



Yi-Hsuan Lai

Big Data Analytics as an Enabler of Trade Facilitation

Cranfield School of Management
Logistics & Supply Chain Management

MSc in Logistics and Supply Chain Management
Academic Year: 2024 - 2025

Supervisor: Dr. Asefe Forghani
September 2025



Cranfield School of Management
Logistics & Supply Chain Management

MSc in Logistics and Supply Chain Management

Academic Year 2024 - 2025

Yi-Hsuan Lai

Big Data Analytics as an Enabler of Trade Facilitation

Supervisor: Dr. Asefe Forghani
September 2025

This thesis is submitted in partial fulfilment of the requirements for
the degree of MSc Logistics and Supply Chain Management

© Cranfield University 2025. All rights reserved. No part of this
publication may be reproduced without the written permission of the
copyright owner.

Academic integrity declaration

I declare that:

- The thesis submitted has been written by me alone.
- The thesis submitted has not been previously submitted to this university or any other.
- That all content, including primary and/or secondary data, is true to the best of my knowledge.
- That all quotations and references have been duly acknowledged according to the requirements of academic research.

I understand that to knowingly submit work in violation of the above statement will be considered by examiners as academic misconduct.

Abstract

This thesis investigates the role of big data analytics (BDA) in supporting maritime trade facilitation through a Systematic Literature Review (SLR). Against the backdrop of rising global pressures for efficient, sustainable, and resilient supply chains, the review systematically examines how BDA has been applied in port operations and customs procedures. The study maps available tools and applications, evaluates reported benefits, identifies adoption barriers, and assesses the maturity of decision-support frameworks.

A structured SLR process was implemented, analysing 55 peer-reviewed studies published between 2015 and 2025. To ensure synthesis beyond descriptive reporting, findings were clustered into three thematic domains: efficiency (e.g., port optimisation, customs automation, forecasting), sustainability (e.g., emissions reduction, energy efficiency), and resilience (e.g., disruption detection, vulnerability analysis). Evidence confirms strong contributions of BDA to operational efficiency, but integration into sustainability and resilience strategies remains limited. Common barriers include weak data interoperability, regulatory restrictions, and organisational inertia. The review also identifies a shortage of empirically validated decision-support models.

The contribution of this thesis lies in offering a structured and replicable synthesis of the academic evidence base, alongside a typology of BDA tools and applications relevant to maritime trade. By aligning closely with the goals of an SLR, the study not only consolidates fragmented research but also exposes critical gaps and future research priorities. It provides actionable insights for stakeholders such as port authorities, customs agencies, and policy-makers, especially in relation to scaling real-world implementations and fostering cross-sectoral collaboration.

Keywords:

Big Data Analytics, Maritime Trade Facilitation, Port Operations, Customs Automation, Supply Chain Resilience, Sustainability, Decision-Support, Digitalisation, Trade Logistics

Acknowledgements

I am sincerely grateful to my supervisor, Dr. Asefe Forghani, for her guidance, support, and encouragement throughout this research, and for the time and patience she invested at every stage. I also wish to thank my peers for their help during this period—their feedback, discussion, and support made this work possible.

Table of Contents

Academic integrity declaration	i
Abstract.....	ii
Acknowledgements	iii
List of Figures	vi
List of Tables	vii
List of Abbreviations	viii
1 Introduction	1
1.1 Research Background and Context.....	1
1.2 Problem Statement.....	2
1.3 Research Aim and Objectives.....	3
1.4 Research Questions	4
1.5 Justification and Relevance of the Topic	5
1.6 Structure of the Thesis	6
2 Methodology: Systematic Literature Review Chapter title	8
2.1 Introduction & Justification for Using a Systematic Literature Review.....	8
2.2 Search Strategy	9
2.3 Inclusion & Exclusion criteria.....	11
2.4 Screening and PRISMA flow diagram.....	13
2.5 Data Extraction and Synthesis	14
2.6 Limitation of the Review Method	16
3 Descriptive Analysis	18
3.1 Descriptive Findings	18
3.1.1 Chronological Distribution.....	19
3.1.2 Geographical Distribution	20
3.1.3 Research Scale.....	21
3.2 Research Types.....	22
3.3 Conclusion	23
4 Thematic Analysis	24
4.1 Introduction	24
4.2 Thematic Cluster 1: Big Data Analytics for Operational Efficiency	24
4.2.1 Port/Terminal Operations Optimisation	25
4.2.2 Customs Clearance & Inspection Automation	26
4.2.3 Predictive Analytics & Forecasting	27
4.2.4 Synthesis Commentary – Efficiency.....	29
4.3 Thematic Cluster 2: Big Data Analytics for Environmental Sustainability	30
4.3.1 Emissions Monitoring & Reduction.....	30
4.3.2 Fuel consumption optimization	31
4.3.3 Sustainable Supply Chain Practices	33
4.3.4 Synthesis Commentary – Sustainability.....	34

4.4 Thematic Cluster 3: Big Data Analytics for System Resilience	35
4.4.1 Disruption Detection & Response	35
4.4.2 Network Vulnerability & Cascading Failures	36
4.4.3 Synthesis Commentary – Resilience	37
4.5 Cross-Theme Analysis: Barriers and Enablers.....	38
4.6 Synthesis Summary.....	39
5 Discussion	42
5.1 Introduction	42
5.2 Addressing the Research Questions	42
5.2.1 RQ1 – Mapping of BDA Tools, Data Sources, and Applications.....	42
5.2.2 RQ2 – Impact on Efficiency, Sustainability, and Resilience	43
5.2.3 RQ3 – Barriers to BDA Adoption.....	44
5.2.4 RQ4 – Decision-Support Frameworks and Models	44
5.3 Key Gaps, Contradictions, and Emerging Themes	45
5.4 Practical Implications and Academic Relevance	46
6 Conclusion	47
6.1 Summary of Findings.....	47
6.2 Limitations	48
6.3 Recommendations for Practice and Future Work.....	50
References.....	52
Appendices	57

List of Figures

Figure 2-1: PRISMA Flow Diagram..... 14

Figure 3-1: Chronological Distribution 19

Figure 3-2: Geographical Distribution 20

Figure 3-3: Research Scale Distribution 21

Figure 3-4: Research Types Distribution 22

List of Tables

Table 2-1: Inclusion & Exclusion Criteria 11

List of Abbreviations

BDA	Big Data Analytics
SCM	Supply Chain Management
ML	Machine Learning
ARIMA	Autoregressive Integrated Moving Average
MSTGNN	Maritime-knowledge-fused Spatio-Temporal Graph Neural Network
NN	Neural Network
WSN	Wireless Sensor Networks
ETA	Estimated Time of Arrival
LNG	Liquefied Natural Gas
QAP	Quadratic Assignment Problem
TBCC	Two-Hop Biconnected Component
API	Application Programming Interface

1 Introduction

1.1 Research Background and Context

Maritime transport remains the cornerstone of global commerce, accounting for more than 80% of world trade by volume and connecting production regions with consumer markets through intricate logistics networks (Brouer Berit Dangaard and Karsten, 2016). Despite its critical contribution to economic development, the sector continues to grapple with persistent challenges, including inefficiencies in operations, fragmented information flows, and reliance on outdated processes. Paper-based documentation, incompatible digital systems, and inconsistent stakeholder communication frequently result in delays, elevated transaction costs, and reduced responsiveness to disruptions.

The rapid proliferation of big data sources is reshaping this environment. Automatic Identification System (AIS) signals, sensor-equipped vessels and port machinery, meteorological monitoring platforms, and customs databases collectively generate vast data streams, providing unprecedented visibility of maritime activities. Harnessing these resources through Big Data Analytics (BDA)—which encompasses predictive modelling, prescriptive optimisation, and machine learning—offers a powerful avenue for creating digitally integrated, adaptive, and resilient supply chains.

At the same time, the notion of trade facilitation has expanded considerably. Rather than being limited to accelerating customs or port clearance procedures, modern frameworks adopt a broader perspective. As Wilson et al. (2005) note, trade facilitation now incorporates both “border” measures and “behind-the-border” elements, spanning institutional reforms, regulatory harmonisation, and the service infrastructure that supports interoperability and data sharing. Within this broader lens, digitalisation and analytics are positioned as central drivers of efficiency, transparency, and resilience in international trade.

Global initiatives further highlight this strategic transformation. Programmes such as China’s Belt and Road Initiative, the ASEAN Single Window, and the European Union’s Digital Single Market place digital connectivity and data-

sharing mechanisms at the centre of cross-border trade integration. Yet, despite these high-level policy efforts, there remains a significant gap between vision and practice. While the academic literature demonstrates strong engagement with algorithmic design and technical modelling, less attention has been directed toward multi-stakeholder collaboration, standardised data governance, and the validation of decision-support frameworks in operational contexts.

Within this context, undertaking a Systematic Literature Review (SLR) is both appropriate and necessary. The SLR approach enables a transparent, replicable, and methodologically rigorous synthesis of existing evidence. It provides structured insights into how BDA has been applied within maritime trade facilitation, identifies areas of concentrated progress, and exposes persistent shortcomings. By consolidating findings from diverse studies, the review demonstrates the transformative potential of BDA while also clarifying the barriers—technological, organisational, and regulatory—that continue to hinder widespread adoption. This methodological approach ensures that the present study does not simply summarise existing work but critically evaluates it, thereby advancing both academic knowledge and practical understanding of digitalisation in maritime trade.

1.2 Problem Statement

Although the potential of Big Data Analytics (BDA) to transform maritime trade facilitation is widely recognised, its adoption in practice remains fragmented, narrow in scope, and largely exploratory. A significant proportion of existing studies focus on single-use cases such as berth allocation, Estimated Time of Arrival (ETA) prediction, or congestion monitoring. However, few contributions demonstrate how such applications can be scaled, integrated, and embedded into the wider maritime logistics ecosystem. The result is a research and practice landscape dominated by experimental prototypes and isolated models, rather than interoperable and sustainable decision-support frameworks.

This gap becomes increasingly critical in light of evolving global trade dynamics. Geopolitical realignments, climate-induced disruptions, and tightening regulatory requirements are exerting unprecedented pressure on maritime

networks to modernise operational systems as well as governance structures. Despite these demands, persistent barriers constrain progress. Data integration across heterogeneous sources remains limited, particularly when attempting to merge structured datasets such as customs records with unstructured streams including AIS transmissions, meteorological data, or predictive maintenance logs. Institutional inertia, commercial sensitivities, and unresolved legal and regulatory concerns around data ownership and sharing further obstruct the establishment of robust cross-stakeholder governance mechanisms.

At the same time, large-scale initiatives such as China's Belt and Road maritime corridors or the development of Arctic shipping routes have attracted substantial global attention. Yet, research has not systematically examined how BDA could facilitate customs streamlining, enable multimodal transport integration, or monitor environmental performance across these corridors. The absence of empirically validated, governance-aware, and sector-wide frameworks raises serious questions about the maritime industry's readiness to deliver the promised efficiency, transparency, and resilience of digital trade facilitation.

These shortcomings underline the need for a more integrated research agenda that moves beyond fragmented technical contributions and advances scalable, multi-stakeholder solutions. A Systematic Literature Review (SLR) provides the appropriate vehicle for this task, offering methodological rigour, transparency, and replicability. By systematically mapping the existing body of work, an SLR enables identification of structural gaps, clarifies persistent barriers, and highlights pathways through which data-driven decision-support systems can be embedded across maritime supply chains.

1.3 Research Aim and Objectives

The primary aim of this study is to systematically evaluate the role of Big Data Analytics (BDA) in advancing operational efficiency, environmental sustainability, and resilience within maritime trade facilitation. By adopting a Systematic Literature Review (SLR) approach, the research ensures methodological rigour, transparency, and replicability, enabling a comprehensive and structured synthesis of current knowledge.

To achieve this aim, the study pursues four interrelated objectives:

1. Mapping and categorisation – Compile and classify the range of BDA tools, data sources, and applications reported in the literature on maritime trade facilitation, thereby providing a structured overview of the technological landscape.
2. Synthesis of benefits – Identify, synthesise, and critically evaluate the reported benefits of BDA adoption, with a particular focus on improvements in cost efficiency, time savings, emission reductions, and recovery from disruptions.
3. Examination of barriers – Investigate the key barriers hindering BDA adoption, encompassing technical constraints, organisational challenges, and governance-related limitations, while considering variations across regional and stakeholder contexts.
4. Critical review of frameworks – Assess the decision-support frameworks and models proposed in the literature for BDA implementation, with emphasis on their data requirements, scalability, validation status, and practical relevance for both public and private sector stakeholders.

