

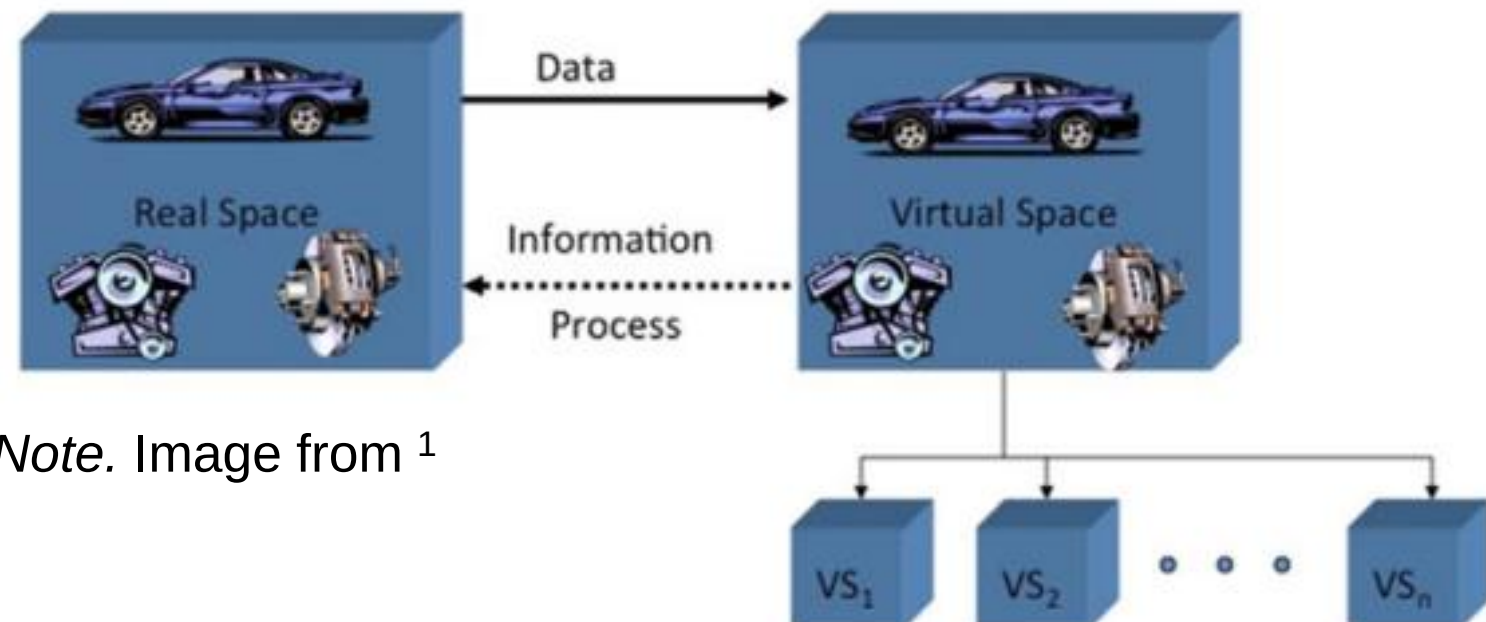
Digital Twin Roles in AI/ML Training, Testing, and Data-Driven Decision-Making: A Systematic Literature Review

Sheri Leder, Dr. Ana Wooley*, Dr. Vineetha Menon, Dr. Howard Chen, Gracie Gray, Dr. Bryan Mesmer, and Dr. Nicholas Loyd

