REVIEW

Evidence Boundaries and Scale-Up for Knowledge Extraction for Industrial Evidence: A Cross-Domain Methods Perspective

Jeffrey Davis, Douglas Wilson, Raymond Brown*

* Corresponding author: Raymond Brown

Abstract

This critical review examines evidence boundaries and scale-up in knowledge extraction for industrial evidence, with links to intelligent manufacturing and evidence governance. The organizing question is which measurements and assumptions remain valid when a laboratory component is translated into a production decision. Ten or more scholarly and user-supplied records are synthesized through a mechanism-to-decision framework spanning system boundaries, measurement, representation, evaluation, translation, and governance. No experiment, participant dataset, effect estimate, or production result is invented. Particular attention is given to component-level success being reported as end-to-end readiness. The review argues that credible translation requires source-level traceability, explicit validity domains, failure-aware evaluation, and revalidation triggers. Google Scholar-supplied records are retained in structured APA form, and no missing DOI or pagination field is completed by conjecture.

Keywords: artificial intelligence, systems, evidence synthesis, reliability, provenance

1. Introduction

Research on evidence boundaries and scale-up in knowledge extraction for industrial evidence, with links to intelligent manufacturing and evidence governance often begins with an apparently successful component. Practical value, however, depends on a longer chain: inputs must be measured reproducibly, representations must preserve decision-relevant variation, models or physical mechanisms must be evaluated under plausible shift, and outputs must reach an accountable decision maker with their uncertainty intact. The central question is therefore which measurements and assumptions remain valid when a laboratory component is translated into a production decision.

This article uses a structured narrative synthesis. It compares how the cited works define mechanisms, observations, computational representations, manufacturing conditions, and use constraints; it does not pool effect sizes across incompatible designs. This separation matters because a central risk is component-level success being reported as end-to-end readiness. A source is used only for the proposition it can directly support, while cross-domain implications are presented as design inferences rather than empirical findings.

2. System Boundaries and Evidence Architecture

A defensible system boundary includes material or data inputs, preparation and preprocessing, environmental and equipment conditions, the transformation mechanism, measurement, decision rules, and downstream outcomes. Excluding any of these elements can make a local improvement appear to be a system-level benefit. Boundary choices should therefore be reported as part of the result rather than treated as neutral setup.

Zhang et al. (2025) develop unsupervised anomaly detection for cloud-native microservices through cross-service temporal contrastive learning. In this synthesis, the work supplies mechanism-level evidence or a source-domain design analogy for evidence boundaries and scale-up in knowledge extraction for industrial evidence, with links to intelligent manufacturing and evidence governance. Where the source and application domains differ, the connection is methodological rather than empirical. The inference is bounded to the reported mechanism, dataset, or operating condition and does not by itself establish system readiness. Translation requires an explicit bridge among measurement, uncertainty, and the decision that follows.

Hollnagel, Woods, and Leveson (2006) define resilience as the capacity to anticipate, monitor, respond, and learn. In this synthesis, the work supplies mechanism-level evidence or a source-domain design analogy for evidence boundaries and scale-up in knowledge extraction for industrial evidence, with links to intelligent manufacturing and evidence governance. Where the source and application domains differ, the connection is methodological rather than empirical. The inference is bounded to the reported mechanism, dataset, or operating condition and does not by itself establish system readiness. Translation requires an explicit bridge among measurement, uncertainty, and the decision that follows.

Anderson (2020) treat dependable security as a lifecycle property of socio-technical systems. In this synthesis, the work supplies mechanism-level evidence or a source-domain design analogy for evidence boundaries and scale-up in knowledge extraction for industrial evidence, with links to intelligent manufacturing and evidence governance. Where the source and application domains differ, the connection is methodological rather than empirical. The inference is bounded to the reported mechanism, dataset, or operating condition and does not by itself establish system readiness. Translation requires an explicit bridge among measurement, uncertainty, and the decision that follows.

Together, these sources show that evidence architecture is not merely documentation. It determines which comparisons are valid, which variations remain visible, and where uncertainty can enter. A useful review asks whether the evidence preserves the distinctions that matter for the intended decision and whether a claimed analogy has a defensible physical, computational, or operational basis.