Through these objectives, the study not only consolidates fragmented evidence but also identifies gaps and limitations that must be addressed to enable wider, more integrated, and sustainable application of BDA in maritime trade facilitation.

1.4 Research Questions

Building on the stated objectives, this study formulates the following research questions. They are designed to progress systematically from identification and classification to synthesis, evaluation, and practical insights, in line with the structured approach of a Systematic Literature Review (SLR):

1. Mapping and classification – What types of BDA tools, data sources, and applications have been investigated in relation to maritime trade facilitation, and how are these categorised within the literature?

2. Assessment of benefits – What benefits have been reported from the application of BDA in maritime operations, particularly in relation to efficiency, sustainability, and resilience?
3. Examination of barriers – What barriers to the adoption of BDA are most frequently highlighted, and how do these differ across regional contexts and stakeholder groups?
4. Evaluation of frameworks – Which decision-support frameworks or analytical models have been proposed to guide BDA adoption in maritime trade facilitation, and how do they compare in terms of scope, scalability, validation, and practical relevance?

Collectively, these research questions ensure that the review not only identifies and synthesises existing contributions but also evaluates their applicability and limitations, thereby advancing both academic and practical understanding of BDA in maritime trade.

1.5 Justification and Relevance of the Topic

This research is academically significant because it addresses both methodological and thematic gaps in the existing literature on Big Data Analytics (BDA) within maritime trade facilitation. Previous reviews have often concentrated on narrow technical dimensions, such as algorithmic design or optimisation models, while neglecting the broader ecosystem of adoption that includes organisational processes, governance mechanisms, and cross-sectoral integration. By applying a Systematic Literature Review (SLR), this study provides a transparent and replicable synthesis that spans technical, organisational, and policy perspectives, thereby offering a more holistic understanding of the current state of knowledge and its limitations.

From a practical standpoint, the study generates insights that are directly relevant for industry stakeholders, including port authorities, customs agencies, and supply chain operators. By identifying which BDA applications deliver measurable benefits—and under what contextual conditions—they can make better-informed decisions regarding the design, investment, and scaling of data-

driven solutions. In doing so, the research supports more evidence-based innovation, operational efficiency, and resilience in maritime logistics.

At the policy level, the findings inform ongoing trade facilitation reforms by clarifying the enabling conditions necessary for successful BDA implementation. These conditions include harmonised data standards, supportive legal and regulatory frameworks, and incentive structures that promote cross-sector collaboration. Such insights are particularly valuable for emerging economies and strategic trade corridors that are undergoing rapid digital transformation, including those shaped by initiatives such as the Belt and Road framework.

Taken together, the academic, practical, and policy relevance of this research underscores its contribution not only to scholarly debates but also to real-world efforts aimed at building more efficient, transparent, and resilient maritime trade systems.

1.6 Structure of the Thesis

The remainder of this dissertation is organised into five main chapters, each aligned with the principles of the Systematic Literature Review (SLR) methodology and designed to progressively build from evidence collection to synthesis and evaluation:

- Chapter 2: Methodology – Describes the SLR approach adopted in this study, including the development of search strings, database selection, inclusion and exclusion criteria, quality assessment procedures, and the PRISMA-guided screening process. This chapter establishes the rigour, transparency, and replicability of the review.
- Chapter 3: Thematic Literature Overview – Presents the descriptive findings of the reviewed studies, such as publication trends, geographical distribution, data sources, and the analytics techniques applied. This chapter provides the empirical foundation for subsequent synthesis.
- Chapter 4: Synthesis and Analysis – Integrates and interprets the evidence by clustering the literature into three thematic domains—efficiency, sustainability, and resilience. It evaluates the contributions of

BDA in each area, highlights barriers to adoption, and reviews proposed decision-support frameworks.

- Chapter 5: Discussion – Directly addresses the research questions by situating the findings within wider academic debates and policy contexts. It also discusses the implications for practitioners and decision-makers in maritime trade and logistics.
- Chapter 6: Conclusion – Summarises the key contributions of the research, reflects on its methodological and conceptual limitations, and identifies avenues for future investigation and practical application.

2 Methodology: Systematic Literature Review

2.1 Introduction

Research on Big Data Analytics (BDA) in maritime trade is extensive yet dispersed across disciplinary boundaries. Operational optimisation is typically discussed within engineering and operations management studies, whereas digital infrastructure and governance issues are more often addressed in information systems or policy-focused literature. This disciplinary separation fragments the knowledge base, limiting a holistic understanding of how BDA contributes to trade facilitation as an integrated, multi-stakeholder process.

To overcome this fragmentation, the present study adopts a Systematic Literature Review (SLR) as its methodological foundation. Unlike narrative reviews, which often rely on subjective judgement and selective citation, an SLR applies a transparent, replicable, and scientifically rigorous process to collect, appraise, and synthesise evidence. This structured approach reduces bias, enhances comparability across studies, and provides a reliable basis for drawing generalisable conclusions.

The methodological distinctiveness of SLRs lies in their reliance on pre-specified protocols that make every decision explicit and auditable. These protocols include the formulation of clear inclusion and exclusion criteria, the construction of reproducible search strings, and the application of standardised quality appraisal tools. By documenting each stage of the review process, the study ensures full traceability and replicability of findings (Tranfield et al., 2003; Gough et al., 2012; Lame, 2019).

Consistent with established guidance, this study implements an eight-step SLR framework (Lame, 2019; Okoli, 2015): (1) defining research questions, (2) developing inclusion and exclusion criteria, (3) conducting comprehensive searches, (4) screening retrieved records, (5) appraising study quality, (6) extracting data using structured forms, (7) conducting thematic synthesis, and (8) interpreting the findings in relation to the research aims. The PRISMA flow

diagram is employed to document the screening and selection process, further strengthening transparency.

This methodological choice enables systematic mapping of BDA applications in maritime trade, the synthesis of reported benefits such as time and cost savings, emission reductions, and disruption management, as well as the critical evaluation of adoption barriers spanning institutional, technical, and governance dimensions. Moreover, by organising the synthesis into the three core themes of efficiency, sustainability, and resilience, the study moves beyond fragmented, case-specific insights towards conclusions with broader academic and practical relevance.

2.2 Search Strategy

To guarantee transparency and reproducibility, a structured search protocol was designed in advance of conducting any database queries. The protocol defined the search string, database scope, time horizon, language filters, and document type restrictions before the review commenced. Establishing these parameters upfront allowed for a consistent process in which any deviations could be documented and scrutinised, thereby ensuring methodological rigour and replicability (Okoli, 2015).

The review employed three academic databases—Scopus, Web of Science, and EBSCO—to secure broad disciplinary coverage. Scopus and Web of Science were selected for their extensive indexing of peer-reviewed research across logistics, supply chain management, and operations domains, while EBSCO was included to capture trade-oriented and SCM-focused studies that might otherwise be underrepresented. This combination reflects prior methodological recommendations that combining general citation indexes with domain-specific repositories improves retrieval accuracy and minimises bias.

A Boolean search string was developed to capture the intersection of four conceptual domains: analytics, maritime logistics, trade facilitation, and performance outcomes. The string was applied consistently across platforms,

with minor adjustments to reflect each database's syntax (e.g., TITLE-ABS-KEY in Scopus, TS= in Web of Science).

Boolean String
("Big Data Analytics" OR "data analytics" OR "machine learning" OR "predictive analytics") AND (maritime* OR "port logistics" OR shipping* OR "sea transport") AND ("trade facilitation" OR customs OR "international trade") AND (efficiency OR resilience OR sustainability OR "decision making")

Multi-word terms were placed in quotation marks to ensure precise retrieval, while wildcards (e.g., maritime*, shipping*) expanded recall to include lexical variations. All searches were executed on the same calendar date, and the full string was archived with the review protocol to enable exact replication.

The review was restricted to peer-reviewed journal articles, published in English between 2015 and 2025, and situated in the context of supply chain management or logistics. Conference proceedings, technical reports, books, and studies outside the defined scope were excluded to maintain alignment with the research objectives and to ensure a consistent threshold of methodological quality.

All retrieved records were exported in RIS format and imported into Zotero. Automated deduplication, based on DOI, title, first author, and publication year, was followed by a manual review to detect near-duplicates arising from minor title variations or author discrepancies. For records retrieved from multiple databases, bibliographic details such as abstracts, keywords, and citation fields were consolidated into a single complete entry. Each stage of removal and refinement was logged, ensuring traceability in line with systematic review reporting standards (Gough et al., 2012; Okoli, 2015).

By combining diverse databases, a carefully structured Boolean string, and robust documentation practices, the search strategy ensures both comprehensive coverage and methodological rigour. The process reduced bias,

prioritised relevance, and produced a fully auditable record that can be independently replicated in future studies.

2.3 Inclusion & Exclusion criteria

To ensure methodological consistency and transparency, predefined inclusion and exclusion criteria were applied during the initial screening stage and maintained throughout the review process. These criteria were established in line with the principles of systematic reviewing, providing a structured and replicable framework for identifying relevant studies. They reflect three priorities: (i) direct relevance to the research questions on BDA and maritime trade facilitation; (ii) methodological robustness and transparency; and (iii) practical accessibility to enable full-text assessment.

Table 2-1. Inclusion & Exclusion Criteria

Criteria	Inclusion	Exclusion	Rationale
Language	English	Non-English	To ensure consistency and avoid misinterpretation due to translation limitations. Also, the leading academic papers and journals are published primarily in English.
Publication date	From 2015 to 2025	Pre-2015	To capture evidence from the past decade, reflecting current developments and preserving analytical validity.
Publication type	Peer-reviewed journal articles in SCM or	Conference papers, books, technical reports, and editorials	To ensure methodological consistency, transparency, quality, and reliability.

	logistics-related		
Domain relevance	Linking BDA to maritime trade facilitation	Outside the maritime logistics or trade facilitation context	To keep the focus on the specific scope and synthesis aligned with the research question.
Access	Full-text available	Full-text unavailable	To ensure each study can be critically reviewed and allows for a comprehensive, detailed analysis within the review.

As shown in Table 2-1, restricting the review to English-language publications ensures consistency and avoids potential misinterpretation arising from translation, while also acknowledging that most high-impact journals publish in English. The chosen time horizon (2015–2025) reflects the period in which operational BDA applications began to emerge more widely in maritime and trade contexts, thereby ensuring that the dataset remains up to date and analytically relevant.

Limiting the dataset to peer-reviewed journal articles helps safeguard methodological quality and comparability across studies. Grey literature (e.g., conference papers, reports, editorials) was excluded because of inconsistent review standards and variable methodological transparency. The domain relevance criterion was applied to focus only on studies explicitly linking BDA to maritime logistics or trade facilitation; studies centred on manufacturing, terrestrial transport, or general data science were excluded unless a direct connection to maritime trade facilitation was evident. Finally, requiring full-text availability enabled comprehensive appraisal and synthesis during subsequent stages of the review.

Each screening decision, along with its justification, was logged in accordance with PRISMA guidelines to ensure traceability, procedural rigour, and transparency.

2.4 Screening and PRISMA flow diagram

The study employed the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) framework to document and visualise the multi-stage screening and selection process. The PRISMA flow diagram provides both a reporting tool and a methodological safeguard, ensuring that each decision is traceable, reproducible, and systematically justified. As emphasised by Lane (2019), PRISMA helps to mitigate selection bias by recording the number of studies identified, screened, excluded, and retained, together with the rationale for exclusion at each stage.

The workflow followed three sequential phases:

1. Deduplication – removal of duplicate records across databases.
2. Title and abstract screening – exclusion of irrelevant studies based on predefined inclusion and exclusion criteria.
3. Full-text eligibility assessment – in-depth evaluation to confirm compliance with all selection requirements.

At the initial search stage, a total of 3,058 records were retrieved: 1,189 from Scopus, 1,860 from EBSCO, and 9 from Web of Science. After applying inclusion and exclusion filters (removing conference papers, books, technical reports, and news sources), the dataset was reduced to 900 records from Scopus, 648 from EBSCO, and 7 from Web of Science.

All records were exported in RIS format and imported into Zotero for reference management. Automatic duplicate detection, based on DOI and title matching, removed 75 duplicates, leaving 1,480 records for initial screening.