*acw0047@uah.edu

Background & Motivation

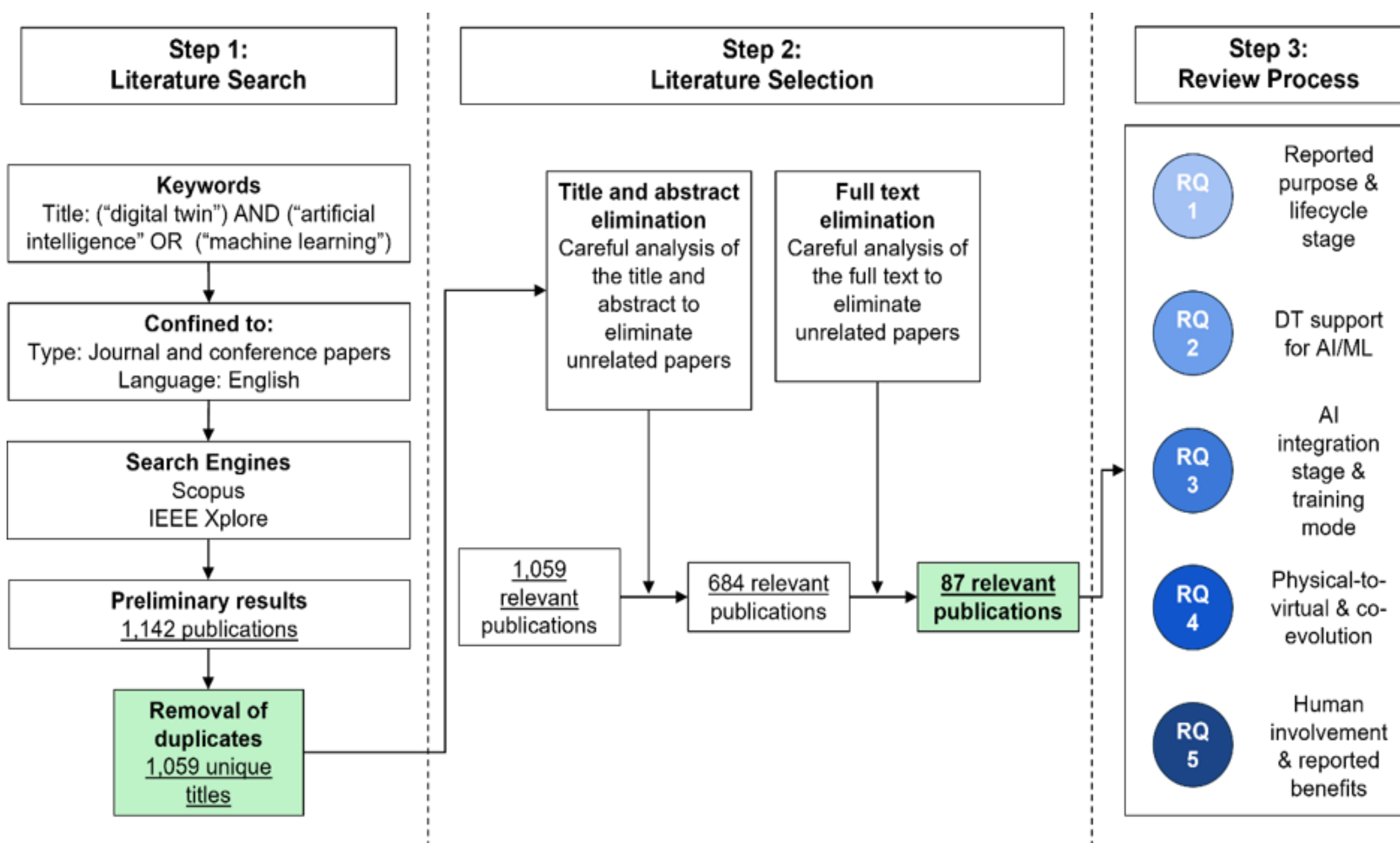
A digital twin (DT) is a virtual representation of an object, process, or system that spans its lifecycle and is composed of (1) physical space, (2) virtual space, and the (3) real-time data flow between the spaces.^{1, 2}



Note. Image from ¹

DTs are increasingly integrated with artificial intelligence (AI) and/ or machine learning (ML).³⁻⁶ DTs can enable AI/ML development and evaluation in virtual environments.⁷⁻⁹ Prior reviews mostly examine AI/ML characteristics (e.g., tasks, algorithms, modeling approaches).^{10, 11} However, less is known about how DTs support AI/ML training, testing, updating, and adaptation.

Research Questions and Methodology



1. What purposes motivate DT-AI/ML integration, and at what lifecycle stages are those systems implemented?
2. What are the roles DTs play in supporting AI/ML development and evaluation?
3. How is the AI/ML component incorporated, trained, and updated within DT-AI/ML systems?
4. What evidence is provided for physical-virtual connection and adaptive behavior in DT-AI/ML systems?
5. What evidence is provided for human involvement, and what benefits are reported for DT-AI/ML systems?

