



Prifysgol Cymru
Y Drindod Dewi Sant
University of Wales
Trinity Saint David

Research Methods and Professional Development

Student Name: Suhaana Kapali

Student ID: 2538421

MSc Data and Artificial Intelligence

Module: L7 Research Methods and Professional Development

Module Code: ACFA7004

Lecturer: Mohammed Ben Farah

Table of Contents

1. Introduction.....	4
1.1 Context and Background.....	4
1.2 Research Problem	6
1.2.1 Problem Statement	6
1.2.2 Current Challenges	6
1.3 Research Aim and Objectives	7
1.4 Significance of the Research	9
2. Literature Review	10
2.1 Introduction to the Literature Review	10
2.2 Thematic Review of Literature	10
2.2.1 Regulatory Compliance Frameworks.....	11
2.3 Key Contributions and Gaps	12
2.4 Theoretical and Technical Frameworks	12
2.5 Relation to Study	13
2.6 Literature Summary.....	13
3. Research Design and Methodology	14
3.1 Research Design.....	14
3.1.1 Research Approach	14
3.1.2 Justification of the Selected Design	14
3.2 Research Methods and Materials	15
3.2.1 Overview of Research Methods.....	15
3.2.2 Research Strategy.....	15
3.2.3 Proposed System Architecture and Dashboard Design	16
3.2.4 Data Analysis and Evaluation Metrics	17
3.3 Materials.....	17
3.3.1 Research Data	17
3.3.2 Justification of Materials	18
3.3.3 Ethical Considerations Related to Data.....	18
3.3.4 Software Requirements	18
3.3.5 Hardware Requirements.....	19
4. Expected Outcome.....	20
4.1 Expected Results.....	20
4.2 Implications	20
4.3 Limitations	21
5. Timeline	22
REFERENCES.....	23
Appendices	25
Research Ethics Self-Assessment Form	25

Table of Figures

Figure 1 SUEZ	4
Figure 2 Dashboard Architecture and Design	16
Figure 3 Gantt Chart.....	22

1. Introduction

1.1 Context and Background

SUEZ plays one of the leading roles in the global water and waste treatment market, with operations around the world in 40 countries, an employee base of 40,000 people, and Smart Water systems serving over 1,500 clients throughout a wide region. As it understands that mastery of data management and workflow automation is the key to achieving environmental change in local areas, the company has become a target in the field of digital environmental solutions. This dedication corresponds with a large scale change in the waste management structures in the UK, as the industry is now undergoing a digitization process driven mainly by the legal AMP8 regulatory cycle. The cycle will bring in for the first time a major investment of £104 billion, which is to result in an increase of 77% compared to the AMP7 previous phase [1]. Under this new context, methods like telemetry, automated controls, and advanced data analytics are among the most effective ways used to enhance operational performances, detect resource leaks, and provide better customer experiences in water and waste management. In reality, the law setting out how companies report sustainable results has changed in a number of ways like businesses nowadays must comply more strictly with the rules if they want to be sustainable on an international scale and also be in keeping with the latest global standards, such as the EU directive about corporate sustainability reporting and the EU Taxonomy. These regulations have laid down a high threshold which will be a target for generating and verifying sustainability data with an audit trail [2]. Basically, EU Taxonomy is a classification model of economy related activities which sets compliance targets for six ecological aspects. These include climate change, water protection, circular transition, prevention and control of pollution, and protection and restoration of ecosystems and biodiversity.



Figure 1 SUEZ

In addition, CSRD has broadened the sustainability requirements internationally, and the directive brought into effect by the European Sustainability Reporting Standards has the effect of an associated corporate disclosure regime. For that reason, international companies must show environmental performance, quantifiable socio economic footprint, and governance accountability across multifaceted, fragmented, and often data sparse supply chain systems. SUEZ has consistently demonstrated its dedication to sustainable governance by earning an EcoVadis Gold Award for global supply chain sustainability and acquiring the ISO 20400 standard for sustainable purchasing practices [3]. It's in house finding system is mostly used to perform the environmental, social, and company related checks as part of the supplier relationship management. But doing away with these disclosure requirements still involves a fair amount of manual work, thorough checking, and gathering of data in different formats. Such a situation of inefficiency clearly shows how much an enterprise would benefit if these tasks were done using an AI based system that is connected to its operations and automated to an extent.

