E-ISSN: 2708-4523
P-ISSN: 2708-4515
AJMC 2025; 6(2): 205-209

www.allcommercejournal.com
Received: 16-05-2025
Accepted: 18-06-2025

Venkatesh K Iyengar
Research Scholar, Jamnalal Bajaj Institute of Management Studies & Research, Mumbai, Maharashtra, India

Dr. Sarika Mahajan
Programme Coordinator, Jamnalal Bajaj Institute of Management Studies & Research, Mumbai, Maharashtra, India

Corresponding Author:
Venkatesh K Iyengar
Research Scholar, Jamnalal Bajaj Institute of Management Studies & Research, Mumbai, Maharashtra, India

Unlocking data science adoption in MSMEs: A systematic literature review using the SPAR-4-SLR based systematic review and thematic analysis

Venkatesh K Iyengar and Sarika Mahajan

DOI: https://www.doi.org/10.22271/27084515.2025.v6.i2c.651

Abstract
This study presents a comprehensive Systematic Literature Review (SLR) aimed at identifying and synthesizing the key factors influencing the adoption of data science among Micro, Small, and Medium Enterprises (MSMEs). Drawing upon 192 peer-reviewed articles published between 2018 and 2024, the review adopts the SPAR-4-SLR framework a structured, transparent, and replicable protocol for high-quality literature synthesis alongside a structured thematic classification. The analysis identifies seven core thematic clusters: Technology adoption models, digital transformation, business intelligence and big data analytics, Industry 4.0/5.0, emerging technologies, trust in technology adoption, and methodological tools such as PLS-SEM. The review uncovers significant conceptual, methodological, and contextual gaps in current scholarship, offering actionable insights and future research directions. This work enhances understanding of the dynamics shaping data science adoption in MSMEs and underscores the need for context-sensitive, strategic interventions.

Keywords: SPAR-4-SLR, MSMEs, data science, systematic literature review, digital transformation, technology adoption

1. Introduction
The adoption of data science in MSMEs is a growing area of interest with significant implications for competitiveness, innovation, and digital resilience. Despite its advantages, adoption remains limited due to strategic, operational, and cultural barriers. To investigate this challenge, the paper conducts a rigorous Systematic Literature Review (SLR) to evaluate and synthesize existing research, uncover knowledge gaps, and inform future inquiry.
SLRs are structured, replicable, and transparent methods of reviewing literature that contribute meaningfully to theory and practice. They are typically categorized as domain-based, theory-based, or method-based. Domain-based reviews most prevalent in management and technology research include subtypes such as structured, framework-driven, bibliometric, hybrid, and model-building approaches, often guided by frameworks like ADO, TCCM, and 6W. Theory-based reviews examine literature through specific theoretical lenses, while method-based reviews focus on particular research techniques. Meta-analytical reviews, a quantitative extension, statistically synthesize findings to reveal broader patterns. A summary of various SLR approaches is presented in Table 1.
In order to structure and analyse the selected literature effectively, this study references five widely recognized SLR frameworks ADO, TCCM, 6W, TCM, and 7P. Each offers a unique analytical lens: ADO supports causal structuring through Antecedents Decisions Outcomes; TCCM guides methodological and theoretical mapping; 6W enables contextual and situational exploration; TCM aids in identifying narrative-based future research directions; and 7P facilitates a holistic, qualitative review.
These frameworks collectively enhance the rigor and comprehensiveness of the thematic synthesis, as summarized in Table 2.
The primary objective of this study is to identify and analyzes gaps in the existing body of literature pertaining to the critical factors influencing the adoption of data science in Micro, Small, and Medium Enterprises (MSMEs).

To achieve this, the study employs the SPAR-4-SLR protocol (Paul & Criado, 2020) [7], ensuring a structured, transparent, and replicable approach to synthesizing scholarly insights. A total of 192 peer-reviewed articles published between 2018 and 2024 form the basis of this review. The structure of the paper is as follows: Section 2 outlines the research methodology; Section 3 presents the descriptive and thematic analyses; Section 4 identifies key research gaps and proposes future research directions; and Section 5 concludes with implications for academic research, policy formulation, and MSME practice.