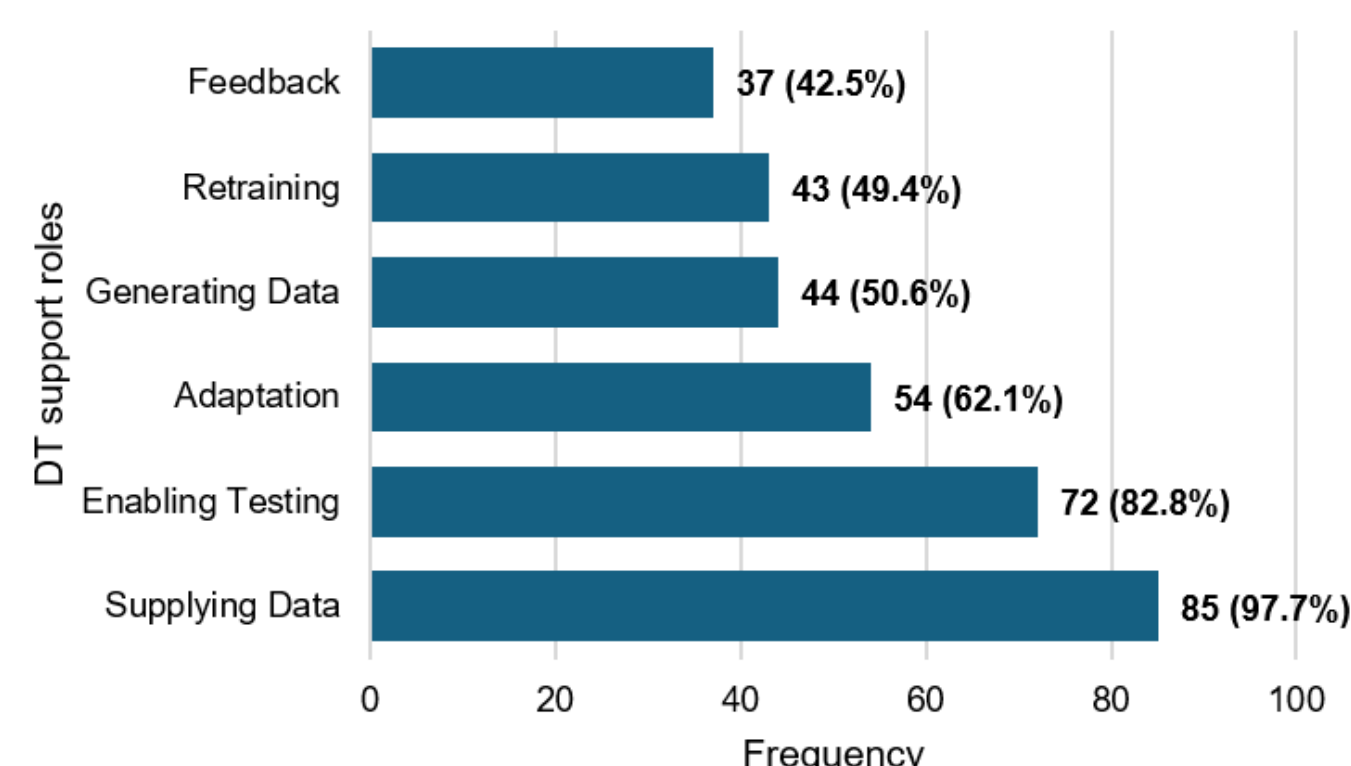
Results

RQ1. Reported purposes and lifecycle stages

Reported Purpose	Lifecycle Stage			
	Design	Production	Operation	Maintenance
Prediction (n = 50)	4 (8.0%)	9 (18.0%)	41 (82.0%)	24 (48.0%)
Optimization (n = 34)	3 (8.8%)	8 (23.5%)	29 (85.3%)	7 (20.6%)
Control (n = 22)	1 (4.5%)	5 (22.7%)	21 (95.5%)	5 (22.7%)
Training (n = 12)	1 (8.3%)	2 (16.7%)	11 (91.7%)	3 (25.0%)
Detection (n = 10)	—	2 (20.0%)	8 (80.0%)	6 (60.0%)
Diagnosis (n = 5)	—	—	3 (60.0%)	5 (100.00%)
Classification (n = 3)	1 (33.3%)	—	1 (33.3%)	2 (66.7%)

Key findings: Prediction and Optimization are the most common purposes and applications are strongly concentrated in the Operation stage.