This company's technological development plan mainly focuses on installing AI processing units systematically on design, operational, and analytical stages, where the emphasis is placed on agent based AI structures that allow for increased automation and contextual understanding. By consuming and distributing millions of operational details extracted from the Internet of Things environment, the systems transform data into decisions that could help improve operations. Computational tools, as the head of data at SUEZ points out, demand strict data management strategies and well focused group capabilities to achieve significant results which only technically effective AI can be realized if it operates on a solidly planned data system [4].

1.2 Research Problem

1.2.1 Problem Statement

Although SUEZ is digitally cultured and has put in place effective methods of maintaining sustainability, when it comes to the full chain creation, tracking, and verification of the environmental, social, and governance data across different supply chains and the company's internal operations, the data continue to be spread out, not centralized, and relying mostly on manual administrative work [5]. This fragmentation that has been systematically happening leads to loss of performance, restricts real time compliance validation, and limits the organization's ability to convert operation telemetry into honesty about its sustainable practices. In essence, it is exposing risks related to regulatory compliance and ability to differentiate in the market.

1.2.2 Current Challenges

This core research problem is defined by four clear technical and structural challenges:

- **Data Fragmentation and Operational Storage:** The dependence of historical business practices on a combination of disconnected spreadsheets and isolated data sheets has been the reason for reporting inconsistencies and analytical inconsistencies being the standard within functions of a business. AI based gaining systems have, in part, addressed the problem of the lack of communication or integration among various systems on a basic level [6]. Though the key environmental metrics remain hidden and isolated, which means they are still not available in the main operations of the company and its different parts now. This situation reflects the wider trend of the industry as per one study, more than 50% of enterprises still face considerable operational challenges when they try to combine as well as make use of environmental data through integration based systems [7].
- **Supply Chain Data Deficiencies:** SUEZ uses specialized data resources to identify the potential effects of the social and environmental aspects of the suppliers from whom the company sources. Still, SUEZ conducts the study in a time limited and manual manner. The fact is also revealed that a lack of completeness, consistency, and standard format of the data is the major factor

that constantly hinders the systematic checking of performance by the suppliers [5]. The study further reveals that around 48 percent of chief obtaining officers have cited the lack of proper or partial emissions data as why an organization is stopped from rolling out a fully automated solution for sustainability workflows [1]. It is also quite understandable from this limitation that the data must be in a certain way for it to produce the required results from it; in other words, garbage in, garbage out.

- **Complexity of Social Value Quantifying Metrics:** SUEZ generated a recorded social value creation of £2.7 billion, below an internal goal of £3 billion [1]. However, the precise impact of distributed supply chains on localized social value indicators is mathematically very complex to model and track, needing integration and alignment of unequal data points from different suppliers in very different geographical locations.
- **Regulatory Compliance and Materiality Thresholds:** The EU Taxonomy requires hard data, so each company must keep fully transparent and auditable records to verify that their economic activities meet the environmental criteria. But manual data collection exposes the organization to significant audit risks and excessive compliance costs. Recent regulatory proposals include materiality thresholds, where activities below 10% of capital expenditure or corporate revenue are excused from eligibility assessments, reducing total data requirements by about 70% [8] [9] [10]. However, the execution of automated reporting still requires advanced data planning models and baseline data readiness before any technical AI application can generate reliable results.

1.3 Research Aim and Objectives

- **Research Aim**

The primary objective of this work is to develop an integrated architecture for AI-enabled data arrangement and dashboards for automated monitoring, verification, and analysis of environmental, social, and governance measures in internal operations and supply networks, decoupling enterprise value from waste volume and enabling auditable compliance reporting.