3. Mechanisms, Representation, and Process Design

Mechanistic reasoning should connect an intervention or design choice to an observable pathway. In materials systems this may involve transport, surface interaction, catalytic activity, separation, or electrochemical stability. In computational systems it may involve sequence generation, relational extraction, representation learning, or real-time control. In manufacturing it includes contamination pathways, batch variability, metrology, and feedback.

Vaswani et al. (2017) establish the attention architecture underlying modern sequence and multimodal models. In this synthesis, the work supplies comparative evaluation evidence or a source-domain design analogy for evidence boundaries and scale-up in knowledge extraction for industrial evidence, with links to intelligent manufacturing and evidence governance. Where the source and application domains differ, the connection is methodological rather than empirical. The inference is bounded to the reported mechanism, dataset, or operating condition and does not by itself establish system readiness. Translation requires an explicit bridge among measurement, uncertainty, and the decision that follows.

Kingma and Ba (2015) introduce adaptive moment estimation for stochastic optimization. In this synthesis, the work supplies comparative evaluation evidence or a source-domain design analogy for evidence boundaries and scale-up in knowledge extraction for industrial evidence, with links to intelligent manufacturing and evidence governance. Where the source and application domains differ, the connection is methodological rather than empirical. The inference is bounded to the reported mechanism, dataset, or operating condition and does not by itself establish system readiness. Translation requires an explicit bridge among measurement, uncertainty, and the decision that follows.

Hu et al. (2022) show how low-rank updates can adapt language models efficiently. In this synthesis, the work supplies comparative evaluation evidence or a source-domain design analogy for evidence boundaries and scale-up in knowledge extraction for industrial evidence, with links to intelligent manufacturing and evidence governance. Where the source and application domains differ, the connection is methodological rather than empirical. The inference is bounded to the reported mechanism, dataset, or operating condition and does not by itself establish system readiness. Translation requires an explicit bridge among measurement, uncertainty, and the decision that follows.

The common design principle is modularity with explicit interfaces. A component can perform correctly while the complete pathway fails because inputs drift, a proxy ceases to represent the intended construct, an update changes calibration, or a downstream user acts outside the validation domain. Interface tests should therefore include degraded modes, missing information, conflicting sensors, and conditions that require abstention or shutdown.

4. Evaluation, Calibration, and Error Analysis

Evaluation should follow the decision pathway instead of ending with one technical score. A layered plan covers analytical validity, robustness to plausible shift, calibration, decision utility, and monitoring after release. Average performance should be accompanied by error categories, regime-specific results, uncertainty intervals, and sensitivity to preprocessing, material batches, equipment settings, and comparator choice.

Rasmussen and Williams (2006) establish Gaussian processes as probabilistic models with explicit predictive uncertainty. In this synthesis, the work supplies translation evidence or a source-domain design analogy for evidence boundaries and scale-up in knowledge extraction for industrial evidence, with links to intelligent manufacturing and evidence governance. Where the source and application domains differ, the connection is methodological rather than empirical. The inference is bounded to the reported mechanism, dataset, or operating condition and does not by itself establish system readiness. Translation requires an explicit bridge among measurement, uncertainty, and the decision that follows.

Linkov and Trump (2019) connect resilience measurement with practical risk management. In this synthesis, the work supplies translation evidence or a source-domain design analogy for evidence boundaries and scale-up in knowledge extraction for industrial evidence, with links to intelligent manufacturing and evidence governance. Where the source and application domains differ, the connection is methodological rather than empirical. The inference is bounded to the reported mechanism, dataset, or operating condition and does not by itself establish system readiness. Translation requires an explicit bridge among measurement, uncertainty, and the decision that follows.

Error costs are asymmetric in decision support for knowledge extraction for industrial evidence. A missed degradation mode, unsafe operating recommendation, unstable ranking, contaminated batch, or false assurance cannot be exchanged at a universal rate. Thresholds and escalation rules should be connected to the operating context. When evidence is sparse or metadata is incomplete, the scientifically honest output is a bounded hypothesis or a request for additional measurement rather than a definitive conclusion.

5. Translation, Monitoring, and Governance

Translation requires evidence that evaluated conditions resemble the intended environment closely enough for the proposed decision. Dataset shift, material variability, sensor drift, environmental conditions, changing incentives, and software or equipment failures can each break that connection. A release decision should define its validity domain and identify observations that trigger revalidation.

