



IRIS

Insight, Record, & Intelligence System

Shaun J Davies (2317135)

Tata Steel UK

Research context

- Steel manufacturing presents significant Safety, Health & Environmental (SHE) risks to employees, communities, the environment and organisational reputation.
- These risks are managed through management systems, standards, procedures and digital technologies.
- Analysis and reporting of SHE trends can provide insights that support SHE management (SHEM).
- This project focuses on two digital reporting areas: incident reporting and employee competency reporting.
- Intended users are SHE representatives, managers, senior leaders, directors and employees across the organisation.



Research Problem



Tata Steel UK data is stored across separate third-party platforms, limiting integration between related datasets.



Incident data alone cannot show whether factors including training completions or staff and contractor competency are associated with incidents.



Existing systems create duplicated effort, inconsistent data quality and reduced analytical capability because of the limited integration between systems.

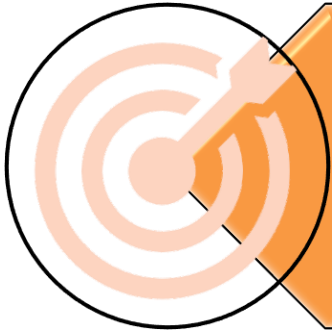


In both incident and competency reporting, the focus is largely descriptive and retrospective.

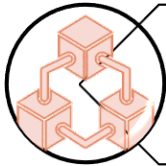


This drives the need to integrate and analyse key SHE datasets for a deeper insight and future AI enabled analysis.

Research Aim & Objectives



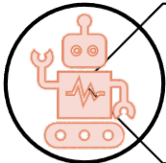
To design, develop and implement a proof of concept integrated SHE data solution by July 2027 that combines key datasets to support consolidated data management, analytical insight, and investigation of future SHE risks with a software implementation using SHE incident and competency data.



Design, develop and implement a consolidated data structure that integrates incident and competency data, applying appropriate data quality, standardisation, and governance controls.



Design, develop and implement an analytical solution that transforms the integrated data into standardised SHE performance measures, trends, and relationships across relevant sites and functions.



Develop and evaluate a predictive machine learning model using historical incident and competency data to determine whether identifiable patterns provide meaningful indicators of future SHE risk.

Literature Review

- SHEM has progressed from a focus on compliance into a performance based and proactive approach [1].
- Research shows that the steel industry can include fragmented SHEM systems across sites and local practices [2].
- Historical incident data can reveal recurring patterns, associations and risk factors but is often an underutilised analytical tool [5].
- The usefulness of incident data is heavily reliant on reporting quality, completeness, and data quality factors [7].
- Competency and training can improve factors related to SHEM but there is limited evidence of its effect on incident reductions [8].