- **Research Objectives**

To achieve this central research aim, the study will fulfill the following measurable, sequential objectives:

1. **System Audit:** Assess SUEZ's current data management architecture, data flows and reporting workflows to uncover operational blocks caused by data silos, poor data quality and fractured integration across finding and sustainability divisions.
2. **Data Model Design:** Construct a comprehensive, multi dimensional relational data model that maps enterprise operational telemetry and supply chain activities directly to standardized environmental performance metrics and localized social impact indicators.
3. **AI Pipeline Engineering:** Develop and tune automated ML pipelines for analyzing unstructured supplier disclosures using NLP algorithms and automatically highlighting compliance errors or differences with specialized irregularity detection models.
4. **Dashboard Prototype & Verification:** Design a technical proof of concept dashboard architecture leveraging the engineered data pipeline to automate compliance reporting against EU Taxonomy and UK Net Zero frameworks, testing its performance using explicit material recovery and contamination reduction metrics.
5. **Critical Evaluation:** Define clear parameters for industrial deployment by evaluating the completed prototype architecture against current legal regulations, system data traceability requirements, and computational performance constraints.

1.4 Significance of the Research

The research makes an important contribution at the intersection of regulatory compliance frameworks, corporate data science, and circular economy infrastructures. With growing legal disclosure requirements around the world, organizations that can automatically generate and validate sustainability metrics will have a massive competitive advantage.

Methodologically, SUEZ is an illustrative case study with a pre existing digital infrastructure, a large operational scale, and a well documented commitment to environmental leadership. This study presents an academic framework for automated data arrangement with a reproducible model that facilitates the move of large scale industrial operations from manual compliance models to systematic data engineering.

- **Automated Regulatory Reporting:** The engineered system architecture will help to directly translate complex regulatory frameworks, such as the EU Taxonomy and CSRD, into automated data workflows, reducing manual audit risks and compliance costs.
- **Predictive Sustainability Analytics:** Observed data demonstrates that specialized AI implementations can predict corporate sustainability report publication times with more accuracy and isolate client firm parameters with more accuracy, matching human capabilities. This architecture allows dynamic mitigation of risk for complex commercial networks and prediction of resource modelling.
- **Operational Cost Reduction:** The architecture is aimed at large reductions in administrative overheads by replacing manual verification with automated data processing pipelines. Similar automated integration frameworks deployed on asset heavy infrastructure projects have shown baseline labor cost savings of 20.15%, proving the financial viability of moving from traditional volume based service paradigms to high margin, data enabled circular intelligence models.

2. Literature Review

2.1 Introduction to the Literature Review

The main aim of this literature review is to explore the existing academic and industrial literature regarding the integration of automated workflows and advanced telemetry in the sectors of environmental services.

As corporate compliance requirements start to replicate globally, the management of sustainability metrics has evolved from a voluntary corporate social responsibility effort to a complex data engineering challenge. This review systematically evaluates the existing research trends, theoretical frameworks, and technological methodologies to identify the key capabilities and operational limitations of sustainability reporting infrastructures [9]. The scope of this review is limited to literature published in the last five years, with a focus on themes of enterprise data orchestration, artificial intelligence applications in resource logistics, and automated regulatory reporting frameworks. These domains are pulled together to point out the foundation for the analysis of the systemic data gaps that currently affect large scale circular economy operators.

2.2 Thematic Review of Literature

Transitioning from outdated waste management practices to automated circular economy frameworks requires a paradigm shift in the collection and use of operational data. Under traditional waste paradigms, organizations treat data as an afterthought to physical logistics, often isolating information in isolated sources within different divisions [11]. Current research shows that data organization, which is automatic binding, processing, and synchronization of data streams from different software environments, is needed to build contextually aware operational models. Recent studies emphasize that artificial intelligence deployment in corporate resource models cannot deliver tangible performance results without baseline data readiness and strong underlying pipeline architectures [12]. This is a common architectural blockage often referred to as the garbage in, garbage out problem. Automated analytical models can only overcome systemic validation errors when data orchestration is considered. The combination of Internet of Things (IoT) hardware facilities with real time telemetry analytics represents a considerable development for advanced management of

assets. Scientific research shows that, by employing smart sensor arrays within distributed collection networks, companies have been able to constantly track the level of fill, material contents, and the overall equipment condition. The blend of live data and machine learning methods allows asset operators to move from fixed, reactive operating methods to ever changing, predictive analytics. Through algorithmic optimization, fleet logistics have been optimized by academics with the result that vehicles can cut back on the unnecessary number of trips and, as a result drop fuel expenses by almost a third in the form of 10%-15% [13]. However, the combination of the latest deep learning models with robotic sorting systems has been found to improve material recovery rates by 20% on average because of this, computerized data treatment is a big part if one wants to increase efficiency in the resource recovery context [14].