Table 1: Overview of SLR types and subcategories

SLR Type	Definition / Focus	Subcategories / Key Frameworks
Domain-based	Focus on a specific topic area or subject domain	Thematic
→ Framework-based	Uses established frameworks (e.g., ADO, TCCM, 6W) to structure the review	ADO (Antecedents-Decisions-Outcomes) TCCM (Theory-Construct-Context-Methodology) 6W (Who, What, When...)
→ Structured Review	Organized around methodologies, theories, and constructs	Content often presented via structured tables/figures
→ Bibliometric	Uses statistical tools to map literature	Categorization by author, country, theory, method, journal, time
→ Hybrid	Combines narrative + framework or bibliometric + structured	TCM-guided narrative or mixed bibliometric & structured
→ Model/Framework Dev.	Builds new conceptual models/frameworks from literature synthesis	Theory-building without empirical testing
Theory-based	Synthesizes literature based on a specific theory	Examples: Transaction Cost Theory, Resource-Based Theory
Method-based	Synthesizes literature using a specific method	Examples: Event Study, Discriminant Validity Testing
Meta-analytical	Statistical aggregation of quantitative studies	Correlations, effect size, moderator impact

Table 2: Key frameworks in domain-based SLRS

Framework Name	Components	Purpose
ADO	Antecedents-Decisions-Outcomes	Causal/Decision-focused structure
TCCM	Theory-Construct-Characteristics-Method	Methodological and conceptual mapping
6W	Who-What-When-Where-Why-How	Contextual and situational exploration
TCM	Theory-Context-Method	Future research direction within narrative
7P	Problem, Purpose, Perspective, Process, Participants, Phenomena, Place	Holistic review framework

2. Methodology

2.1 Search Strategy and Study Selection

This study adheres to the SPAR-4-SLR protocol (Paul, Lim *et al*., 2021) [8] to ensure scientific rigor, transparency, and replicability in the systematic review process. The literature search was conducted using the ProQuest database suite, which included ABI/INFORM, e-book Central, and Publicly Available Content DB.

The primary search string incorporated a combination of keywords such as:

"SME" and "UTAUT2" and "Trust" and "Technology Adoption" and ("Data Science" or "Artificial Intelligence" or "Business Analytics" or "Machine Learning").

To enhance the quality and relevance of the literature, the following filters were applied:

- **Document type:** Peer-reviewed journal articles and conference proceedings
- **Language:** English
- **Publication period:** 2018 to 2024

The search and selection process followed a multi-stage refinement protocol (S1-S22) using Boolean logic. Each step progressively narrowed the results based on source type, time period, language, and field-specific criteria. Table 3 presents a summary of the search iterations and corresponding output at each stage.

Table 3: Summary of search strategy levels

Level	Search Approach	Results
S1	Basic search combining SME, UTAUT2, Trust, adoption of new technology with Data Science / AI / BA / ML	110,757
S2	S1 + filtered by source types (Scholarly Journals, Conference Papers, etc.) and PEER (yes)	110,757
S3	S2 + time period filter (2017-2023)	102,224
S4	S3 with narrower source filter	102,223
S5	S4 without 'Working Papers'	102,220
S6	Filtered for non-scholarly sources + Peers	855
S12	S7 + language filter: English	834
S13	Changed search field from full-text to summary; broader conceptual terms added	128,611
S14	S13 + filter for 'Conference Papers & Proceedings'	1,193
S17–S18	Focused on Technology Adoption as a summary term with filters	86,875
S19–S22	Narrowed version of S17 with conference papers, year/language filters	874
S23	Narrowed down version of S22-basis thematic approach	192

2.2 Thematic clusters derived from theory-driven and empirically recurrent dimensions in MSME data science adoption

To synthesize insights from the 192 selected studies, a structured thematic analysis was conducted, guided by the TCCM framework.

(Theory Context Construct Methodology). This approach enabled a comprehensive classification of the literature across conceptual foundations, empirical settings, key constructs, and methodological choices. The seven thematic clusters were derived from patterns in the literature and validated through key references, as summarized in Table 4. The Year-wise Distribution of Research Themes (2018-2024) is summarized in Table 5.

The themes identified in this review are grounded in both theoretical significance and empirical recurrence across the selected literature. Drawing from established models such as TAM, UTAUT, TOE, and RBV, as well as contemporary frameworks related to trust, digital transformation, and emerging technologies, these clusters represent the core dimensions influencing data science adoption in MSMEs. Each theme is supported by key references and reflects patterns consistently observed in studies published between 2018 and 2024.

2.2.1 Technology Adoption Models

Technology adoption models remain foundational in understanding how MSMEs adopt data science. Widely used models include TAM, UTAUT, TOE, and DOI (Venkatesh *et al*., 2003; Tornatzky & Fleischer, 1990) [2, 1]. Studies commonly investigate constructs such as perceived usefulness, ease of use, and behavioral intention, often relying on PLS-SEM and cross-sectional survey methods.

Gap: These models are rarely adapted to the unique contexts of MSMEs in developing economies, limiting their real-world applicability.

2.2.2 Digital Transformation in SMEs

This theme conceptualizes digital transformation as a strategic renewal process involving cultural, technological, and operational shifts. Studies use theoretical lenses such as Dynamic Capabilities and Digital Maturity Models and explore constructs like digital readiness and managerial leadership.

Gap: Longitudinal evidence on transformation pathways in micro and service-based MSMEs remains limited.