Kipf and Welling (2017) provide a scalable formulation for graph representation learning. In this synthesis, the work supplies governance evidence or a source-domain design analogy for evidence boundaries and scale-up in knowledge extraction for industrial evidence, with links to intelligent manufacturing and evidence governance. Where the source and application domains differ, the connection is methodological rather than empirical. The inference is bounded to the reported mechanism, dataset, or operating condition and does not by itself establish system readiness. Translation requires an explicit bridge among measurement, uncertainty, and the decision that follows.

Radford et al. (2021) demonstrate transferable visual representations learned from language supervision. In this synthesis, the work supplies governance evidence or a source-domain design analogy for evidence boundaries and scale-up in knowledge extraction for industrial evidence, with links to intelligent manufacturing and evidence governance. Where the source and application domains differ, the connection is methodological rather than empirical. The inference is bounded to the reported mechanism, dataset, or operating condition and does not by itself establish system readiness. Translation requires an explicit bridge among measurement, uncertainty, and the decision that follows.

Governance is part of the technical architecture because it assigns authority and preserves recoverable records. Useful controls include versioned data and models, batch and sample provenance, decision logs, human override, escalation paths, and change thresholds. Each control should be tied to a named hazard and an accountable owner. Monitoring should test the evidence chain, not merely whether a service remains online or a headline metric remains stable.

6. Research and Reporting Priorities

Future studies should predefine the intended decision, primary outcomes, exclusions, and analysis plan when feasible. They should report missingness, sensitivity analyses, negative findings, and the cost of errors to differently situated stakeholders. Comparative studies need baselines that isolate the contribution of each added component. External validation should introduce meaningful chemical, environmental, temporal, institutional, or operational shifts.

Evidence packages should preserve citation and data provenance. Readers should be able to identify which source supports each mechanism, evaluation choice, and governance recommendation. Bibliographic fields should never be silently completed from conjecture. Authorship, title, venue, year, pagination or article number, and persistent identifiers should be checked against the best available record.

7. Conclusion

The literature supports a disciplined pathway for evidence boundaries and scale-up in knowledge extraction for industrial evidence, with links to intelligent manufacturing and evidence governance: define the decision, preserve evidence boundaries, test consequential failure modes, and connect uncertainty to control. The strongest contribution is not a universal ranking but a transparent account of what is known, what remains conditional, and what would justify the next stage of use in decision support for knowledge extraction for industrial evidence.

Declarations

Ethics approval and consent: Not applicable. This article synthesizes previously published or user-supplied records and reports no new human, animal, or identifiable participant data.

Funding: No external funding is reported for this synthesis.

Competing interests: No competing interests are reported.

AI-assisted writing: Generative AI supported drafting, source organization, and language editing. Accountable authors and editors must verify every claim, citation, and bibliographic record before publication.

Data, code, and materials availability: No new dataset or experimental code was generated. Sources are identified in the reference list.

References

Anderson, R. (2020). Security engineering: A guide to building dependable distributed systems (3rd ed.). Wiley.

Hollnagel, E., Woods, D. D., & Leveson, N. (Eds.). (2006). Resilience engineering: Concepts and precepts. Ashgate.

Hu, E. J., Shen, Y., Wallis, P., Allen-Zhu, Z., Li, Y., Wang, S., Wang, L., & Chen, W. (2022). LoRA: Low-rank adaptation of large language models. In International Conference on Learning Representations.

Kingma, D. P., & Ba, J. (2015). Adam: A method for stochastic optimization. In International Conference on Learning Representations.

Kipf, T. N., & Welling, M. (2017). Semi-supervised classification with graph convolutional networks. In International Conference on Learning Representations.

Linkov, I., & Trump, B. D. (2019). The science and practice of resilience. Springer. https://doi.org/10.1007/978-3-030-04565-4

Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G., Askell, A., Mishkin, P., Clark, J., Krueger, G., & Sutskever, I. (2021). Learning transferable visual models from natural language supervision. In Proceedings of the 38th International Conference on Machine Learning (pp. 8748-8763). PMLR.

Rasmussen, C. E., & Williams, C. K. I. (2006). Gaussian processes for machine learning. MIT Press.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. Advances in Neural Information Processing Systems, 30.

Zhang, Z., Liu, W., Tao, J., Zhu, H., Li, S., & Xiao, Y. (2025, December). Unsupervised anomaly detection in cloud-native microservices via cross-service temporal contrastive learning. In 2025 5th International Symposium on Artificial Intelligence and Big Data (AIBDF) (pp. 221-226). IEEE.