2.2.1 Regulatory Compliance Frameworks

Environmental laws are pushing global companies to change their internal reporting systems. Corporate Sustainability Reporting Directive implementation and the EU Taxonomy's strict classification criteria have brought about the legal requirements of companies' transparency and the need to make their environmental impact report auditable and verified by the companies. These legal shifts call for keeping track of a whole range of metrics within six main environmental categories, with pollution control and circular economy as two examples. From recent research, gathering data manually is an operational risk, and the method of human review is very slow and error prone. Because of this, it cannot meet the need for real time audits. Regulatory changes have come with materiality thresholds that have been used to exclude low revenue activities, thereby, in theory, decreasing total compliance data points by 70%. Despite this, the remaining compliance burden does not allow even the best qualified companies to meet the requirements without a reliable automated validation solution, and because of this, they will have to rely on the long term legal viability it gives [15].

2.3 Key Contributions and Gaps

In the field, special machine learning models have made significant progress in tackling environmental problems. One example is the use of forecasting methods that predicted the issue date of corporate reports with more precision. Scientists have conducted numerous theoretical and practical studies to explore how blockchain can enhance supply chain transparency. It turned out that the use of blockchain could decrease labour costs by 20.15% for large scale infrastructure projects that use a lot of resources [16]. One huge research area that lacks is the systematic way to combine various technical solutions for creating the right corporate systems. The problem currently is that one area of research is not connected with another one very well. Computer science studies most of the time focus on the effectiveness of individual pieces of software, for example, robotic sorting accuracy, while business literature is mostly concerned with theoretical setups of policy. Research that can clearly show a major circular economy operator how to share operational data with multiple public dashboards that contain different types of regulations that are constantly being updated is very scarce [17]. Through this paper, a complete model that is data driven, unified, and orchestrates all the elements into a single picture, thereby filling the identified gap is proposed.

2.4 Theoretical and Technical Frameworks

This study is theoretically based on Design Science Research, a framework dedicated to the development and assessment of novel technical artifacts aimed at addressing identified organizational issues. DSR offers the structured framework essential for the design, development, and critical evaluation of the proposed data pipeline and user interface [17].

The study will use Natural Language Processing text mining frameworks, specifically leveraging transformer based tokenization architectures to systematically extract structural carbon data from unstandardized vendor papers. This is coupled with an algorithmic anomaly detection framework, which quantitatively defines operating baselines and autonomously identifies reporting entries that diverge from previous compliance trends.

2.5 Relation to Study

This work makes a unique link between operational resource logistics and real time corporate compliance automation. Compared to generic business intelligence tools used in standard industrial implementations for the summarization of static data, this study proposes a novel real time data planning architecture tailored for large scale waste management networks. What's unique about this study is the development of automated data verification pipelines that transform raw physical metrics, fill levels, and contamination rates directly to standardized, auditable regulatory reporting formats. This shifts the industrial case study from manual, reactive auditing to automated, continuous compliance validation.

2.6 Literature Summary

According to the literature, global environmental regulations are expanding, which requires a fundamental transformation of corporate compliance processes. Real time IoT sensors and dedicated machine learning algorithms are very effective in optimizing discrete operational components, but current systems are hampered by data fragmentation and manual reporting blockages. There is no coherent architectural framework in the existing knowledge base that can ingest real time physical telemetry into a verifiable corporate dashboard. This research proposal tackles this structural limitation head on by developing and validating an integrated AI driven data planning platform.

3. Research Design and Methodology

3.1 Research Design

3.1.1 Research Approach

This study adopts the Design Science Research paradigm, which focuses on the development and evaluation of innovative artifacts designed to solve specific, accepted organizational problems [17]. Given that the core research challenge involves addressing structural data fragmentation and manual reporting inefficiencies at SUEZ [1], the creation of a technical artifact is required. The DSR framework provides a systematic, iterative methodology to engineer two distinct artifacts:

- The underlying data on the system architecture
- The customer dashboard interface

By focusing on artifact construction, the study shifts from traditional descriptive research toward an applied engineering approach that generates actionable technical solutions for the resource management sector.