2.2.3 Business Intelligence and Big Data Analytics

Studies in this cluster utilize the Resource-Based View (RBV) to explore how MSMEs develop and deploy analytics capabilities for agility and performance. Constructs such as analytical culture, decision quality, and strategic alignment are prominent.

Gap: Few studies quantify return on investment (ROI), and actionable frameworks for MSMEs remain underdeveloped.

2.2.4 Industry 4.0 and 5.0

This theme investigates how MSMEs engage with cyber-physical systems, automation, and human-centric digital technologies. Theoretical grounding includes Sociotechnical Systems and Smart Manufacturing.

Gap: Research often overlooks informal and micro enterprises that face structural limitations in adopting advanced technologies.

2.2.5 AI, Blockchain, and Emerging Technologies

Emerging technologies such as AI, Blockchain, and IoT are increasingly relevant to MSME transformation. The TOE framework is widely used to study perceived risk, innovation diffusion, and technology readiness (Dwivedi *et al*., 2021) [9].

Gap: Empirical evidence on real-world adoption, scalability, and sector-specific implications is still sparse.

2.2.6 Trust in Technology Adoption

Trust is a pivotal behavioral factor in both internal and external adoption processes. Key constructs include system transparency, vendor credibility, and social influence, guided by Trust Theory and UTAUT2.

Gap: Studies insufficiently address how cultural and psychological trust dynamics shape technology acceptance in resource-constrained MSMEs.

2.2.7 Methodological Tools and PLS-SEM

PLS-SEM is the dominant analytic approach, particularly valued for its applicability in exploratory studies with smaller sample sizes (Hair *et al*., 2017; Sarstedt *et al*., 2022) [3, 10].

Gap: Inconsistent application particularly in reporting model fit, testing multi-group validity, and sample justification limits methodological robustness.

Table 4: TCCM mapping of thematic clusters

Thematic Cluster	Theory	Context	Constructs / Characteristics	Methodology
Technology Adoption Models	TAM, UTAUT, TOE, DOI	Developing countries, MSMEs, ICT sectors	Perceived usefulness, ease of use, behavioral intention	Surveys, PLS-SEM, cross-sectional studies
Digital Transformation in SMEs	Dynamic Capabilities, Digital Maturity	Manufacturing and service MSMEs	Digital readiness, innovation orientation, transformation barriers	Case studies, longitudinal analysis, hybrid reviews
BI & Big Data Analytics in SMEs	Resource-Based View (RBV)	SMEs across Asia and Europe	Analytics capabilities, data-driven culture, SME agility	Bibliometric reviews, survey-based analysis
Industry 4.0 / 5.0	Smart Manufacturing, Sociotechnical Systems	Automotive and electronics MSMEs	Cyber-physical systems, human-centricity, automation	Framework-based reviews, structured analyses
AI / Blockchain / Emerging Technologies	Technology–Organization–Environment (TOE)	Fintech, retail, logistics MSMEs	AI use, blockchain trust, perceived risk	Quantitative studies, interviews, mixed methods
Trust in Technology Adoption	Trust Theory, UTAUT2	Asian fintech and technology-adopting SMEs	Trust, security, transparency, social influence	PLS-SEM, structural modeling, surveys
PLS-SEM / Methods	Measurement Theory, Model Validation	Business and marketing research in MSMEs	Discriminant validity, reflective constructs, reliability	Meta-analysis, SmartPLS simulation studies

Table 5: Year-wise distribution of research themes (2018-2024)

Year	Tech Adoption Models	Digital Transformation	BI & Big Data Analytics	Industry 4.0 / 5.0	AI / Emerging Tech	Trust in Tech Adoption	PLS-SEM / Methods	Total
2018	9	7	5	4	4	7	5	41
2019	4	2	2	2	2	2	2	16
2020	6	4	3	3	3	2	4	25
2021	7	5	3	5	4	3	5	32
2022	10	8	4	6	6	6	8	48
2023	9	7	3	6	7	4	6	42
2024	3	2	0	1	3	1	1	11
Total	48	35	20	27	29	25	31	192

3. Discussion and Research Gaps

The findings of this review reveal several critical gaps in the literature on data science adoption among MSMEs:

- **Theoretical Gaps:** Heavy reliance on legacy frameworks (TAM, UTAUT, TOE and DOI) fails to address the complexity of MSME contexts. There is a lack of integrative or hybrid models that consider trust, digital culture, and organizational maturity.
- **Methodological Gaps:** Many studies use PLS-SEM without appropriate validation, robustness checks, or longitudinal follow-ups. This limits the ability to derive causal or temporal insights.
- **Construct-Level Gaps:** While constructs such as perceived usefulness and intention dominate, key outcomes like ROI, decision-making quality, data governance, and employee data literacy are underexplored.
- **Contextual Gaps:** MSMEs from developing economies, especially non-manufacturing sectors, are significantly underrepresented, affecting the generalizability of current findings.
- **Trust as a Cross-Cutting Construct:** Trust influences both external (e.g., vendor reliability) and internal (e.g., data handling by employees) adoption processes. In low-digital-literacy environments, trust must be treated as a core precondition rather than a secondary variable.