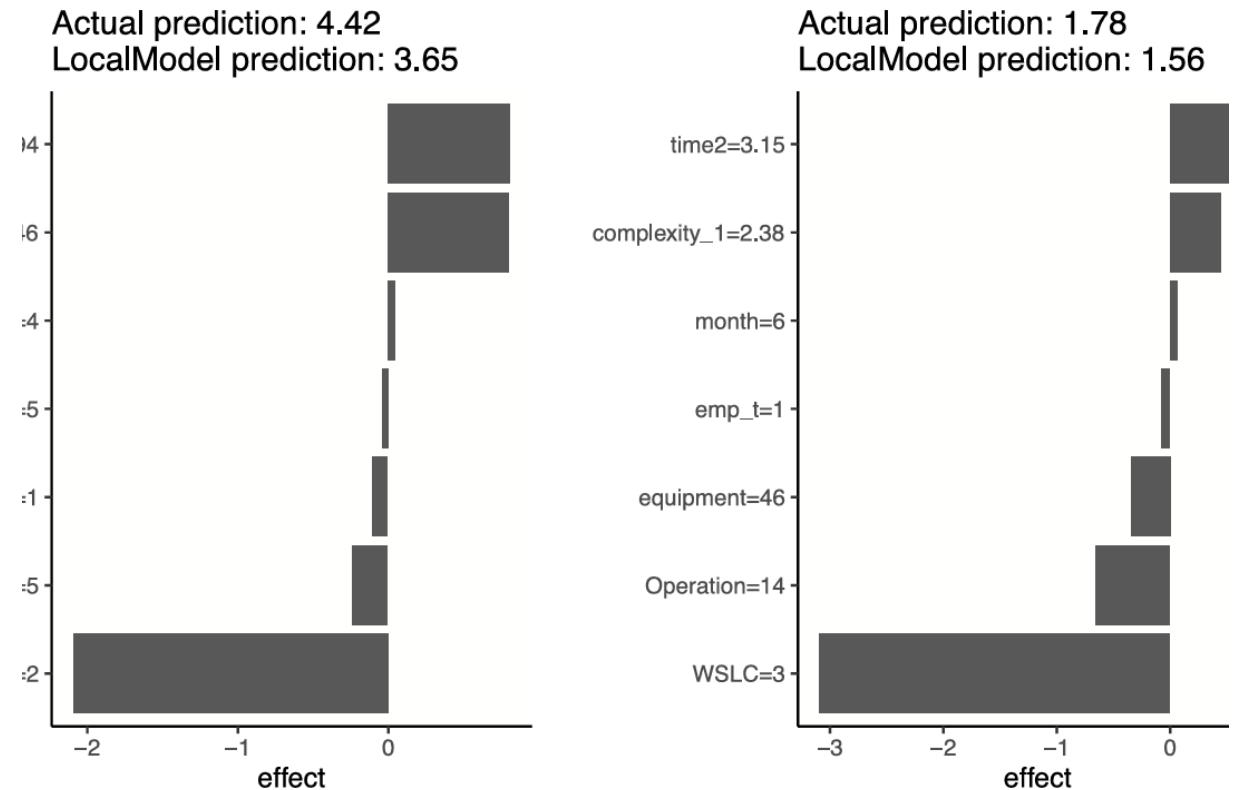
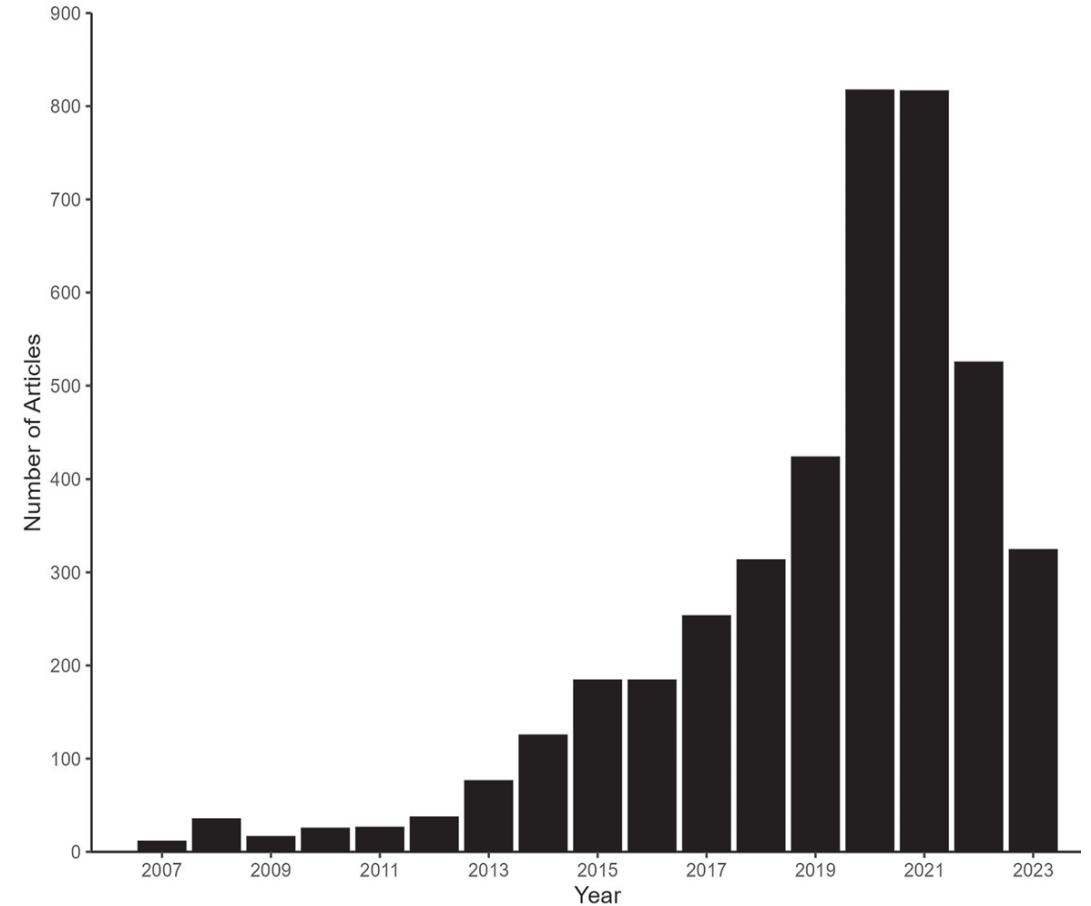


Fig. 5. Local model interpretation.