3.1.2 Justification of the Selected Design

Alternative methodological paradigms such as theoretical behavioural analysis or simply quantitative surveys are not appropriate for this research issue since they cannot deal with system fragmentation or technical operational inefficiencies. SUEZ needs a structural framework to combine data silos across its activities and to automate compliance. The design based technique is very valuable, since it enables the researcher to construct, test, and revise architectural configurations in a controlled case scenario. This observed validation ensures the evaluation of the designed artifact in terms of operational constraints, regulatory disclosure requirements, and industrial performance measures at a specific level.

3.2 Research Methods and Materials

3.2.1 Overview of Research Methods

The methodology utilized for this project is a software engineering and case evaluation methodology, with hybrid approaches. The engineering phase is quantitative, dealing with data integration, speed of script processing, accuracy of algorithms for text extraction, and anomaly detection approaches. The assessment stage is a qualitative compliance review to confirm that the generated reporting outputs are in strict conformance with the legal constraints of the EU Taxonomy and UK Net Zero criteria. This dual technique aims to ensure that the prototype is not only technically sound but also legally valid.

3.2.2 Research Strategy

The research approach takes the form of an intense single case study of SUEZ's UK business. SUEZ offers a perfect study setting, being big scale about 10 million tonnes of garbage collected each year over hundreds of sites. It also has an established Internet of things IoT infrastructure and clear digital management methods, making it a realistic platform to test new procedures. The case study method makes it possible to build the system using real industrial data formats, which allows the results to be realistically applied to large scale infrastructure environments.

3.2.3 Proposed System Architecture and Dashboard Design

The proposed system architecture is designed as a data pipeline that relies into the Customer Waste Intelligence Dashboard:

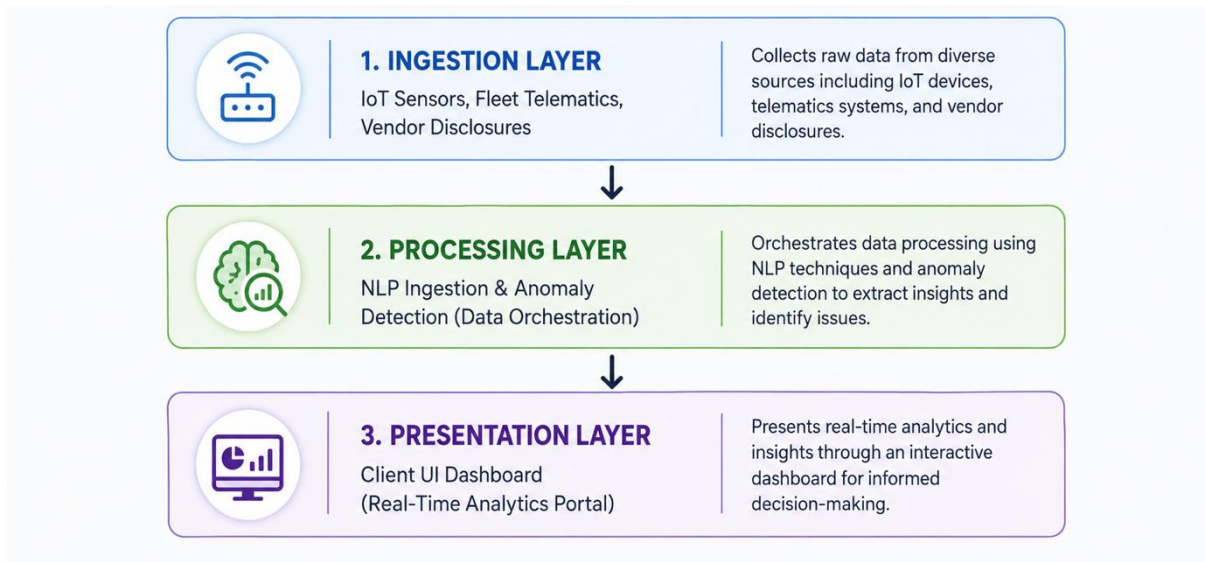


Figure 2 Dashboard Architecture and Design

- **The Ingestion Layer:** Automatically captures raw data streams from IoT smart bin sensors, fleet telematics units, and unstandardised supplier disclosure documents.
- **The Processing Layer:** Utilizes Natural Language Processing (NLP) routines to extract structural carbon metrics from supplier texts and runs anomaly detection algorithms to flag reporting errors or inconsistent entries.
- **The Presentation Layer:** Feeds verified data directly into the client user interface (UI), automatically generating real time summaries of recycling rates, contamination levels, and carbon footprints to satisfy statutory disclosure demands.

3.2.4 Data Analysis and Evaluation Metrics

To validate the technical efficacy of the engineered platform, the prototype will be evaluated against concrete engineering performance metrics:

- **Material Recovery Optimisation:** Measuring if the integration of automated tracking assists sorting systems in achieving the projected 15% to 20% improvement in material recovery.
- **Contamination Reduction Rate:** Assessing the dashboard's ability to drive a targeted 15% reduction in source level batch contamination through real time telemetry alerts and client feedback loop mechanisms.
- **Data Pipeline Latency:** Evaluating the processing time required to ingest raw telemetry data, run the anomaly verification models, and update the regulatory reporting fields on the client interface.

3.3 Materials

3.3.1 Research Data

The materials for this study will be gathered through a dual source data collection strategy combining open source digital data with some corporate requests:

- **Publicly Available Online Data:** Baseline operational data will be compiled directly from SUEZ's official digital publications, including the UK Environmental Performance Reports, published supply chain sustainability statements, and online case studies detailing their existing Smart Water and IoT sensor implementations.
- **Regulatory Open Data:** Unstructured compliance text profiles will be collected from public reporting databases, such as the UK Government's carbon disclosure registry and the EU Corporate Sustainability Reporting Directive open portals.
- **Additional Corporate Request:** A formal academic request will be submitted to SUEZ to secure anonymised interviews and historical supply network emissions data to ground the prototype in real world data scenario.

3.3.2 Justification of Materials

The use of genuine company interviews and documented internet case studies to construct an engineering data model that matches the real regulatory issues of the environmental services sector.

In conjunction with this public infrastructure, a guided inquiry for internal SUEZ data allows testing the AI pipeline against real data fragmentation patterns and realistic supply chain data gaps. The hybrid method is a combination of observed accessibility and industrial authenticity.

3.3.3 Ethical Considerations Related to Data

Because this plan involves public documentation and targeted corporate data requests, clear data ethics protocols will be strictly maintained:

- **Data Minimisation:** The formal request to SUEZ will strictly exclude any exclusive financial projections, customer personal data, or contract specifics.
- **Anonymisation:** Any profiles, individual facility logs, or geographic data provided by SUEZ or found online will be anonymised.
- **Institutional Alignment:** All data combination workflows will be submitted to the University Research Ethics Committee for formal approval before any automated scripts or corporate request letters are issued.

3.3.4 Software Requirements

The prototype architecture will use a software ecosystem to execute data reprocessing and knowledge discovery within the compressed timeframe:

- **Backend Environment:** Python and Jupyter Notebook manage data cleaning, handling missing values, and executing Scikit learn machine learning algorithms.
- **Data Engineering:** Structured Query Language handles database querying, while Pandas and NumPy execute feature engineering and data transformation routines.

- **Dashboard Interface:** Java manages the application logic, while native HTML, CSS, and Google Colab to compile the interactive visualizations and charts for the presentation layer.