The screening process unfolded in two stages. First, title-and-abstract inspection excluded 1,402 records, advancing 78 articles to the full-text review stage. Second, during full-text assessment, a further 23 studies were excluded

due to failing to meet one or more inclusion criteria. Ultimately, 55 peer-reviewed journal articles were retained for final synthesis.

The entire process is summarised in Figure 2-1 (PRISMA Flow Diagram), which provides a transparent record of each decision point. This systematic approach ensured consistency, reduced bias, and preserved full traceability of the dataset, in line with best practice guidelines for systematic reviews (Moher et al., 2009; Tranfield et al., 2003).

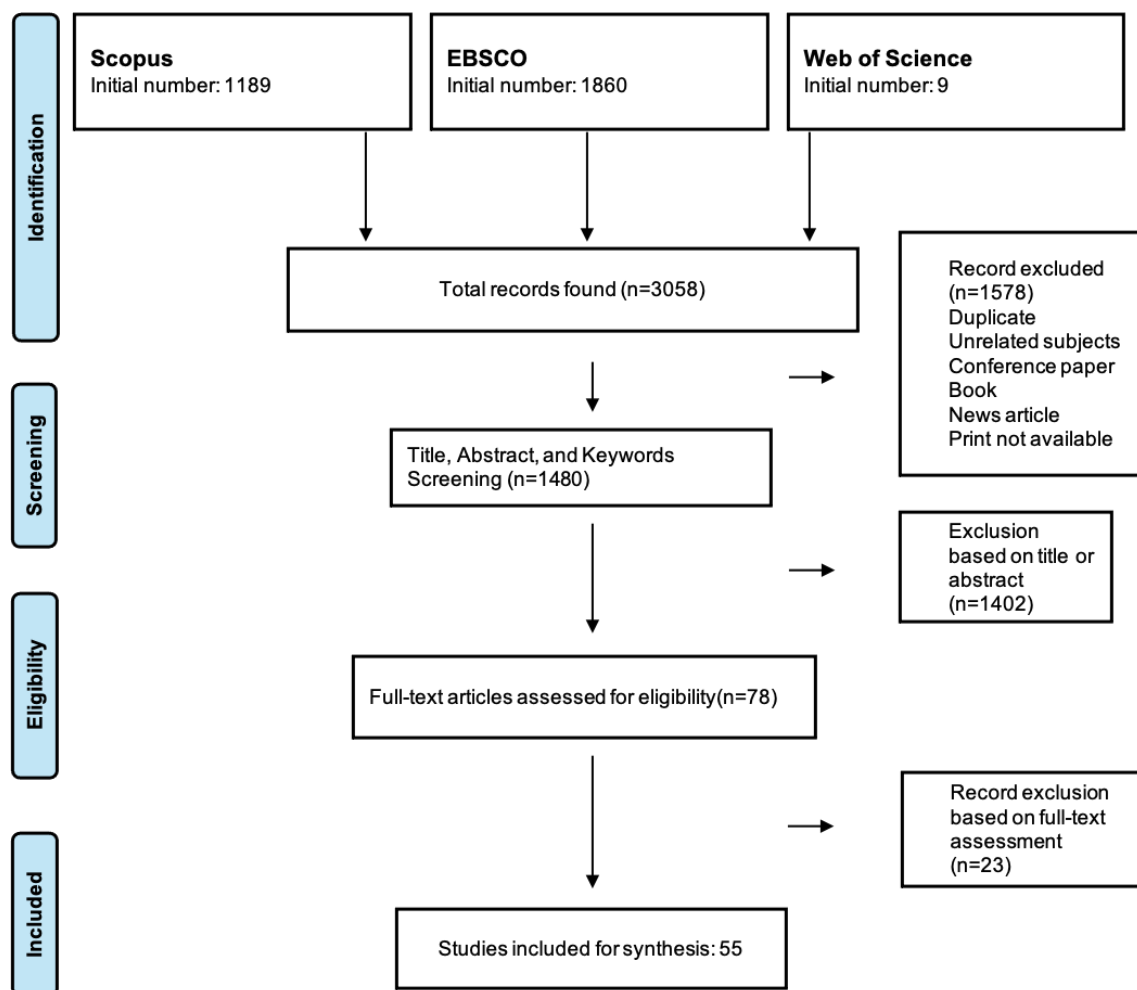


Figure 2-1: PRISMA Flow Diagram

2.5 Data Extraction and Synthesis

After the screening stage, a structured process was employed to extract, organise, and synthesise data from the finalised set of 55 studies. The purpose

of this phase was twofold: first, to capture information relevant to the research questions and objectives, and second, to systematically convert this evidence into analytical insights through descriptive and thematic synthesis.

To promote consistency and transparency, a customised data extraction matrix was designed for this study. Each study was mapped against key variables, including bibliographic details, the type and nature of big data utilised, the analytics techniques applied, the specific stages of trade facilitation addressed, and the outcome category (efficiency, sustainability, or resilience). Additional fields captured contextual elements such as geographic focus, modelling frameworks, indicators of scalability, and reported barriers or enablers to implementation. The design of this matrix reflects the recommendation by Munn et al. (2014) that systematic reviews should move beyond narrative description by using structured and reproducible data extraction methods capable of capturing both descriptive attributes and outcome-based evidence.

The synthesis process followed a two-stage approach. The first stage employed descriptive statistics to provide an overview of the literature, classifying studies by publication year, geographical distribution, data sources, methodological approaches, and types of analytics applied. This mapping established the empirical foundation for Chapter 3, where publication trends, geographic scope, and methodological concentrations are presented systematically.

The second stage involved thematic synthesis, focusing on the interpretation of findings in relation to the three core goals of this research: efficiency, sustainability, and resilience. Reported benefits were categorised accordingly, linking types of BDA applications with specific facilitation outcomes such as reduced clearance times, lower emissions, or improved recovery from disruptions.

By combining descriptive mapping with thematic synthesis, the review adopts an integration strategy suited to both the nature of the literature and the objectives of the SLR. Chapter 3 provides a structured descriptive overview of the field, while Chapter 4 synthesises conceptual insights and thematic relationships. Together, these steps establish a robust basis for analytical

conclusions and support the development of a coherent framework to address the research questions.

2.6 Limitation of the Review Method

While the systematic approach adopted in this review ensures transparency, rigour, and analytical consistency, several methodological limitations must be acknowledged.

First, the review was limited to peer-reviewed journal articles published in English between 2015 and 2025. This restriction strengthened quality control and comparability across studies but excluded potentially relevant contributions from non-English publications, industry reports, and grey literature. As a result, the review reflects the academic discourse more than the full spectrum of practical implementation experiences in port operations and customs environments.

Second, the search strategy was confined to three databases—Scopus, Web of Science, and EBSCO—using a predefined Boolean string aligned to the research objectives. While this combination provided broad coverage across logistics, supply chain management, and operations, it is possible that relevant studies using alternative terminologies or situated in adjacent fields were not captured.

Third, the heterogeneity of the included studies in terms of methodologies, application contexts, and reporting standards meant that a formal meta-analysis could not be conducted. Instead, synthesis relied on descriptive statistics and thematic analysis. These approaches are well suited for capturing diversity and mapping trends, but they do not generate statistically generalisable findings or quantify effect sizes.

Fourth, some degree of subjectivity in coding decisions was unavoidable, particularly when classifying reported benefits into the categories of efficiency, sustainability, and resilience. A customised extraction matrix and consistency checks were employed to mitigate bias, yet interpretive judgement cannot be fully eliminated.

Finally, because digitalisation and data governance in maritime trade are rapidly evolving, certain findings may become time-bound or incomplete in capturing the most recent technological advancements. This inherent limitation underscores the need for continuous updating of systematic reviews in dynamic fields such as BDA.

These limitations do not undermine the validity of the review but should be considered when interpreting the findings. They also highlight opportunities for future research, which will be further discussed in Chapter 5, particularly in relation to expanding sources, refining methodologies, and incorporating longitudinal perspectives.

3 Descriptive Analysis

This chapter presents a structured descriptive analysis of the evidence base derived from the 55 studies selected for full inclusion in the review. As outlined in Chapter 2, the dataset was established through the PRISMA procedure, where 23 articles were excluded at the full-text stage, leaving 55 publications that form the denominator for all descriptive statistics reported here. These studies also constitute the empirical foundation for the subsequent thematic synthesis.

The analysis adheres to established SLR reporting conventions, with studies classified by publication year, geographic context, research scale, and methodological orientation. The aim is to provide a transparent descriptive profile that clarifies the scope and nature of existing research, before moving toward interpretive synthesis in Chapter 4.

3.1 Descriptive Findings

This section provides the descriptive characteristics of the reviewed studies, including their chronological distribution, geographic origin, empirical scale, and research type. It is important to note that concentrations by year, country, or method should be interpreted cautiously, as they are conditioned by the review protocol (database coverage, language, document type, and time window). Nonetheless, this structuring supports methodological traceability and highlights where academic attention has been most visible.

3.1.1 Chronological Distribution

The temporal profile of the selected studies is illustrated in Figure 3-1. Publications prior to 2018 appear only sporadically, with a clear increase observed through the early 2020s. This pattern reflects both the review protocol, which restricted the dataset to studies published between 2015 and 2025, and the broader rise of data infrastructures and analytics tools in maritime logistics during this period.

The distribution should be read as descriptive rather than as evidence of field maturity. Peaks in certain years often coincide with special journal issues, large-scale data releases, or global disruptions that catalyse applied research. The apparent decline in the final year of the window is an artefact of the incomplete publication cycle, representing an undercount rather than a downward trend.

Although fewer in number, the earlier studies retain value in providing the conceptual and methodological groundwork against which claims of efficiency gains, emissions reductions, or resilience improvements can be evaluated.

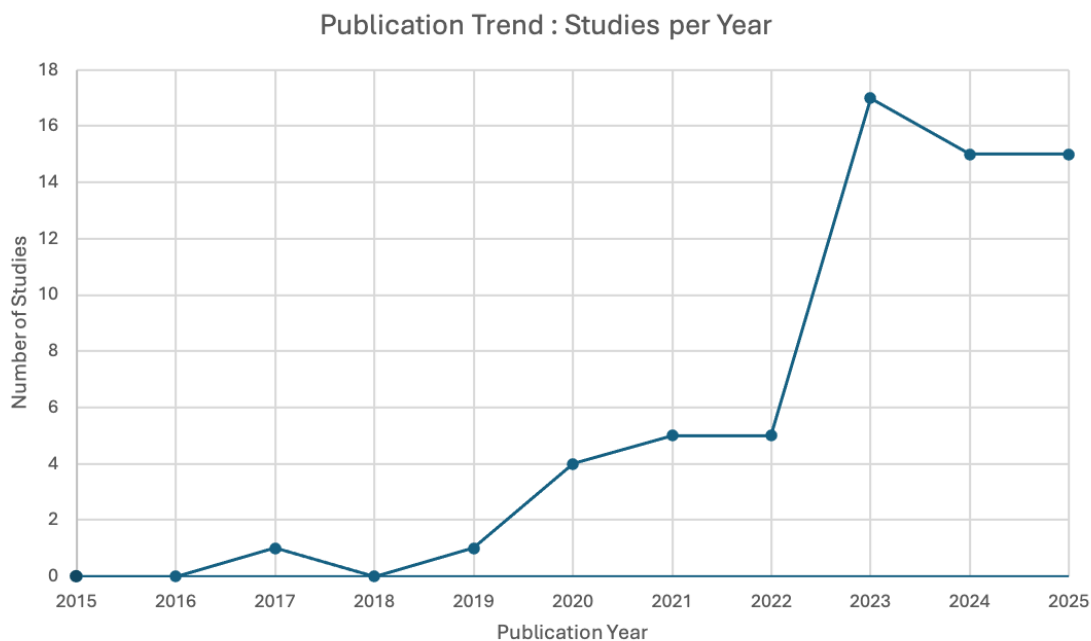


Figure 3-1: Chronological Distribution

3.1.2 Geographical Distribution

The geographic profile of the included studies is shown in Figure 3-2. Author affiliations are used as a proxy for research origin, which indicates where academic outputs are produced rather than where data are collected or solutions implemented.