4. Thematic synthesis using the TCCM framework

This review applies the TCCM framework to structure insights across, Theory, Context, Construct and Method

- **Theory:** Most studies use traditional models such as TAM, TOE, or RBV, but few engage with emerging MSME-specific perspectives like trust-centric or capability-building frameworks. Develop integrative models reflecting trust, absorptive capacity, and data ethics.
- **Construct:** Commonly studied variables include perceived ease of use, digital readiness, and top management support. However, constructs central to data science like data quality, trust, and decision-making effectiveness are insufficiently covered.
- **Context:** The literature is biased toward studies from Europe and Asia, with minimal research on Africa, Latin America, or informal MSMEs. Digitally immature and service-based firms are notably neglected.
- **Method:** Despite the widespread use of PLS-SEM, methodological transparency is often lacking. There is an urgent need for longitudinal tracking, multi-group analysis, and replication studies.

5. Recommendations and future research directions

In light of the identified gaps, the following research directions are proposed:

- **Model Innovation:** Develop integrated adoption models incorporating trust, digital readiness, and data governance tailored to MSME contexts.
- **Broaden Context:** Include underrepresented regions and non-manufacturing sectors in empirical research to improve generalizability.
- **Longitudinal Impact Assessment**: Conduct studies tracking data science ROI, adoption intensity, and transformation outcomes over time.
- **Trust Frameworks:** Explicitly integrate trust as both a mediator and moderator in data science adoption models.
- **Methodological Rigor:** Enhance PLS-SEM transparency with detailed reporting on model fit, reliability, sample logic, and cross-group comparisons.
- **Emerging Concerns:** Investigate data ethics, AI explainability, and vendor trustworthiness as new determinants influencing MSME adoption behavior.

6. Conclusion

This study offers a structured synthesis of 192 peer-reviewed papers (2018-2024) on data science adoption in MSMEs. Using the SPAR-4-SLR protocol and the TCCM framework, it identifies seven recurring themes: technology adoption models, digital transformation, big data and BI, Industry 4.0/5.0, emerging technologies, trust, and methodological tools like PLS-SEM. The review exposes critical gaps overreliance on legacy theories, limited methodological rigor and focus on intention over outcomes, and narrow coverage of sectors and geographies. It calls for context-sensitive models that integrate trust, digital readiness, and governance, and urges future research to embrace diverse methods, regions, and emerging concerns like AI explain ability and data ethics. This review advances theoretical and practical understanding, offering a foundation for future inquiry and action.

7. References

1. Tornatzky LG, Fleischer M. The processes of technological innovation. Lexington: Lexington Books, 1990.
2. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: Toward a unified view. MIS Q. 2003;27(3):425-478. https://doi.org/10.2307/30036540
3. Hair JF, Hult GTM, Ringle C, Sarstedt M. A primer on partial least squares structural equation modeling (PLS-

SEM). 2nd Ed, Thousand Oaks: SAGE Publications, 2017.
4. Li L, Su F, Zhang W, Mao JY. Digital transformation by SME entrepreneurs: A capability perspective. Inf Syst J. 2018;28(6):1129-1157.
5. Tamilmani K, Rana NP, Dwivedi YK. A systematic review of citations of UTAUT2 article and its usage trends. J Enterp Inf Manag. 2019;32(5):836-892. https://doi.org/10.1108/JEIM-06-2018-0136
6. Rozmi ANA, Bakar MIA, Hadi ARA, Nordin AI. Predicting ICT adoption among SMEs in Malaysia using UTAUT. Int J Adv Comput Sci Appl. 2020;11(4):222-230.
7. Paul J, Criado AR. The art of writing literature review: What do we know and what do we need to know? Int Bus Rev. 2020;29(4):101717. https://doi.org/10.1016/j.ibusrev.2020.101717
8. Paul J, Lim WM, O'Cass A, Hao AW, Bresciani S. Scientific procedures and rationales for systematic literature reviews (SPAR-4-SLR). Int J Consum Stud. 2021;45(4):01-15.
9. Dwivedi YK, Hughes DL, Coombs C, Constantiou I, Duan Y, Edwards JS, *et al*. Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. Int J Inf Manag. 2021;57:101994.
10. Sarstedt M, Ringle CM, Hair JF. PLS-SEM: Looking back and moving forward. Long Range Plann. 2022;55(3):101228. https://doi.org/10.1016/j.lrp.2021.102168