3.3.5 Hardware Requirements

- **System Architecture:** Dell G15 laptop platform.
- **System Memory:** 8 gigabytes of RAM to handle data manipulation.
- **Processing Unit:** Central processing unit to run data reprocessing scripts and database queries.
- **Graphics Architecture:** RTX 3060Ti Nvidia Graphics Card

4. Expected Outcome

4.1 Expected Results

The key constructed product of this project, i.e., the data arrangement pipeline and interactive dashboard prototype, is expected to demonstrate the technical viability of automating complex environmental, social, and governance reporting activities. The platform will replace manual, disjointed procedures in spreadsheets with automated machine learning workflows and should deliver considerable improvements to the efficiency of raw material tracking and reporting. In particular, the natural language processing models and anomaly detection modules of the pipeline are predicted to exhibit excellent accuracy in extracting structural carbon indicators from unstandardized vendor texts and pass through data arrays. Furthermore, the use of telemetry for real time tracking will confirm the ability of the presentation layer to automatically identify reporting errors and so provide a robust system design for ongoing auditing of corporate compliance.

4.2 Implications

The theoretical and practical implications of this study span both the academic data engineering domain and the broader environmental services industry:

- **Academic Frameworks:** This research addresses a obvious gap in the design based computer science literature, by outdoing individual machine learning measures, to provide a whole data organization blueprint for large scale circular infrastructures.
- **Operational Decoupling:** In real terms, effective validation of the dashboard prototype will show a working mechanism for fully decoupling corporate profitability from physical waste volumes. The design enables a repeatable path for a traditional volume reliant waste operator to become an essential digital infrastructure partner, monetizing difficult raw indicators as strategic, high profit sustainability intelligence.
- **Regulatory Readiness:** The system provides an auditable compliance engine for corporate management to enable reporting against changing mandates

such as the EU Taxonomy and the Corporate Sustainability Reporting Directive, helping companies navigate evolving legal standards with less risk and cost.

4.3 Limitations

While the proposed system architecture targets high technical performance, several operational and structural constraints must be recognized:

- **Data Integrity Dependencies:** The automated dashboard is expressly reliant on the quality of the entering data, as high rates of missing or corrupted vendor files would unavoidably lower the precision of the underlying text mining models.
- **Contextual Scaling Boundaries:** The machine learning pipelines are specifically calibrated to SUEZ's operational profiles and regulatory structures in the UK and EU. Therefore, the transfer of the prototype to other geographical regions with completely different regulations for environmental reporting or data formats will require considerable re training and system reconfiguration.
- **Computational Limits:** Running continuous natural language processing and transformer based tokenization pipelines across complex multinational supply chains brings with it clear processing and latency challenges that may require specialized high performance hardware scaling for industrial deployment.