The corpus shows a concentration in East Asia, particularly China, Hong Kong, and South Korea, alongside contributions from Europe (notably France) and cross-border collaborations, which are grouped under “Multi-institution”. A residual “Other” category captures smaller contributions from diverse regions.

This distribution highlights the importance of data-rich maritime hubs and well-developed academia–industry ecosystems. However, it also reflects the review’s English-language restriction, which may underrepresent scholarship from non-English journals and Global South contexts. Despite this, the literature demonstrates considerable breadth, with many studies explicitly targeting cross-regional or generalisable trade facilitation challenges rather than single-country cases.

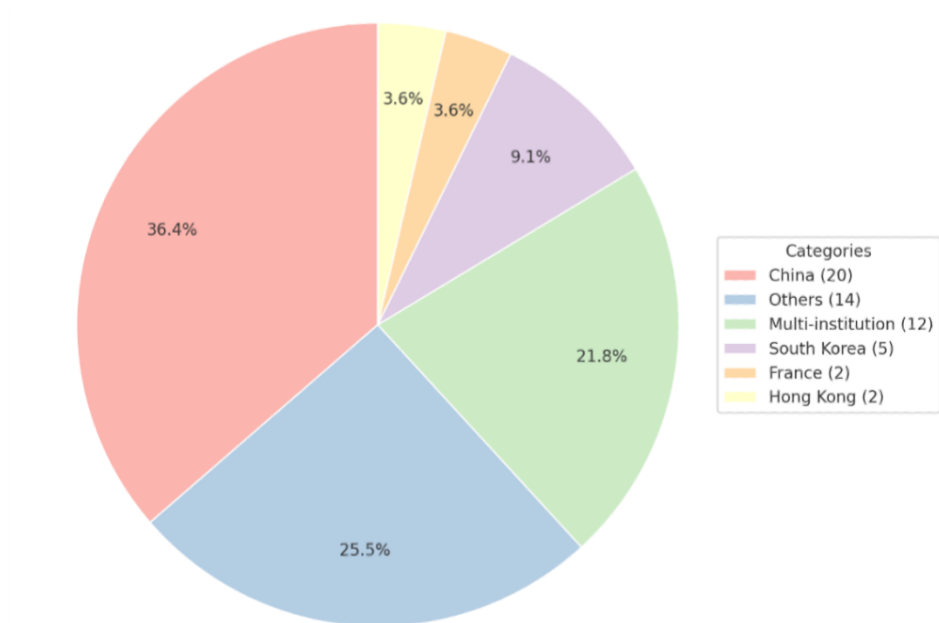


Figure 3-2: Geographical Distribution

3.1.3 Research Scale

Moving beyond author affiliation, Figure 3-3 presents the empirical scale of the studies reviewed. The classification distinguishes between:

1. Regional – covering interconnected locations such as maritime corridors, sea basins, or port clusters.
2. National – confined to a single country's maritime supply chain, port systems, or customs processes.
3. Global – spanning multi-continental trade networks and shipping routes.

Regional studies typically explore dynamics such as congestion propagation or chokepoint exposure, offering granular visibility into inter-port interactions. National-level work frequently examines policy instruments (e.g., single windows, customs risk engines, port community systems) or infrastructure configurations (e.g., gate operations, hinterland connectivity), making them suitable for analysing “inside-the-border” reforms. Global studies adopt a systemic perspective, assessing the strategic role of transport hubs, the cascading effects of delays, and the resilience of shipping networks under disruption.

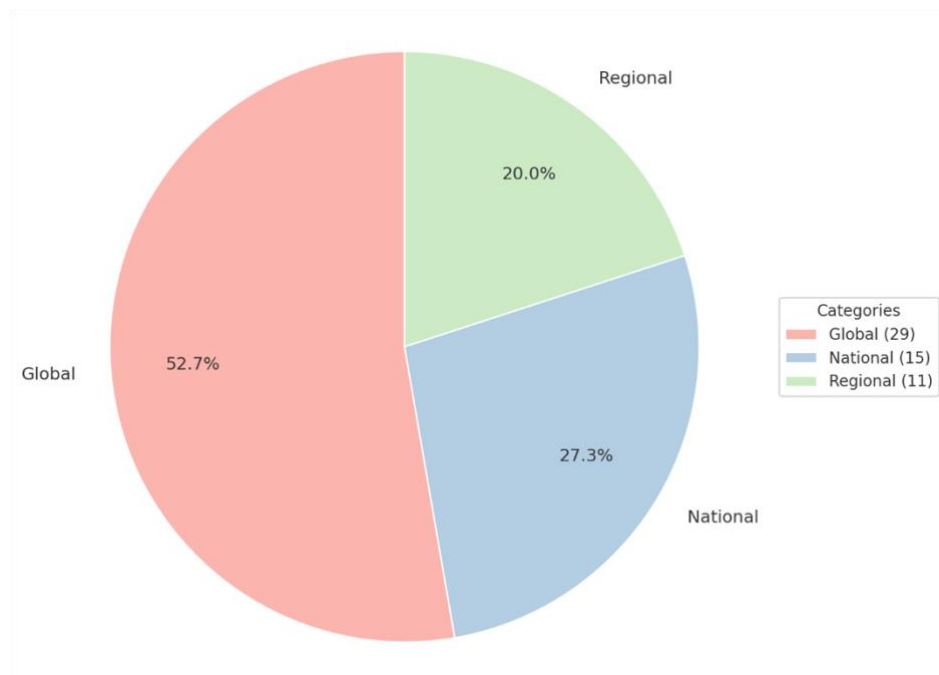


Figure 3-3: Research Scale Distribution

3.2 Research Types

The methodological orientations of the 55 studies are classified using the three-way taxonomy proposed by Heyvaert et al. (2013), distinguishing between quantitative, qualitative, and mixed-methods approaches. As shown in Figure 3-4, the corpus is overwhelmingly quantitative, with 50 studies (90.9%) relying on model-based or data-driven techniques, 1 study (1.8%) adopting a qualitative approach, and 4 studies (7.3%) employing mixed methods.

The predominance of quantitative research reflects the nature of BDA itself, which draws heavily on large-scale datasets such as AIS signals, port transaction logs, and external feeds. However, two caveats must be acknowledged. First, quantitative dominance does not necessarily equate to integration into decision-making; many contributions remain at the prototype or simulation stage. Second, the very limited use of qualitative and mixed approaches suggests that organisational adoption, governance arrangements, and infrastructural contexts are comparatively underexplored in relation to algorithmic performance.

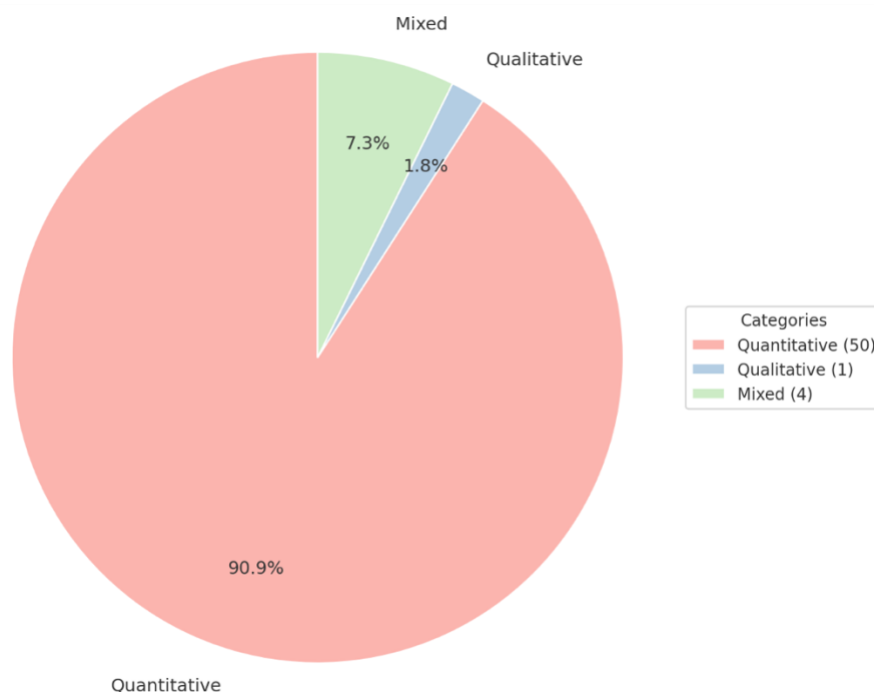


Figure 3-4: Research Types Distribution

3.3 Conclusion

This chapter has provided a descriptive mapping of the 55 studies included in the review, covering publication year, geographic distribution, research scale, and methodological orientation.

- Chronology – Most publications cluster between 2021 and 2023, reflecting both the recent surge in applied research and the predefined search window. This indicates growing academic attention rather than field maturity.
- Geography – Research origin is concentrated in East Asia, particularly China, Hong Kong, and South Korea, supplemented by European and multi-institutional contributions. These reflect where research is produced, not always where solutions are implemented.
- Scale – The studies span regional, national, and global levels, offering insights into corridor-specific, country-specific, and systemic maritime challenges.
- Methods – The literature is dominated by quantitative approaches (90.9%), with limited qualitative and mixed-method contributions. This imbalance suggests that organisational and governance dimensions remain comparatively underexplored.

Overall, the descriptive analysis provides a transparent profile of the current research landscape. It establishes the empirical base for the thematic synthesis in Chapter 4, where the contributions of BDA to efficiency, sustainability, and resilience are examined in depth.

4 Thematic Analysis

4.1 Introduction

This chapter presents the thematic synthesis of peer-reviewed studies on the role of Big Data Analytics (BDA) in enabling maritime trade facilitation. The analysis directly addresses the research questions, structured across four dimensions:

1. Mapping BDA tools, data sources, and application domains.
2. Assessing reported contributions to efficiency, sustainability, and resilience.
3. Identifying the barriers and enablers that shape adoption.
4. Reviewing decision-support frameworks in terms of their scope, scalability, and practical relevance.

The selected articles are organised into thematic clusters—efficiency, sustainability, and resilience—using a coding framework refined during full-text review. Within each cluster, subsections are defined by the dominant empirical applications identified in the literature. Cross-cutting issues such as barriers and enablers are synthesised separately, followed by an appraisal of decision-support frameworks. Findings are reported with explicit recognition of their scope and limitations, while broader theoretical and policy implications are developed in Chapter 5.

4.2 Thematic Cluster 1: Big Data Analytics for Operational Efficiency

The first thematic cluster focuses on how BDA is applied to improve operational efficiency in maritime trade. The selected studies highlight three main domains of application: port and terminal optimisation, customs automation, and predictive analytics. This section begins with an examination of port and terminal operations, where optimisation and machine learning (ML) approaches are most prevalent.

4.2.1 Port/Terminal Operations Optimisation

The literature demonstrates increasing use of BDA to optimise resource allocation, scheduling, and throughput in ports and dry ports, employing methods ranging from optimisation modelling to simulation and machine learning.

Dynamic yard block allocation illustrates how integrating analytical models with ML can deliver measurable efficiency gains. For instance, real-time policies that combine queuing and renewal theory with ML-based quay crane productivity forecasts reduce vessel berth times compared with static allocation, thereby minimising idle time associated with truck travel and yard-crane queues (Abouelrous et al., 2025).

At a broader network level, Industry 5.0 principles and Internet of Things (IoT) sensor data underpin multi-stage optimisation frameworks. One example combines multi-criteria port selection, seasonal ARIMA forecasting, and a bi-objective mixed-integer programming model solved with fuzzy meta-goal programming. Reported outcomes include a 68% reduction in tardiness and a 71% increase in transparency, though these results are contingent on high IoT investment and are demonstrated in a single-country setting (Rekabi & Sazvar, 2025).

AI/ML methods are also used to enhance short- and medium-term scheduling, particularly by improving traffic and workload forecasts for berth, crane, gate, and rail operations. Neural network architectures feature prominently. A maritime-knowledge-fused spatio-temporal graph neural network (MSTGNN), for example, successfully models complex inter-port interactions, outperforming benchmark models and enabling proactive congestion management across dataset (Mei et al., 2025). Similarly, a learning-based robust optimisation scheme for synchromodal freight integrates Bayesian neural networks to quantify uncertainty and construct adjustable robust sets. Applied to short-sea shipping and hinterland integration, this approach generates feasible scheduling policies under variable travel times, thereby enhancing terminal operations

under uncertainty and aligning with the goals of trade facilitation (Filom & Razavi, 2025).