RQ2. Roles of DT supporting AI/ML

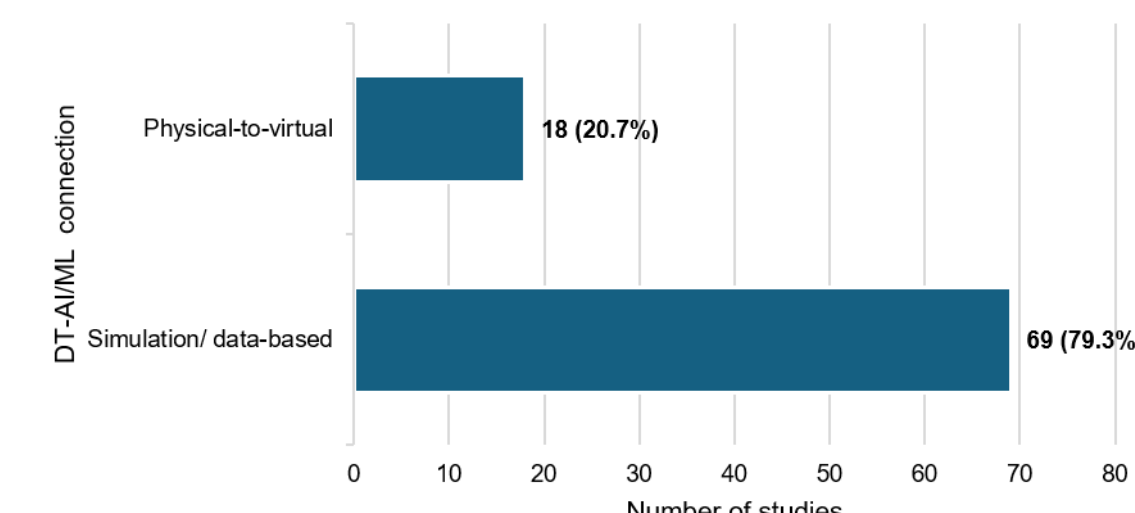


49 studies (56.3%) used the DT as both a training environment and testbed, while 38 studies (43.7%) used it as a training environment only

RQ3. AI/ML incorporation, training, and model updating

- Initial model development primarily offline: 60 studies (69%).
- Subsequent model updating not demonstrated: 38 studies (43.7%).
- AI/ML was incorporated during DT development in only: 9 studies 10.3%.
- Hybrid initial most often led to continuous updating: 7/11 (63.6%).

RQ4. Physical-virtual connection and adaptive behavior



Only 2 studies (5.3%) adequately demonstrated live co-evolution (i.e., continuous, bidirectional feedback loop).

RQ5. Human involvement and reported benefits

- Human involvement was reported in 22 studies (25.3%)
- Most common roles: decision making, monitoring/oversight, and setup/tuning.
- Main benefits: optimization/decision support, prediction/forecasting, fault or anomaly detection/diagnosis, and cost/time/resource reduction.

References

- [1] M. Grieves and J. Vickers, "Digital twin: Mitigating unpredictable, undesirable emergent behavior in complex systems," in Transdisciplinary Perspectives on Complex Systems: New Findings and Approaches. Cham, Switzerland: Springer, 2016, pp. 85–113.
- [2] M. Javadi, A. Haleem, and R. Suman, "Digital twin applications toward Industry 4.0: A review," Cogn. Robot., vol. 3, pp. 71–92, 2023.
- [3] F. Emmert-Streib, "What is the role of AI for digital twins?," AI, vol. 4, no. 3, pp. 721–728, 2023.
- [4] A. Fuller, Z. Fan, C. Day, and C. Barlow, "Digital twin: Enabling technologies, challenges and open research," IEEE Access, vol. 8, pp. 108952–108971, 2020.
- [5] R. Kaul, C. Ossai, A. R. M. Forkan, P. P. Jayaraman, J. Zelter, S. Vaughan, and N. Wickramasinghe, "The role of AI for developing digital twins in healthcare: The case of cancer care," WIREs Data Min. Knowl. Discov., vol. 13, no. 1, Art. no. e1480, 2023.
- [6] J. Lee, M. Azamfar, J. Singh, and S. Shihpour, "Integration of digital twin and deep learning in cyber-physical systems: Towards smart manufacturing," IET Collab. Intell. Manuf., vol. 2, no. 1, pp. 34–36, 2020.
- [7] U. Dahmen, T. Osterloh, and J. Roßmann, "Generation of virtual test scenarios for training and validation of AI-based systems," in Proc. 2021 IEEE Int. Conf. Progress Informatics and Computing (PIC), 2021, pp. 64–71.
- [8] P. V. Lopes, L. Silveira, R. D. Guimaraes Aquino, C. H. Ribeiro, A. Skoogh, and F. A. N. Verri, "Synthetic data generation for digital twins: Enabling production systems analysis in the absence of data," Int. J. Comput. Integr. Manuf., vol. 37, nos. 10–11, pp. 1252–1269, 2024.
- [9] D. L. Van Bossuyt, D. Allaire, J. F. Bickford, T. A. Bozada, W. Chen, R. P. Cutitta, R. Cuzner, K. Fletcher, R. Giachetti, and B. Hale, "The future of digital twin research and development," J. Comput. Inf. Sci. Eng., vol. 25, no. 8, Art. no. 080801, 2025.
- [10] T. Kreuzer, P. Papadimitrou, and J. Zdravkovic, "Artificial intelligence in digital twins—A systematic literature review," Data Knowl. Eng., vol. 151, Art. no. 102304, 2024, doi: 10.1016/j.datak.2024.102304.
- [11] M. M. Rathore, S. A. Shah, D. Shukla, E. Bentataf, and S. Bakiras, "The role of AI, machine learning, and big data in digital twinning: A systematic literature review, challenges, and opportunities," IEEE Access, vol. 9, pp. 32030–32052, 2021.

Funding



Research was sponsored by the Leonard Wood Institute (LWI) Army Research Laboratory and was accomplished under Cooperative Agreement Number W871NF2420158.