Literature Review

- The research shows a shift towards predictive analytics with adoption not yet mature [11].
- Only a limited proportion of the reviewed research contains proactive or leading indicators, such as “safety training history.” [11]
- Machine learning (ML) has been applied to injury classification, hazard prediction, accident prediction and risk identification [12].
- Sufficiently complete and consistent data is required to enhance predictive capability [11] [12] [15].
- The key research gap in the steel industry literature is that it does not demonstrate integration of workforce competency records with historical SHE incidents to analyse and investigate future SHE risk.

A. M. FOREMAN ET AL.



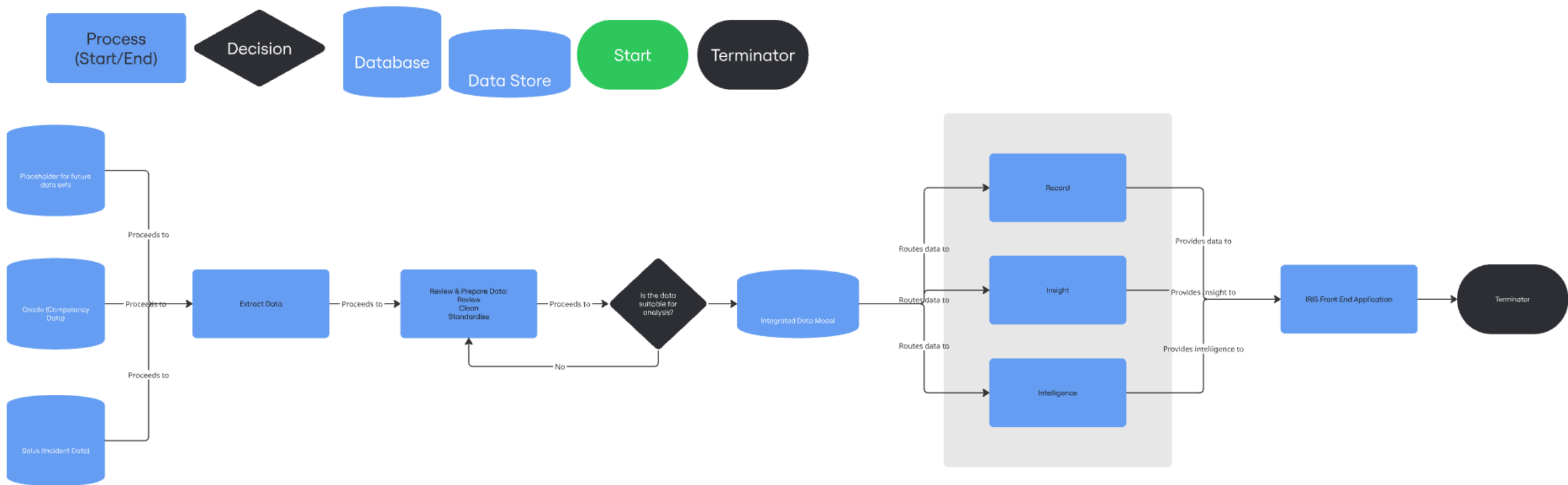
Number of Web of Science search results from 2007 to 2022 for the search 'safety analytics' in the article title, abstract and key words.



Research/Development Methodologies

- Five years of historical SHE incident and competency data including variables such as site, area, location, incident details and training records.
- Rules will be implemented to clean, validate and standardise the data. Once finalised, these rules will be documented and discussed with stakeholders for review [15].
- Descriptive, comparative and relationship analysis will be conducted to identify frequencies, trends, patterns and associations [17].
- Validated data will be split into training and testing datasets to investigate the suitability of logistic regression, decision trees, random forests and gradient boosting ML models [18].
- Predictive performance will be evaluated using the F1 score, ROC-AUC, and other methods followed by stakeholder evaluations of outputs [14].

Project Workflow



Resources/Materials

- Subject to access and business preference from business owners, existing Tata Steel UK software infrastructure and deployment platforms will be used (likely, Google Cloud Platform).
- Access to incident and competency datasets.
- Stakeholder support from SHE department including senior leaders, Learning & Development, IT & Data Privacy teams.
- University support for academic guidance and project supervision.
- No additional project funding is required. The planned financial budget is £0.



Project Plan

- Planning, design and stakeholder engagement: gather further requirements and stakeholder feedback.
- Implementation and system development: core feature development including data model, dashboards, ML implementation and testing.
- Deployment and user adoption: final security and stakeholder checks, deployment into a Tata environment, producing user guidance and workshops and initial user/admin support.
- Validation: confirm the platform has reached a stable operational state
- Review and evaluation: document project to capture if requirements have been met, review with stakeholders, document limitations and areas for future development.