Alongside optimisation and ML scheduling, IoT-enabled cold chain monitoring represents another strand of operational efficiency improvement. Wireless sensor networks and RFID, integrated with cloud-based systems, provide continuous container monitoring, issuing proactive alerts when conditions deviate from thresholds. This enhances responsiveness, reduces spoilage risk, and supports resource allocation in temperature-sensitive logistics (Cil et al., 2022).

Taken together, these studies demonstrate the potential of combining IoT sensing, predictive ML, and prescriptive optimisation to improve berth turnaround times, equipment utilisation, and inland coordination. Nonetheless, most applications remain context-specific, with scalability limited by data integration challenges, organisational readiness, and the costs of IoT deployment. Interpretability of ML models also remains a barrier to their acceptance in operational decision-making.

4.2.2 Customs Clearance & Inspection Automation

A second domain of efficiency gains concerns the automation of customs clearance and inspection, where Big Data Analytics (BDA) is increasingly applied to enhance risk targeting, accelerate verification processes, and support AI-driven screening.

Recent developments show a shift away from heuristic inspection rules toward data-driven risk profiling. In port state control, domain-informed machine learning models such as monotonic XGBoost have been shown to improve both fairness and detection accuracy. When combined with officer scheduling, deficiency detection rates increase by more than 20% relative to current practice, narrowing the gap to benchmark performance standards (Yan et al., 2021). These approaches demonstrate how prediction can be integrated with operational decision-making, although the evidence remains geographically narrow and is constrained by imbalanced inspection datasets.

Building on this, some studies adopt decision-focused learning to improve vessel selection by directly optimising inspection outcomes. At the Port of Hong Kong, such models outperform traditional risk scores and baseline methods, delivering stronger inspection payoffs (Tian et al., 2023). While this demonstrates methodological innovation, the approach remains limited by data quality and technical scalability, with findings drawn from a single-port dataset.

Automation of identity and document verification is another area of progress. Trials combining electronic vehicle identification, RFID, multimodal biometrics, and genetic algorithms for real-time inspection dispatch report up to 30% shorter waiting times and 25% higher utilisation of customs resources. Yet, effectiveness depends heavily on biometric accuracy in variable environments and on the ability to maintain performance under complex, high-volume customs conditions (Zhou et al., 2024).

Adoption outcomes are also shaped by institutional and legal factors. Cross-country comparisons reveal that customs automation performance correlates with governance quality and institutional maturity. Stronger regimes achieve measurable efficiency gains, whereas weaker systems face procedural and legal bottlenecks that reduce the benefits of technical innovations (Ben Haj Ahmed et al., 2023).

Overall, classification models and decision-focused frameworks improve inspection targeting, while automation technologies reduce bottlenecks in clearance processes. However, most applications remain at pilot or experimental scale, constrained by interoperability challenges, data quality issues, and legal requirements for due process. Hybrid approaches—combining interpretable risk factors with machine learning outputs—appear the most practical pathway for balancing efficiency with transparency in customs decision-making.

4.2.3 Predictive Analytics & Forecasting

Within the broader framework of intelligent transport systems, predictive analytics play a central role in supporting maritime trade facilitation. By enabling

forecasts of estimated time of arrival (ETA), congestion levels, throughput volumes, and condition-related risks, predictive models enhance coordination across ports, terminals, and hinterland connections, thereby improving operational efficiency.

Recent advances in deep learning on maritime networks have strengthened traffic-flow forecasting. For example, the application of a maritime-knowledge-fused spatio-temporal graph neural network (MSTGNN) has been shown to outperform baseline models by capturing multi-step vessel traffic flows. These improved forecasts enhance berth and yard planning, enabling ports to manage congestion more proactively. However, their generalisability beyond the datasets on which they are trained remains uncertain (Mei et al., 2025).

Another line of research focuses on short-term throughput forecasting. At China's two largest ports, an integrated model combining ARIMA with a zonotopic Kalman filter reduces lag and noise in monthly throughput predictions. This supports improved labour scheduling and equipment deployment but relies heavily on the stability of port dynamics and the availability of reliable historical records (Zhang & Cao, 2025).

Forecasting has also been extended to the market domain, linking freight rates and trade indices with port operations. One study uses an ensemble of a convolutional neural network, a bi-directional long short-term memory (Bi-LSTM) network, and an attention mechanism, with commodity-futures indicators as inputs. The model predicts the Shanghai Containerised Freight Index with high accuracy, while interpretability is enhanced through Shapley Additive Explanations (SHAP). These results suggest that financial signals can inform earlier adjustments to terminal and network operations than operational data alone. However, dependence on volatile financial markets can decouple forecasts from actual physical flows during crises (Su et al., 2025).

Although less common, predictive analytics have also been applied to equipment and fuel-risk forecasting. Physics-informed neural networks, coupled with mixed-integer optimisation, improve fuel-consumption modelling, supporting ETA control while retaining interpretability. Nonetheless, these

models require vessel-specific calibration, limiting their portability across fleets (Wang et al., 2023).

In summary, predictive analytics offer considerable potential to enhance foresight in berth allocation, yard scheduling, throughput management, and multimodal coordination. Yet, challenges remain regarding dataset coverage, model transferability, interpretability, and institutional readiness. These limitations highlight the importance of hybrid approaches that combine operational data, market indicators, and domain knowledge to balance predictive power with robustness and scalability.

4.2.4 Synthesis Commentary – Efficiency

The reviewed studies collectively demonstrate that Big Data Analytics (BDA) can deliver substantial improvements in operational efficiency across maritime trade facilitation, particularly in the domains of port optimisation, customs automation, and predictive forecasting. Integrated frameworks that combine IoT-enabled sensing, machine learning, and optimisation models consistently report measurable gains in vessel turnaround times, equipment utilisation, and inland coordination. Within this body of work, robust synchromodal planning and spatio-temporal graph neural networks (MSTGNN) are especially notable for their capacity to model complex interdependencies and anticipate congestion.

Hybrid modelling approaches also show promise. For example, combining queuing theory with ML, or ARIMA with Kalman filters, improves resilience to noise and variability, while ensemble deep learning extends forecasting horizons through integration of market signals. Meanwhile, decision-focused learning and physics-informed neural networks help address concerns over interpretability by embedding domain knowledge into predictive pipelines.

Despite these advances, generalisability remains limited. Most implementations are restricted to single ports or national systems, leaving multi-port coordination, cross-border customs automation, and equipment-risk prediction relatively underexplored. Furthermore, the integration of predictive outputs into operational decision rules is often conceptual rather than demonstrated in live

environments. The predominance of black-box ML methods also raises concerns over transparency, with statistical models vulnerable to structural breaks in dynamic trade environments.

Looking forward, progress in this domain requires broader validation across diverse contexts, stronger coupling with operational governance frameworks, and the design of explainable, audit-ready ML systems. Future research should also pursue the development of modular, policy-aware analytics architectures that can support interoperability and scale across ports and customs authorities. Such advances are critical to moving beyond isolated pilot projects toward the systemic deployment of efficiency-enhancing BDA solutions.

4.3 Thematic Cluster 2: Big Data Analytics for Environmental Sustainability

Focusing on the environmental dimension of maritime digitalisation, this cluster highlights how BDA supports emissions monitoring and reduction, fuel consumption optimisation, and the advancement of sustainable supply chain practices.

4.3.1 Emissions Monitoring & Reduction

Within the sustainability dimension of maritime trade facilitation, Big Data Analytics (BDA) is increasingly applied to track and reduce emissions by fusing diverse operational and external data streams. These applications not only strengthen environmental monitoring but also inform regulatory compliance and operational decision-making.

At the network level, analytics enable emissions surveillance even when direct fuel consumption data are unavailable. For example, a structural analysis employed AIS-derived origin–destination links combined with graph deep learning to identify critical nodes, map traffic flows, and expose resilience gaps. This demonstrates how network-scale observables can support emissions monitoring and highlight vulnerabilities in maritime corridors (Mei et al., 2024).

At the ship and fleet level, monitoring increasingly relies on integrating heterogeneous data sources such as noon reports, AIS tracks, and weather forecasts to build auditable models of fuel consumption and emissions. An innovative approach combines physics-informed neural networks with mixed-integer optimisation, balancing predictive accuracy with interpretability. This design enhances transparency in forecasting fuel use and emissions, enabling more accountable and reliable monitoring systems (Wang et al., 2023).

Complementary methods include hybrid prediction–optimisation frameworks. For instance, a two-stage model that integrates random forest forecasting with speed optimisation improves fuel consumption prediction for dry bulk ships. By merging noon report and weather data, this approach achieves 2–7% reductions in both fuel use and CO₂ emissions, demonstrating the operational value of machine learning in maritime energy efficiency (Yan et al., 2020).

Beyond monitoring, decision-support frameworks directly link analytics to reduction strategies. One study integrates a carbon tax into an optimisation framework, using Bayesian neural networks to capture uncertainty in travel times. Scenario-based modelling demonstrates how carbon taxation encourages modal shifts toward greener transport options, offering evidence for both regulatory design and freight planning (Filom & Razavi, 2025).

Despite these advances, several limitations persist. AIS and log data remain prone to gaps and inconsistent standards; black-box components within ML models undermine auditability; and many applications are confined to specific fleets or regional contexts, restricting generalisability. Wider deployment will require standardised data pipelines, interpretable algorithms, and cross-context validation to ensure that emissions-focused analytics are not only technically robust but also operationally and regulatorily relevant.

4.3.2 Fuel consumption optimisation

Emissions monitoring and reduction provide the foundation of measurement and compliance, whereas fuel consumption optimisation serves as its operational extension, embedding predictive models into routing and speed

management. Complementing emissions-focused forecasting, Yan et al. (2020) extend predictive modelling into operational optimisation. Their two-stage framework translates fuel rate predictions into speed adjustments, aligning voyage planning with energy efficiency goals. By embedding ETA constraints and environmental inputs, the model supports actionable strategies for reducing fuel use, reinforcing machine learning's role in sustainable maritime operations. The study by Wang et al. (2023), their framework supports fuel-efficient routing by translating model outputs into actionable speed adjustments. The integration of domain-informed constraints ensures transparency in decision-making, enabling practical deployment of AI-driven strategies for reducing fuel consumption without compromising navigational performance.

As noted earlier, the leverage of fused data streams provides the foundation that is here extended into prescriptive fuel optimisation. Yan et al. (2024) combine domain-informed neural networks with bi-objective and market-based optimisation models to translate these multi-source data into strategic speed planning, yielding measurable cost savings and up to 15% reductions in emission charges, thereby supporting sustainable voyage optimisation. And analyses based on large-scale AIS pairs with ocean environment data further reveal speed baselines under different operating conditions and highlight inefficient practices (e.g. waiting near ports), which inflate fuel consumption and congestion (Oh et al., 2021).

However, method choice defines trade-offs. Physics-informed or explainable models support traceability, while black-box regressors are easier to deploy but require monitoring (Wang et al., 2023). Limitations remain, narrow ship and route coverage, and sensitivity to the quality of atmospheric and oceanic condition data (e.g., wind, wave, current, and temperature). In addition, without reliable benchmarks relative to operator baselines, claimed optimisation benefits remain difficult to validate in operational terms, limiting both external validity and industry adoption.

4.3.3 Sustainable Supply Chain Practices

With sustainability moving to the forefront of global supply chain agendas, digital technologies are increasingly deployed to align eco-efficiency with operational resilience. In container ports, IoT-enabled cold chain monitoring is a prominent example, reducing spoilage and improving traceability for temperature-sensitive goods. By integrating wireless sensor networks (WSN) and RFID systems, ports enable real-time data capture and decision support, which strengthens logistics planning while advancing environmental goals. These innovations demonstrate how operational performance and sustainability objectives can be mutually reinforcing (Cil et al., 2022).

Beyond the port level, predictive analytics extend sustainability efforts to network-wide planning. A gravity-inspired deep learning framework, trained on trade flows and network metrics, has been shown to forecast vessel movements with high accuracy. Its outputs have practical applications in ecological risk management, such as assessing the spread of invasive species through ballast water, thus linking traffic prediction with sustainable maritime policy design (Song et al., 2024).

