 2025).

h. Funding Constraints

Polytechnics face financial limitations which prevent them from upgrading computer systems and hiring analytics staff or buying software licenses. The limited financial resources prevent institutions from investing in complete BDA ecosystems according to (Zian et al., 2024) and (Azmi et al., 2024). The limited budget prevents institutions from acquiring analytics platforms and skilled personnel while also delaying the update of organizational digital infrastructure which slows down the adoption of broad-based systematic big data initiatives. Small and rural polytechnic campuses experience these constraints repeatedly because they lack the financial means to implement scalable data-driven solutions.

4. Discussion

The integration of big data in Malaysian polytechnics remains in its early stages, but growing national interest in digital transformation, TVET reform, and data governance provides fertile ground for future development. This chapter outlines key trends expected to shape big data adoption in the polytechnic context and offers strategic recommendations for successful implementation.

a. AI-Powered Learning Analytics

The future of large education data will increasingly depend on artificial intelligence (AI) and machine learning (ML) technologies. These technologies will be used to track student engagement in real time and to develop predictive models of academic achievement and adaptive feedback systems. Machine learning predictive models are very good at predicting employability as (*Azida - Malaysian Community College Graduates Employability Prediction Model Using Machine Learning Approach*, n.d.) work shows—an end use potential that could be further extended across polytechnics.

b. Smart TVET Campuses and IoT Integration

Smart campus initiatives—such as IoT sensors, biometric attendance, and adaptive learning spaces—are being piloted. These generate valuable data on student behavior and institutional operations (Rani et al., 2024).

c. Industry Collaboration and Workforce Analytics

Polytechnics are expected to increasingly collaborate with industry partners for graduate tracking, internship analytics, and demand forecasting. Real-time labour market dashboards and AI-matched job placement systems will support curriculum responsiveness and graduate readiness (Zian et al., 2024).

d. Establish a Centralized Data Governance Unit

A Data Governance Unit (DGU) should be created by each polytechnic to handle data policy compliance and analytics system integration and inter-agency data alignment. The proposed national governance model in TVET 2030 should have this unit operating under it.

e. Educators and Data Stewards need professional development

The research of (Md Hani et al., 2024) indicates that targeted professional development should be implemented to build digital competence among educators and administrative staff. The training should focus on data interpretation and ethical analytics use and dashboard design capabilities.

f. Launch Pilot Projects for Learning Analytics

Polytechnics should initiate their analytics journey through small-scale pilot projects to enhance trialability and observability. The initial analytics projects should concentrate on developing early alert systems and tracking assignment completion rates and student class participation metrics. The organization will gain confidence in scaling up after publishing internal results.

g. Integrate Graduate Tracer Data with Analytics Tools

The integration of graduate tracer studies with analytics dashboards should replace traditional static surveys to provide continuous employability insights. The institutions can use this data to identify programme-level gaps and align with national skills development priorities (Azmi et al., 2023; (*Azida - Malaysian Community College Graduates Employability Prediction Model Using Machine Learning Approach*, n.d.).

h. Promote Inter-Polytechnic Data Collaboration

The establishment of a TVET Data Sharing Consortium would allow polytechnics to share data for benchmarking purposes and resource sharing and innovation transfer. The sharing of datasets including dropout rates and teaching workload and assessment trends will enable institutions to develop policies through data analysis.

5. Conclusion

Higher education institutions use big data to transform their learning management and operational processes and strategic planning approaches. The review examined Malaysian polytechnics which lead the TVET sector to assess their current big data technology adoption status and future development potential. Polytechnics lag behind global institutions and local public universities in analytics and smart campus solution adoption because they face limited infrastructure and digital skills and fragmented systems. The paper used TOE–DOI frameworks to identify three main barriers which include poor integration and underdeveloped data cultures and funding constraints. National initiatives such as the TVET Digitalisation Strategic Plan and Dasar TVET Negara 2030 show increasing dedication toward digital transformation. The future development of education will depend on artificial intelligence implementation and centralized data governance and stronger industry partnerships. Polytechnics need to achieve infrastructure and governance improvements and staff development investments and analytics project piloting and strategic review integration of graduate employability data and inter-

institutional collaboration development to fully benefit from big data. Malaysian polytechnics can develop into agile data-driven institutions through national support and clear direction to fulfill education and workforce demands.