NO.	TASK	START	END
1	Planning, Design & Stakeholder Engagement		
1.1	Confirm project scope, requirements, deliverables, resources, dependencies, risks and success criteria	01/10/2026	12/12/2026
1.2	Design system and data diagrams including UML, data flow, user stories, functional requirements and MVP definition	01/09/2026	15/09/2026
1.3	Engage with Heads of Departments, SHE stakeholders, IT, data owners, information security and policy stakeholders to review requirements and designs	15/09/2026	10/10/2026
1.4	Obtain approval and sign off with: scope, requirements, data access, security considerations, design, success metrics and KPIs	11/10/2026	12/12/2026
1.5	Design and build prototype	15/09/2026	10/10/2026
2	Implementation & System Development		
2.1	Establish the required development environment and technical architecture	12/12/2026	25/01/2027
2.2	Develop core IRIS platform with the required functionality	14/12/2026	25/01/2027
2.3	Implement data ingestion transformation, validation and storage processes	14/01/2027	25/01/2027
2.4	Develop data model, relationships, measure and analytical functionality	17/01/2027	25/01/2027
2.5	Develop user facing interfaces, dashboards and reporting functionality	23/01/2027	25/01/2027
2.6	Integrate Record, Insight, & Intelligence sections	23/01/2027	25/01/2027
2.7	Conduct a series of functional and non-functional testing	23/01/2027	25/01/2027
2.8	Resolve defects found in testing stage	23/01/2027	25/01/2027
3	Deployment & User Adoption		
3.1	Complete final technical, security and stakeholder checks	25/01/2027	25/02/2027
3.2	Deploy into the environment suited for Tata identified in planning stage	01/02/2027	06/02/2027
3.3	Produce user documentation, guides and supporting materials related to user training	07/02/2027	10/02/2027
3.4	Deliver workshops to relevant users and stakeholders	10/02/2027	27/02/2027
3.5	Provide initial user support including issues arising during deployment	15/02/2027	27/02/2027
4	Validation		
4.1	Monitor deployment against predefined requirements and key indicators developed in the planning stage	01/03/2027	09/03/2027
4.2	Validate data transformations, calculations and outputs through a series of reconciliation methods	01/03/2027	09/03/2027
4.3	Confirm system performance meets predefined standards	10/03/2027	14/03/2027
4.4	Implement necessary improvements	15/03/2027	27/03/2027
4.5	Confirm the platform has reached a stable operational state	01/04/2027	28/04/2027
5	Review & Evaluation		
5.1	Evaluate project against its aims, objectives and key indicators	01/05/2027	15/05/2027
5.2	Evaluate the functional and non-functional requirements have been met	01/05/2027	15/05/2027
5.3	Engage with stakeholders to gather their feedback	01/05/2027	30/05/2027
5.4	Document limitations, lessons learned and outstanding issues.	01/06/2027	25/06/2027
5.5	Evaluate recommendations and opportunities for further development	01/06/2027	25/06/2027

Project Impact & Limitations

IMPACT

- A consolidated view could create greater value from existing data.
- Analysis of both datasets could support more targeted resource allocation, training decisions and incident prevention discussions. An enterprise solution would assist SHEM through a connected view of major influencing factors of SHE through the use of data.
- The project supports a shift from predominantly reactive reporting to proactive SHEM.

LIMITATIONS

- The project does not represent a complete enterprise solution for SHE and Tata Group.
- Data availability and quality restricts the reliability of output from the proof of concept and enterprise solution.
- The relationship and associations between data may show correlation but cannot establish causation.
- Incident and competency data may be unrelated. The project may imply a connection between the data that is not reflective of reality.

Conclusion

- SHEM is moving towards an integrated analytical and proactive approach [1] [2] [3] [4].
- Current methods are separated by multiple isolated systems, limiting combined analysis and insight [2].
- The project will integrate incident and competency data to demonstrate the value of connecting previously isolated SHE data sources.
- The project will assess whether this connection will provide meaningful relationships and indicators of future SHE risk [8] [11] [12] [14].
- Successful delivery could provide a foundation for broader integrated SHE analytics and more proactive SHEM [1] [4] [11].