Geopolitical disruptions further highlight the importance of resilient and adaptive energy logistics. Leveraging AIS trajectory data, GDELT news datasets, and quadratic assignment procedure (QAP) regression, one study examined the impact of the Russia–Ukraine conflict on the liquified natural gas (LNG) marine supply network. Results reveal shifts in centrality, competition, and resilience, underscoring LNG’s role as a transitional low-carbon fuel. This interdisciplinary application illustrates how data-driven approaches can inform both sustainability governance and supply chain adaptation under systemic shocks (Xiao et al., 2025).

Taken together, these studies illustrate how BDA enables sustainability by linking operational monitoring with network-level forecasting. IoT-based traceability strengthens cold chains, deep learning improves flow forecasting for ecological and environmental management, and AIS-based network analysis captures the systemic effects of geopolitical crises. However, benefits remain

contingent on sensor quality, robust transmission systems, and interoperable data standards. Generalisability is limited, with adoption constrained by governance challenges around data sharing, cybersecurity, and regulatory alignment.

4.3.4 Synthesis Commentary – Sustainability

The sustainability cluster demonstrates that BDA offers strong potential to advance environmental goals in maritime trade facilitation, particularly in emissions monitoring, fuel optimisation, and supply chain resilience. Evidence of effectiveness includes IoT-enabled cold chain monitoring, fuel consumption models incorporating emissions objectives, and LNG network analyses that integrate energy and environmental perspectives.

Despite methodological sophistication—ranging from physics-informed neural networks to gravity-inspired forecasting models—most studies remain confined to pilot projects or region-specific deployments. Fragmented data sources, inconsistent standards, and infrastructure disparities continue to limit model transferability and scalability. As a result, many frameworks remain at the simulation or proof-of-concept stage, rather than integrated into live decision-making systems.

Interdisciplinary approaches are emerging as a promising direction. For example, the integration of AIS data, global event datasets such as GDELT, and QAP regression offers a means of quantifying resilience and geopolitical shocks in ways that combine environmental, energy, and policy insights (Xiao et al., 2025). Similarly, optimisation frameworks that internalise carbon costs show how policy mechanisms can be directly linked with operational planning (Filom & Razavi, 2025).

A key finding is that sustainability is often treated as a secondary outcome within studies primarily focused on efficiency or resilience. However, a growing body of work positions sustainability as a primary objective, with emissions reduction explicitly embedded into planning models.

In conclusion, the literature reflects a dual trend: operational success in specific contexts and broader strategic ambition in conceptual frameworks. The critical integration challenge lies in developing auditable data pipelines, explainable models, and policy-aware systems that can be scaled across jurisdictions. Without stronger institutional alignment and harmonised governance structures, the full potential of BDA to drive maritime sustainability will remain only partially realised.

4.4 Thematic Cluster 3: Big Data Analytics for System Resilience

This cluster examines how Big Data Analytics (BDA) strengthens resilience in maritime trade facilitation, focusing on early disruption detection, network vulnerability assessment, and adaptive strategies to contain cascading failures.

4.4.1 Disruption Detection & Response

As disruption risks intensify, BDA is increasingly applied to support event sensing and dynamic response. Real-time detection of shocks—such as severe weather, chokepoint closures, or targeted disruptions—relies on whether models can detect events early enough to alter operations and whether responses can adapt quickly.

AIS-driven analytics have been used to identify where disruptions propagate most rapidly and which ports are structurally vulnerable. Operational tools such as case-based reasoning combine historical records with live AIS feeds to detect vessel delays from storms or congestion, enabling earlier interventions and mitigating downstream costs (Kim et al., 2017).

Graph-based deep learning extends predictive capacity beyond hazard-driven monitoring. Building on its role in port optimisation, the MSTGNN framework captures inter-port interactions and dynamic fluctuations, enabling proactive congestion detection and early re-planning across multiple vessel types and regions (Mei et al., 2025).

Resilience also requires adaptive optimisation. A learning-based robust optimisation model for synchromodal freight transport integrates Bayesian neural networks with adjustable robust sets, producing flexible shipment-matching policies that mitigate travel-time variability and strengthen responsiveness to unforeseen delays (Filom & Razavi, 2025).

The concept of Industry 5.0 has also been extended to resilience. One framework embeds IoT-enabled monitoring and robust optimisation into a dry port–seaport system, reducing tardiness and increasing transparency by closing the loop from detection to response (Rekabi & Sazvar, 2025).

At the network scale, AIS and graph deep learning have been applied to the European LNG supply chain, showing that targeted node losses degrade connectivity far more severely than random failures. These findings emphasise the importance of diversification, port security, and real-time rerouting for systemic resilience (Mei et al., 2024).

Collectively, these approaches highlight the value of linking hazard feeds, predictive analytics, and adaptive optimisation. Yet challenges remain: models trained on one region often underperform elsewhere, and black-box methods lack the interpretability needed for operational trust.

4.4.2 Network Vulnerability & Cascading Failures

While early detection enhances short-term response, resilience also depends on understanding how disruptions cascade across maritime networks.

AIS-based analyses reveal that structural growth does not guarantee robustness. In the European LNG supply chain, shipments and clustering increased between 2018 and 2020, yet targeted removals of critical ports still triggered severe connectivity losses (Mei et al., 2024). Complementary simulations of LNG import networks confirm that intentional disruptions at chokepoints can rapidly fragment the system, with resilience hinging on only a few critical nodes (Savvakis & Moschovou, 2025). Similarly, closures in the Taiwan Strait caused by typhoons propagate delays and costs exceeding USD

320 million per day, underscoring the fragility of global trade when redundancy is limited (Liang et al., 2025).

Advanced techniques extend vulnerability analysis into actionable diagnostics. Methods such as community detection and the two-hop biconnected component (TBCC) algorithm identify critical clusters, backup routes, and redundancy levels (Liu et al., 2023). At the liner-network level, cascade models that incorporate realistic port-rotation behaviour replicate congestion spillover and quantify critical sets of ports whose simultaneous disruption would have the largest cascading effects (Xu et al., 2025).

These approaches shift resilience planning from isolated asset assessment to system-wide stress-testing. TBCC diagnostics highlight clusters that preserve rerouting capacity, attack simulations guide contingency plans, and commodity-specific analyses (e.g., LNG networks) inform diversification and buffer strategies. Together, they equip policymakers and operators with tools to strengthen resilience before disruptions occur.

4.4.3 Synthesis Commentary – Resilience

The literature reflects two distinct but complementary layers of resilience:

- Operational resilience: near-term continuity achieved through event sensing, congestion forecasting, and uncertainty-aware re-planning.
- Strategic resilience: long-term system design focusing on diversification of gateways, redundancy at chokepoints, and pre-authorised contingency playbooks.

Empirical evidence demonstrates the effectiveness of graph learning in providing early warnings, robust synchronodal optimisation in mitigating variability, and IoT-based monitoring in reducing detection-to-action cycles. Strategic studies quantify the benefits of redundancy and diversification, though most remain simulations rather than implemented reforms.

Persistent gaps include the absence of codified cross-border coordination protocols, limited availability of longitudinal datasets, and the uneven maturity of

resilience analytics. Detection models are close to deployment readiness, predict-then-optimize policies are mid-stage but context dependent, while stress-testing methods are advanced analytically but lack institutional uptake.

In sum, resilience requires not only technical analytics but also governance structures that translate detection into actionable, cross-stakeholder responses.

4.5 Cross-Theme Analysis: Barriers and Enablers

Across the three thematic clusters—Efficiency (§4.2), Sustainability (§4.3), and Resilience (§4.4)—a recurring pattern emerges: technical innovation is advancing, but systemic adoption is constrained by challenges of data interoperability, organisational alignment, and governance frameworks.

Technical barriers include fragmented data systems, inconsistent identifiers, and quality gaps such as AIS dropouts and sensor drift. For efficiency, fragile feedback loops hinder integration between forecasts and terminal control. For sustainability, black-box models complicate emissions accounting. For resilience, cascade simulations are limited by static schedules and simplified assumptions.

Organisational barriers arise from siloed processes and uneven analytics capacity. Port, yard, and inland teams often work with divergent indicators. Efficiency gains remain expert-driven rather than institutionalised. Sustainability suffers from weak integration between energy management and operational planning. Resilience studies model scenarios but rarely translate them into codified activation thresholds or tested protocols.

Governance barriers are perhaps the most systemic. Data-sharing is restricted by privacy, competition, and sovereignty concerns. Regulatory bodies impose varying requirements for interpretability. Efficiency is hampered by contracts that penalise schedule changes. Sustainability requires auditable models, but black-box predictors fall short. Resilience lacks formal cross-border agreements for capacity sharing or rerouting authority.

A useful framing is the Sense–Predict–Decide–Act loop:

- Sense: event-driven platforms integrate IoT telemetry, AIS, and dry port–seaport data streams, providing traceable visibility.
- Predict: models transform signals into foresight, with physics- or knowledge-informed ML ensuring interpretability and audit trails.
- Decide: uncertainty-aware optimisation and decision-focused learning generate robust policies rather than point solutions.
- Act: automation of documentation, pre-tested contingency templates, and rapid execution protocols shorten the cycle from detection to intervention.

In synthesis, BDA has shown measurable improvements in berth times, emissions reduction, and disruption recovery. Yet the transition from pilot projects to systemic deployment depends on three priorities:

1. Standardising the data backbone across maritime stakeholders.
2. Embedding model governance to ensure transparency and accountability.
3. Codifying shared action rights to enable coordinated responses across jurisdictions.

Only by addressing these systemic enablers can BDA move from technical promise to operational transformation in maritime trade facilitation.

4.6 Synthesis Summary

This chapter synthesised 22 core papers, drawn from the 55 studies selected through the review, into three outcome-oriented clusters: efficiency, sustainability, and resilience. Cross-cutting barriers, enablers, and decision-support frameworks were also examined. The analysis directly aligns with the study’s four research questions, which address tool and data mapping (RQ1), impact assessment (RQ2), barrier identification (RQ3), and implementation frameworks (RQ4).

For RQ1 (Mapping), Sections 4.2 to 4.4 catalogued the BDA methods reported in the literature, spanning optimisation, simulation, and AI/ML applications. Use cases included berth scheduling, customs targeting, traffic forecasting, and fuel

modelling, supported by diverse data sources such as AIS, IoT sensors, sea-weather data, operational logs, and market indicators.

For RQ2 (Impact), evidence was consolidated across both operational and systemic levels. Documented improvements included reduced berth times, enhanced inspection yields, and improved coordination, alongside broader insights such as congestion avoidance and detour feasibility derived from network analytics.

For RQ3 (Barriers), recurring challenges were identified in three categories: technical (interoperability, data quality), organisational (siloe processes, skill shortages), and governance-related (legal constraints, fragmented authority). These constraints consistently limited the scalability of otherwise promising approaches.

For RQ4 (Frameworks), the review identified decision-support models that link prediction to operational execution, including robust and synchmodal optimisation policies. These frameworks emphasise scalability and decision relevance, although most remain at prototype or pilot stage.

Three cross-cutting patterns emerged across clusters. First, integrated systems consistently outperform isolated pilots, particularly when sensing, prediction, and decision-making are embedded into operations. Second, model transparency matters: while black-box methods often achieve high accuracy, interpretable or physics-informed approaches are more viable in institutional settings. Third, network structure is decisive: resilience depends not on reactive rerouting but on pre-planned redundancy and feasible detours. Despite these insights, scaling remains difficult. Localised successes rarely generalise across multiple ports, and efficiency gains built on bespoke infrastructures often diminish without standardised governance.

The evidence base can be divided into areas of strength and weakness. Stronger evidence supports: (i) flow detection and nowcasting using knowledge-infused graph learning; (ii) decision-linked improvements where forecasts directly shape dispatch, inspection, or scheduling; and (iii) network stress-

testing that identifies critical nodes and quantifies redundancy payoffs. Weaker evidence surrounds: (i) cross-border coordination models that formalise triggers, capacity sharing, and decision rights; (ii) longitudinal datasets that capture policy, climate, or behavioural regime shifts; and (iii) systematic adoption at institutional scale, particularly in jurisdictions where standards of explainability and accountability vary.

Overall, the findings suggest that progress depends less on algorithmic innovation alone and more on architectural and institutional alignment. Scalability improves when organisations standardise data infrastructures, implement model governance, and operationalise analytics through a Sense–Predict–Decide–Act loop that links data streams to execution. Where interoperable data, coordinated procedures, and shared action rights are present, BDA delivers sustained improvements. In their absence, results remain fragmented, case-specific, and short-lived.