References

Azida Mansor, & Othman, Z. (2024). Malaysian community college graduates employability prediction model using machine learning approach. *Journal of Telecommunication, Electronic and Computer Engineering, 16*(3), 1–7. https://doi.org/10.54554/jtec.2024.16.03.001

Azmi, N. A., Ismail, N. A. A., & Rosman, A. F. (2024). The adoption and impact of building information modelling (BIM) towards facilities management (FM) in Malaysia. *Malaysian Journal of Sustainable Environment, 11*(3), 241–266. https://doi.org/10.24191/myse.v11i3.3926

Da, W. C. (2025). TVET in Malaysia's human resource development: Plans, realities and "game changers." *Journal/Proceedings details pending*.

Harun, N. B., Jalil, H. A., & Zolkepli, M. (2022). Technological, organizational and environmental factors influencing user intention towards big data technology adoption in Malaysian educational organizations. *Accounting, 8*(4), 403–408. https://doi.org/10.5267/j.ac.2022.6.002

Jayashree, S., Hassan Reza, M. N., Malarvizhi, C. A., & Alias, M. B. (2023). Big data analytics and educational sustainability: A Malaysian scenario. In *2023 2nd International Conference on Electrical, Electronics, Information and Communication Technologies (ICEEICT)*. IEEE. https://doi.org/10.1109/ICEEICT56924.2023.10157261

Kamal, M., Ahmad, F., & Rosnan, H. (n.d.). Overcoming challenges in Malaysia's technical and vocational education: A path forward for TVET. *International Journal of Research and Innovation in Social Science*. https://doi.org/10.47772/IJRISS

Md Hani, M. H., Ismail, A., Razali, S. S., Abu Bakar, M. N., Rosman, N. L., Nasir, N., & Mohamad, Z. (2024). The development of digital competence model among TVET educators towards digitization: A concept paper. *Journal of TVET and Technology Review, 2*(1), 1–9. https://doi.org/10.30880/jttr.2024.02.01.001

Rani, M., Kumar, A., Sisodia, D. R., Perwej, Y., Kakde, A. C., Rakhimjonovna, F. M., & Professor, A. (2024). Transforming education through IoT and AI: Opportunities and challenges. *Educational Administration: Theory and Practice, 30*(5), 11610–11622. https://doi.org/10.53555/kuey.v30i5.4982

Ridzuan, F., & Ismail, N. A. (n.d.). Exploring TVET higher education towards learning analytics in Malaysia. *ResearchGate*. https://www.researchgate.net/publication/374923202

Samsul, S. A., Yahaya, N., & Abuhassna, H. (2023). Education big data and learning analytics: A bibliometric analysis. *Humanities and Social Sciences Communications, 10*(1), 1–13. https://doi.org/10.1057/s41599-023-02176-x

Sani, K., Ali, J. K., Sahid, N. Z., Yacob, Y., & Lajuni, N. (2024). Integrating technology by Malaysian polytechnic students in utilizing institutional repositories: A continuous study employing the UTAUT framework. *International Journal of Academic Research in Business and Social Sciences, 14*(2), 143–160. https://doi.org/10.6007/ijarbss/v14-i2/20837

Zian, L. Q., Zulkarnain, N. Z., & Kumar, Y. J. (2024). Challenges in big data adoption for Malaysian organizations: A review. *Indonesian Journal of Electrical Engineering and Computer Science, 33*(1), 507–517. https://doi.org/10.11591/ijeecs.v33.i1.pp507-517