References

- [1] N. J. Duijm, C. Fie ´vez, M. Gerbec, U. Hauptmanns and M. Konstandinidou, "Management of health, safety and environment in process industry," *Safety Science*, vol. 46, no. 6, pp. 908-920, 2008.
- [2] M. Koivupalo, M. Sulasalmi, P. Rodrigo and V. Seppo, "Health and safety management in a changing organisation: Case study global steel company," *Safety Science*, vol. 74, pp. 128-139, 2015.
- [3] D. Huang, C. D. Dinga, Z. Wen and D. Razmadze, "Industrial-environmental management in China's iron and steel industry under multiple objectives and uncertainties," *Journal of Environmental Management*, vol. 310, p. 114785, 2022.
- [4] M. AlMarri, M. Al-Ali, M. Alzarooni, K. Al-Ali, Z. Bahroun and A. AlTeneiji, "Enterprise Resource Planning Systems for Health, Safety, and Environment Management: Analyzing Critical Success Factors," *Sustainability*, vol. 17, no. 7, pp. 1-29, 2025.
- [5] A. Verma, S. D. Khan, J. Maiti and O. Krishna, "Identifying patterns of safety related incidents in a steel plant using association rule mining of incident investigation reports," *Safety Science*, vol. 70, pp. 89-98, 2014.
- [6] G. Cao, L. Yang, Z. Ma, J. Wang, J. Bi and L. Liu, "Environmental incidents in China: Lessons from 2006 to 2015," *Science of The Total Environment*, vol. 633, pp. 1165-1172, 2018.
- [7] J. Guan, Y. Zixuan, A. P. Chan, T. Choi and Y. Yang, "Factors affecting learning from incidents: A cross-industry review," *Journal of Loss Prevention in the Process Industries*, vol. 89, pp. 1-17, 2024.
- [8] D. Hutchinson, G. Luria, S. Pindek and P. Spector, "The effects of industry risk level on safety training outcomes: A meta-analysis of intervention studies," *Safety Science*, vol. 152, p. 105594, 2022.
- [9] D. Pollitt, "Corus forges new approach to safety and health," *Human Resource Management International Digest*, vol. 19, no. 1, pp. 7-9, 2011.
- [10] L. H. Tarigan, M. Cifuentes, M. Quinn and D. Kriebel, "Prevention of Needle-Stick Injuries in Healthcare Facilities: A Meta-Analysis," *infection control & hospital epidemiology*, vol. 36, no. 7, pp. 823-829, 2015.
- [11] A. M. Foreman, J. E. Friedel, M. E. Ezerins, R. Matthews, R. E. Nicholson, L. Wellersdick, S. Bergman, Y. Açıkgoz, T. D. Ludwig and O. Wirth, "Establishment-level safety analytics: a scoping review," *International journal of occupational safety and ergonomics*, vol. 30, no. 2, pp. 559-570, 2024.
- [12] S. Maheronnaghsh, H. Zolfagharnasabb, M. Gorgich and J. Duarte, "Machine learning in Occupational Safety and Health – a systematic review," *International Journal of Occupational and Environmental Safety*, vol. 7, no. 1, pp. 14-32, 2023.
- [13] A. O. Oyedele, A. O. Ajayi and L. O. Oyedele, "Machine learning predictions for lost time injuries in power transmission and distribution projects," *Machine Learning with Applications*, vol. 6, p. 100158, 2021.
- [14] S. Pandey, A. K. Singh, S. Parhi and S. K. Jha, "Towards safer steel operations with a multi model framework for accident prediction and risk assessment simulation," *Scientific Reports*, vol. 15, no. 1, p. 13293, 2025.
- [15] Q. Fu, G. L. Nicholson and J. M. Easton, "Understanding data quality in a data-driven industry context: Insights from the fundamentals," *Journal of Industrial Information Integration*, vol. 42, p. 100729, 2024.
- [16] S.-E. Park, S.-J. Lee, H. Bae, K.-H. Kim, E.-J. Kang, J.-S. Kim, Y.-M. Park and M.-J. Park, "Comprehensive issue identification for manufacturing data analytics implementation: Systematic literature review and case studies," *Journal of Manufacturing Systems*, vol. 82, pp. 42-57, 2025.
- [17] C. N. Khan and J. D. Kribs, "Data-Driven Analytics in Smart Manufacturing: A Review of Methods, Applications, and Adoption Constraints," *The International Journal of Advanced Manufacturing Technology*, vol. 144, pp. 3061-3076, 2026.
- [18] H. Tercan and T. Meisen, "Machine learning and deep learning based predictive quality in manufacturing: a systematic review," *Journal of Intelligent Manufacturing*, vol. 33, pp. 1879-1905, 2022.
- [19] G. A. D. Lucca and A. R. Fasolino, "Testing Web-based applications: The state of the art and future trends," *Information and Software Technology*, vol. 48, no. 12, pp. 1172-1186, 2006.