5 Discussion

5.1 Introduction

This chapter interprets the thematic synthesis presented in Chapter 4, directly addressing the research questions while reaffirming the overarching aim of the study: to assess how Big Data Analytics (BDA) can enhance efficiency, sustainability, and resilience in maritime trade facilitation.

The discussion proceeds question by question:

- RQ1 examines the mapping of tools, data sources, and application domains.
- RQ2 assesses reported impacts on efficiency, sustainability, and resilience.
- RQ3 analyses cross-cutting barriers to adoption.
- RQ4 evaluates the scope, scalability, and decision relevance of proposed frameworks.

Beyond answering these questions, the chapter highlights cross-cutting insights, identifies gaps and contradictions, and situates findings in both academic and practical contexts.

5.2 Addressing the Research Questions

5.2.1 RQ1 – Mapping of BDA Tools, Data Sources, and Applications

The 55 mapped studies show tool clusters concentrated in graph and network analytics ($\approx 36.4\%$), time-series forecasting and machine/deep learning (each $\approx 21.8\%$), and optimisation and prescriptive methods ($\approx 16.4\%$). Simulation, text mining, and topological analysis appear sporadically ($\leq 3.6\%$). Application domains are heavily weighted towards port/terminal operations and supply-chain planning ($\approx 63\text{--}67\%$), while customs, emissions, network vulnerability, and security applications remain marginal.

Data sourcing follows a multi-label pattern. The most common are market and economic indices (56.4%) and AIS/sensor streams (47.3%), with less frequent

reliance on operational logs, customs records, weather data, schedules, or unstructured text. Studies typically combine one or two categories, most often AIS with either market data or weather signals.

This mapping highlights two important observations. First, technical depth is strongest in operational domains, particularly scheduling and throughput forecasting. Second, governance-related contexts—such as customs risk targeting, cross-border data exchange, and regulatory text mining—are underrepresented. This imbalance suggests that while BDA is maturing technically, its application to institutional and policy dimensions of trade facilitation remains thin.

5.2.2 RQ2 – Impact on Efficiency, Sustainability, and Resilience

Evidence is most robust for efficiency. Gains include reduced berth times, shortened queues, and higher inspection yields, with some pilots reporting large improvements (e.g., 68% reduction in tardiness and 71% increase in transparency through IoT–forecasting–optimisation integration; >20% increase in inspection detection rates via predictive scheduling).

For sustainability, benefits are mixed. Emissions monitoring has improved significantly, but verified reductions remain modest: 2–7% fuel savings from speed optimisation and up to 15% reductions in emission charges under policy-aware planning.

For resilience, the literature offers analytically persuasive evidence (early warning, cascade simulations), but most contributions remain confined to modelling and simulation. Institutionalised playbooks and cross-jurisdiction applications are rare.

Overall, efficiency evidence is abundant and operational; resilience evidence is strong analytically but thin institutionally; and sustainability studies emphasise monitoring more than systemic decarbonisation. Context dependency is clear: short-term forecasts embed easily into dispatching, mid-term models require calibrated risk governance, and network-level resilience hinges on data sharing and pre-agreed decision rights.

5.2.3 RQ3 – Barriers to BDA Adoption

Barriers cluster into technical, organisational, and governance domains:

- Technical: fragmented identifiers, AIS dropouts, sensor drift, and incomplete “predict–optimise–act” loops undermine reliability.
- Organisational: siloed teams optimise for divergent KPIs; analytics capacity is uneven; and decision rights to act on outputs are often undefined.
- Governance: privacy, commercial sensitivities, and heterogeneous interpretability standards constrain data sharing and black-box model adoption.

Variation is systematic. Public authorities emphasise transparency and due process, whereas commercial actors prioritise ROI and seamless integration. Developed regions face high contractual and liability constraints, while developing regions struggle with infrastructure gaps and inconsistent data standards.

The key contradiction is that single-port pilots often report smooth adoption, while multi-jurisdictional settings repeatedly encounter interoperability, legal clearance, and decision-right challenges. Without standardised data spines, explainable models, and codified governance, adoption remains pilot-bound.

5.2.4 RQ4 – Decision-Support Frameworks and Models

Frameworks vary by scope:

- Operational models couple forecasts to policies, e.g., queueing + ML for berth allocation, decision-focused inspection scheduling, and uncertainty-aware synchromodal freight planning.
- Systemic models stress network design, e.g., cascade simulations, TBCC diagnostics, carbon-costed optimisation, and IoT-enabled Industry 5.0 architectures.

Scalability is uneven. Local integration into terminal systems and rosters yields repeatable gains, but transferability across ports is fragile without standardised

data backbones and calibrated uncertainty. Stress-testing generalises better, yet real-world implementation remains scarce.

Validation is a persistent gap. Most studies rely on simulations or historical backtests, with limited cross-jurisdiction trials. Frameworks that embed Sense–Predict–Decide–Act loops with data provenance, interpretability, and pre-authorised action rights show the highest decision relevance, though few examples are institutionalised.

5.3 Key Gaps, Contradictions, and Emerging Themes

Three themes characterise the evidence base:

1. Validation gaps: Many strong models stop at predictive or structural insights without audited outcomes for time, cost, or emissions. Examples include topology-driven analyses that map networks without linking to live KPIs (Liu et al., 2023) or physics-informed forecasting disconnected from operational governance (Song et al., 2024).
2. Divergent outcomes: Similar toolsets yield varying effects. For instance, AIS-based delay detection enables actionable responses (Kim et al., 2017), while governance-focused AI proposals report high model scores but little workflow change (Priya et al., 2024). Success often depends more on data maturity and process integration than algorithm choice.
3. Contradictions in governance: IoT-enabled cold chain monitoring requires sensitive data sharing (Cil et al., 2022), while permissioned architectures safeguard privacy but risk undermining real-time responsiveness (Priya et al., 2024). A practical compromise is minimum-necessary, auditable sharing.

Emerging but under-evidenced directions include LNG network resilience as a tool for corridor diversification (Mei et al., 2024), carbon-aware synchromodal optimisation (Filom & Razavi, 2025), and digital twin applications linking early detection to adaptive optimisation (Cil et al., 2022; Kim et al., 2017).

5.4 Practical Implications and Academic Relevance

- Port operators should assess infrastructure and integration costs carefully. Evidence shows Industry 5.0 frameworks reduce tardiness and boost transparency, but only under high-investment conditions (Rekabi & Sazvar, 2025).
- Customs authorities benefit most when prediction is embedded in inspection scheduling, raising detection yields by over 20% (Yan et al., 2021).
- Policy-makers should prioritise interoperable data spines, role-based access, and pre-authorised decision rights, tested via cross-border simulations. LNG case studies demonstrate how stress-testing supports diversification and green-corridor planning (Mei et al., 2024).
- Technology vendors should deliver modular, API-first solutions with audit trails and explainability, balancing accuracy with interpretability. Recent physics-informed models show that explainability can be deliberately designed (Wang et al., 2023).

Academically, this thesis consolidates fragmented research into a structured framework around efficiency, sustainability, and resilience. It identifies where evidence is strong (efficiency), where it is weak (resilience, governance), and where contradictions persist. Practically, it emphasises a shift from prediction to decision-grade analytics under auditable governance.

Scaling BDA in maritime trade facilitation requires forecast-to-dispatch pilots, standardised transparency protocols, and institutionalised scenario-tested playbooks. These priorities form the foundation for sustained, cross-jurisdictional transformation.

6 Conclusion

6.1 Summary of Findings

This study systematically reviewed 55 peer-reviewed articles to evaluate how Big Data Analytics (BDA) supports maritime trade facilitation, expanding the concept beyond customs clearance to include port operations, governance processes, and systemic resilience. The analysis was guided by a pre-registered systematic review protocol designed to minimise bias and enhance traceability.

The findings highlight that most BDA applications cluster around graph and network analytics, time-series forecasting, and machine learning, applied primarily to port/terminal optimisation, customs targeting, and network modelling. Simulation and text mining appear infrequently. Data sources are dominated by publicly available or aggregated datasets (e.g. AIS, market indices), with limited cross-agency regulatory data—raising concerns about auditability and scalability.

Impacts are most evident in efficiency, including measurable improvements in berth scheduling, throughput forecasting, and inspection targeting. Sustainability evidence is uneven, with stronger results for emissions monitoring and visibility than for audited, causally attributed reductions in emissions or fuel use. Resilience benefits are largely analytical, offering early warnings and cascade simulations, but rarely embedded institutionally. The scalability of these gains across ports, corridors, or jurisdictions remains uncertain.

Barriers recur across three domains: technical (interoperability, data quality, model governance), organisational (siloes KPIs, skill gaps, last-mile integration), and legal/regulatory (privacy, liability, interpretability). Stakeholder priorities differ: public agencies emphasise transparency and due process, whereas private actors focus on ROI and operational continuity. Regional variation further conditions adoption.

Few decision-support frameworks achieve decision-grade validation. Many remain limited to simulations or offline tests. Models that scale successfully do

so by embedding forecasts into executable policies supported by data provenance, interpretability, and pre-authorised decision rights—that is, by operationalising a Sense–Predict–Decide–Act loop. Without these governance anchors, results remain pilot-specific and context-bound.

Overall, BDA’s potential is evident, but its lasting impact depends less on novel algorithms than on standardised data spines, embedded model governance, and cross-border authority to act.

6.2 Contributions of the Study

This thesis demonstrates that BDA can substantially enhance efficiency, sustainability, and resilience in maritime trade facilitation. However, the decisive challenge lies not in developing more advanced algorithms, but in standardising data infrastructures, embedding governance mechanisms, and empowering cross-border decision rights. Meeting these conditions is essential for scaling from promising pilots to repeatable, corridor-level outcomes, and for ensuring that BDA becomes a lasting enabler of global trade facilitation.

This thesis makes three sets of contributions: academic, methodological, and practical/policy-oriented.

Academic contributions:

- It is one of the first systematic reviews to examine Big Data Analytics (BDA) explicitly within the scope of maritime trade facilitation, expanding the focus beyond ports or customs individually to include multi-stakeholder ecosystems.
- It develops a three-cluster framework (efficiency, sustainability, resilience) that integrates fragmented literature and allows structured comparison across diverse BDA applications.
- It highlights underexplored domains such as customs automation, governance-aware resilience planning, and emissions accounting, thereby identifying gaps for future academic inquiry.

Methodological contributions:

- By applying a transparent SLR protocol (search strategy, inclusion/exclusion criteria, PRISMA screening, extraction matrix), this study establishes a replicable evidence base that improves on prior narrative reviews in the field.
- It introduces a Sense–Predict–Decide–Act loop as an evaluative lens for decision-support frameworks, offering a systematic way to assess whether BDA outputs translate into operational action.
- It synthesises evidence into strong-evidence versus limited-evidence categories, differentiating between areas of validated impact (e.g., berth scheduling, inspection targeting) and those requiring further validation (e.g., cross-border coordination, long-term robustness).

Practical and policy contributions:

- For port operators and customs authorities, the study identifies actionable practices for scaling BDA, such as establishing shared data spines, codifying triggers, and piloting forecast-to-dispatch integration.
- For policy-makers, it clarifies enabling conditions—interoperable standards, role-based access, and pre-authorised decision rights—that must be institutionalised to achieve corridor-level benefits.
- For technology vendors, it specifies design imperatives such as API-first modular architectures, embedded governance features, and explainability as essential for adoption across heterogeneous operational contexts.

In combination, these contributions advance both scholarly understanding and practical adoption pathways, positioning BDA not only as a technical toolkit but as an institutional enabler of modernised, resilient, and sustainable maritime trade.

6.3 Limitations

Several limitations frame the scope and interpretation of this review:

- Evidence base: restricted to peer-reviewed, English-language journal articles (2015–2025). This ensured quality and comparability but

excluded grey literature and non-English sources that may capture operational practices and regional innovations.

- Domain scope: confined to maritime trade facilitation. Multimodal or hinterland perspectives were only considered where directly linked to seaborne flows, limiting applicability to fully synchromodal contexts.
- Synthesis approach: thematic clustering relied on interpretive coding across heterogeneous studies. Although supported by a structured extraction matrix and PRISMA workflow, categorising benefits under efficiency, sustainability, or resilience inevitably involved subjectivity. The heterogeneity of study designs precluded formal meta-analysis.
- Methodological balance: the corpus is overwhelmingly quantitative ($\approx 90.9\%$), leaving organisational and governance mechanisms underexplored. Qualitative or mixed-methods research—which often better captures institutional and behavioural dynamics—was scarce.
- Temporal scope: the review spans 2015–2025. Developments in standards, regulation, or digital tools emerging after mid-2025 are not reflected. More broadly, most included studies validated results only through simulations or single-site pilots, constraining generalisability and decision-grade relevance.

These limitations should be borne in mind when interpreting the findings and assessing their transferability.

6.4 Recommendations for Practice and Future Work

For Port and Customs Authorities:

- Treat BDA as a core operational capability, not a peripheral experiment.
- Establish a shared data spine with clean identifiers, API gateways, and traceable data provenance.
- Prioritise forecast-to-dispatch pilots (e.g., linking ETA to berth rules or risk scores to inspection schedules) with baseline performance indicators and explainability.

- Institutionalise codified triggers and decision rights to ensure data-driven decisions translate into action across agencies.

For Policy-Makers:

- Develop corridor-level standards for data models, access controls, and cross-border decision rights.
- Institutionalise scenario-tested playbooks for disruption response and decarbonisation, including pre-authorised triggers for rerouting or reprioritisation.
- Channel funding towards multi-port trials that require explainable models, audit trails, and monitoring-ready outputs, not just algorithmic performance.

For Technology Vendors:

- Deliver modular, API-first solutions with governance features such as provenance tracking, audit logs, and role-based access.
- Design models that balance accuracy with interpretability, combining domain knowledge with ML for institutional acceptability.
- Package capabilities into interoperable modules that can replicate across diverse port and customs systems.

For Researchers:

- Move beyond simulation to decision-grade validation in live operational settings.
- Establish shared outcome definitions (time, cost, emissions, recovery) and uncertainty reporting to enable comparability.
- Undertake comparative, cross-jurisdictional trials and open-benchmark initiatives to test transferability.
- Investigate emerging methods (e.g., graph learning, physics-informed models, decision-focused optimisation) alongside socio-technical questions of trust, accountability, and data-sharing ethics, using mixed-methods approaches.

References

- Abouelrous, A., Dekker, R., Bliet, L., & Zhang, Y. (2025). Real-time policy for yard allocation of transshipment containers in a terminal. *Transportation Research Part B: Methodological*, 192. <https://doi.org/10.1016/j.trb.2024.103138>
- Ben Haj Ahmed, R., Bouzir, A., Benhadj Mbarek, M. H., & Benammou, S. (2023). THE IMPACT OF LOGISTICS PERFORMANCE INDEX ON PORT INFRASTRUCTURE QUALITY: COMPARATIVE STUDY TUNISIA MOROCCO. *Archives of Transport*, 67(3), 71–89. <https://doi.org/10.5604/01.3001.0053.7212>
- Brouer Berit Dangaard and Karsten, C. V. and P. D. (2016). Big Data Optimization in Maritime Logistics. In A. Emrouznejad (Ed.), *Big Data Optimization: Recent Developments and Challenges* (pp. 319–344). Springer International Publishing. https://doi.org/10.1007/978-3-319-30265-2_14
- Cil, A. Y., Abdurahman, D., & Cil, I. (2022). Internet of Things enabled real time cold chain monitoring in a container port. *Journal of Shipping and Trade*, 7(1). <https://doi.org/10.1186/s41072-022-00110-z>
- Filom, S., & Razavi, S. (2025). A learning-based robust optimization framework for synchromodal freight transportation under uncertainty. *Transportation Research Part E: Logistics and Transportation Review*, 195. <https://doi.org/10.1016/j.tre.2025.103967>
- Gough, D., Thomas, J., & Oliver, S. (2012). Clarifying differences between review designs and methods. *Systematic Reviews*, 1, 1–9.
- Heyvaert, M., Maes, B., & Onghena, P. (2013). Mixed methods research synthesis: Definition, framework, and potential. *Quality and Quantity*, 47(2), 659–676. <https://doi.org/10.1007/s11135-011-9538-6>
- Kim, S., Kim, H., & Park, Y. (2017). Early detection of vessel delays using combined historical and real-time information. *Journal of the Operational*

- Research Society*, 68(2), 182–191. <https://doi.org/10.1057/s41274-016-0104-4>
- Lame, G. (2019). Systematic Literature Reviews: An Introduction. *Proceedings of the Design Society: International Conference on Engineering Design*, 1(1), 1633–1642. <https://doi.org/10.1017/dsi.2019.169>
- Liang, M., Cai, Y., Chen, T., Wang, H., & Meng, Q. (2025). Data-driven impact analysis of chokepoint on multi-scale maritime networks: A case study of the Taiwan Strait. *Transportation Research Part E: Logistics and Transportation Review*, 200. <https://doi.org/10.1016/j.tre.2025.104179>
- Liu, X., Xu, Y., Xu, M., Deng, W., Pan, L., & Ng, A. K. Y. (2023). Two-Hop Biconnected Components in the Global Liner Shipping Network Reveal International Trade Statuses. *IEEE Transactions on Network Science and Engineering*, 10(3), 1564–1574. <https://doi.org/10.1109/TNSE.2022.3230740>
- Mei, Q., Li, Z., Hu, Q., Zhi, X., Wang, P., Yang, Y., & Liu, X. (2025). Spatio-temporal graph neural network fused with maritime knowledge for predicting traffic flows in ports. *Regional Studies in Marine Science*, 85. <https://doi.org/10.1016/j.rsma.2025.104106>
- Mei, Q., Qinyou, H., Hu, Y., Yang, Y., Liu, X., Huang, Z., & Wang, P. (2024). Structural analysis and vulnerability assessment of the European LNG maritime supply chain network (2018–2020). *Ocean and Coastal Management*, 253. <https://doi.org/10.1016/j.ocecoaman.2024.107126>
- Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Medicine*, 6(7), e1000097. <https://doi.org/10.1371/journal.pmed.1000097>
- Munn, Z., Tufanaru, C., & Aromataris, E. (2014). JBI's Systematic Reviews. *AJN, American Journal of Nursing*, 114(7), 49–54. <https://doi.org/10.1097/01.NAJ.0000451683.66447.89>

- Oh, M. J., Roh, M. Il, Park, S. W., Chun, D. H., Son, M. J., & Lee, J. Y. (2021). Operational analysis of container ships by using maritime big data. *Journal of Marine Science and Engineering*, 9(4). <https://doi.org/10.3390/jmse9040438>
- Okoli, C. (2015). A guide to conducting a standalone systematic literature review. *Communications of the Association for Information Systems*, 37.
- Poiani, R., Stirbu, C., Metelli, A. M., & Restelli, M. (2024). Optimizing Empty Container Repositioning and Fleet Deployment via Configurable Semi-POMDPs. *IEEE Transactions on Intelligent Transportation Systems*, 25(5), 4704–4711. <https://doi.org/10.1109/TITS.2023.3329677>
- Priya, J. C., Rudzki, K., Nguyen, X. H., Nguyen, H. P., Chotechuan, N., & Pham, N. D. K. (2024). Blockchain-Enabled Transfer Learning for Vulnerability Detection and Mitigation in Maritime Logistics. *Polish Maritime Research*, 31(1), 135–145. <https://doi.org/10.2478/pomr-2024-0014>
- Rekabi, S., & Sazvar, Z. (2025). A smart and agile dry port-seaport logistic network considering industry 5.0: A multi-stage data-driven approach. *Socio-Economic Planning Sciences*, 98. <https://doi.org/10.1016/j.seps.2024.102141>
- Savvakis, K. I., & Moschovou, T. P. (2025). European LNG Import Network Analysis and Investigation of Supply Security. *Energies*, 18(3). <https://doi.org/10.3390/en18030634>
- Song, R., Spadon, G., Pelot, R., Matwin, S., & Soares, A. (2024). Enhancing global maritime traffic network forecasting with gravity-inspired deep learning models. *Scientific Reports*, 14(1). <https://doi.org/10.1038/s41598-024-67552-2>
- Su, Z., Li, J., Pang, Q., & Su, M. (2025). China futures market and world container shipping economy: An exploratory analysis based on deep learning. *Research in International Business and Finance*, 76. <https://doi.org/10.1016/j.ribaf.2025.102870>

- Tian, X., Guan, Y., & Wang, S. (2023). A Decision-Focused Learning Framework for Vessel Selection Problem. *Mathematics*, 11(16). <https://doi.org/10.3390/math11163503>
- Tranfield, D., Denyer, D., & Smart, P. (2003). Towards a Methodology for Developing Evidence-Informed Management Knowledge by Means of Systematic Review. *British Journal of Management*, 14(3), 207–222. <https://doi.org/10.1111/1467-8551.00375>
- Wang, H., Yan, R., Wang, S., & Zhen, L. (2023). Innovative approaches to addressing the tradeoff between interpretability and accuracy in ship fuel consumption prediction. *Transportation Research Part C: Emerging Technologies*, 157. <https://doi.org/10.1016/j.trc.2023.104361>
- Wilson, J. S., Mann, C. L., & Otsuki, T. (2005). Assessing the Benefits of Trade Facilitation: A Global Perspective. *The World Economy*, 28(6), 841–871. <https://doi.org/https://doi.org/10.1111/j.1467-9701.2005.00709.x>
- Xiao, R., Zhao, P., Zhang, M., Kang, T., & Ma, T. (2025). Discovering Changes in the Global Liquefied Natural Gas Marine Supply Network During the Russian-Ukrainian Crisis with Big Data. *Chinese Geographical Science*, 35(2), 295–310. <https://doi.org/10.1007/s11769-024-1467-0>
- Xu, M., Shen, Y., Zhu, Y., Hong, L., Ma, M., Lü, L., & Li, K. X. (2025). Identification of critical sets of ports in cascading failures of global liner shipping network. *Transportation Research Part E: Logistics and Transportation Review*, 200. <https://doi.org/10.1016/j.tre.2025.104184>
- Yan, R., Wang, S., Cao, J., & Sun, D. (2021). Shipping Domain Knowledge Informed Prediction and Optimization in Port State Control. *Transportation Research Part B: Methodological*, 149, 52–78. <https://doi.org/10.1016/j.trb.2021.05.003>
- Yan, R., Wang, S., & Du, Y. (2020). Development of a two-stage ship fuel consumption prediction and reduction model for a dry bulk ship.

Transportation Research Part E: Logistics and Transportation Review, 138.
<https://doi.org/10.1016/j.tre.2020.101930>

Yan, R., Yang, D., Wang, T., Mo, H., & Wang, S. (2024). Improving ship energy efficiency: Models, methods, and applications. *Applied Energy*, 368.
<https://doi.org/10.1016/j.apenergy.2024.123132>

Zhang, X., & Cao, W. (2025). Research on Time Series Forecasting Method Based on Autoregressive Integrated Moving Average Model with Zonotopic Kalman Filter. *Sustainability (Switzerland)*, 17(7).
<https://doi.org/10.3390/su17072993>

Zhou, S., Zhang, N., Duan, Q., Xiao, J., & Yang, J. (2024). Automatic Scheduling Method for Customs Inspection Vehicle Relocation Based on Automotive Electronic Identification and Biometric Recognition. *Algorithms*, 17(11). <https://doi.org/10.3390/a17110483>

Appendices

Appendix A Ethical approval letter



22 May 2025

Dear Ms Lai ,

Reference: CURES/25275/2025

Project ID: 28780

Title: Big Data Analytics as an Enabler of Trade Facilitation

We are pleased to inform you that you have successfully declared that your research project is a **Literature Review – based solely on openly available literature which is in the public domain, and you are undertaking desk-based research not involving any other form of data or information.**

You have also confirmed that your project **does not meet** any of the literature review specified exceptions, listed both within the relevant section of the CURES form and below:

1) Your supervisor has requested that you apply for approval through CURES because:

- The journals or data are in a sensitive area (please discuss this with your supervisor)
- The project will be embargoed i.e. **will not be publicly available via the Cranfield library immediately or longer term**

2) Approval is/will be specifically required by another external body e.g. journal publishers

Therefore, **you do not require ethical approval** and your CURES application will be automatically closed.

Please keep a copy of this letter safe, if this exception is in relation to your thesis project, you will need to include a copy with your final thesis submission.

If you have any queries, please contact CURES Support.

We wish you every success with your project.

Regards,

CURES Team