5. Timeline

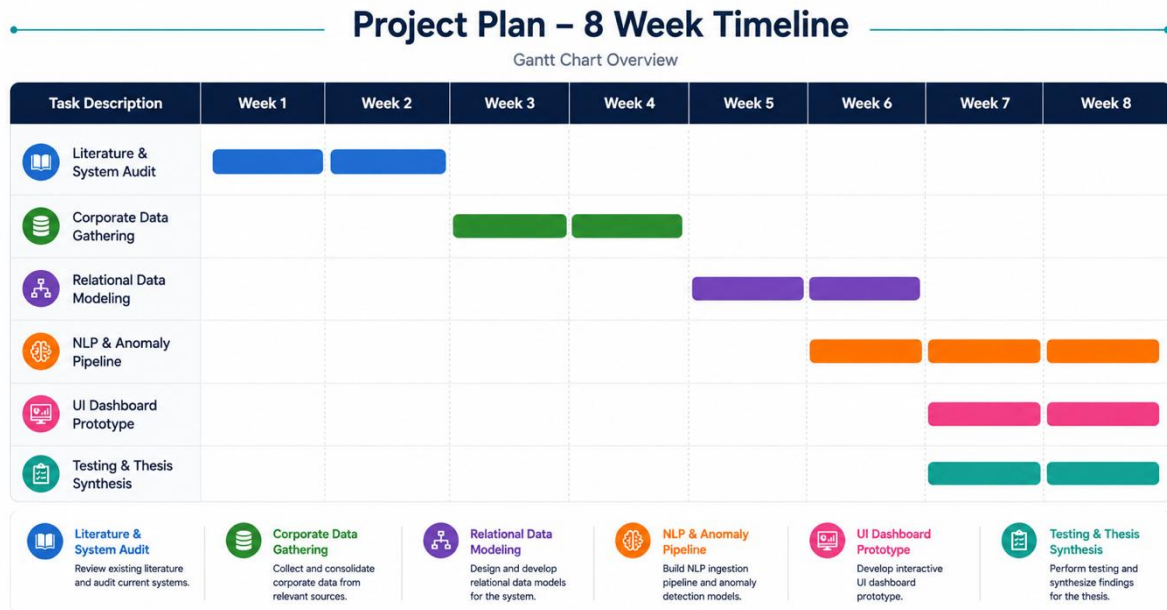


Figure 3 Gantt Chart

The research project will be executed on a compressed 8-week time frame, separated into overlapping functional phases to ensure fast artifact delivery. Weeks 1 and 2 will be used to start a literature review and conduct a system audit to analyze SUEZ software configurations. Weeks three and four are dedicated to the gaining of corporate data and the harvesting of essential open data from public registers and supplier channels. Relational data modelling design efforts transform these collected data sources into structured plans in weeks five and six.

During week six, system engineering scales up dramatically, with the natural language processing and irregularity pipeline operations going through to week eight building out the core processing backend. Weeks seven and eight: front-end interface building overlaps to assemble the interactive user interface dashboard prototype for final presentation. During weeks seven and eight, the sprint involves extensive testing of the platform against target metrics, as well as the synthesis of engineering findings for the final thesis submission.

REFERENCES

- [1] S. Group, "Presentation of the SUEZ Group," SUEZ SA, nd nd 2025. [Online]. Available: <https://www.suez.com/en/group/about-us>. [Accessed 27 06 2026].
- [2] European Commission, "EU taxonomy for sustainable activities," Directorate General for Financial Stability, Financial Services and Capital Markets Union, nd nd 2024. [Online]. Available: https://finance.ec.europa.eu/sustainable-finance/tools-and-standards/eu-taxonomy-sustainable-activities_en. [Accessed 01 07 2026].
- [3] International Organization for Standardization, "ISO 20400:2017 Sustainable procurement," ISO Central Secretariat, nd 04 2017. [Online]. Available: <https://www.iso.org/standard/63026.html>. [Accessed 3 07 2026].
- [4] Avanade Group, "SUEZ, AI & IoT Case Study," Avanade Inc., nd nd 2025. [Online]. Available: <https://www.avanade.com/en/insights/clients/suez-ai-iot>. [Accessed 09 07 2026].
- [5] SUEZ Group, "2025 SUEZ Sustainability Statement," SUEZ SA, 03 06 2026. [Online]. Available: <https://www.suez.com/-/media/suez-global/files/publication-docs/pdf-english/suez-csr-d-sustainability-statement-2025-en.pdf>. [Accessed 09 07 2026].
- [6] McKinsey & Company, "Buying into a more sustainable value chain," McKinsey & Company Operations Practice, 24 09 2024. [Online]. Available: <https://www.mckinsey.com/capabilities/operations/our-insights/buying-into-a-more-sustainable-value-chain>. [Accessed 09 07 2026].
- [7] IBM Institute for Business Value, "Beyond checking a box: How to create business value with sustainability," IBM Corporation, 14 04 2024. [Online]. Available: <https://www.ibm.com/insights/sustainability/esg-data-challenges>. [Accessed 09 07 2026].
- [8] Deloitte Group, "European Omnibus Proposal," Deloitte Ireland LLP, nd 02 2025. [Online]. Available: <https://mkto.deloitte.com/rs/712-CNF-326/images/EU%20Omnibus%20Proposals%20-%20What%20it%20Means%20for%20Sustainability%20Reporting.pdf?version=0>. [Accessed 09 07 2026].
- [9] KPMG International, "New amendments simplify EU Taxonomy," KPMG International Sustainability Practice, nd nd 2025. [Online]. Available: <https://kpmg.com/xx/en/our-insights/ifrg/2025/EU-taxonomy-amendments.html>. [Accessed 09 07 2026].
- [10] Arthur Cox LLP, "EU Taxonomy: Simplification measures adopted," Arthur Cox Limited, nd nd 2025. [Online]. Available: <https://www.arthurcox.com/insights/eu-taxonomy-simplification-measures-adopted/>. [Accessed 09 07 2026].
- [11] World Economic Forum, "Harnessing the Fourth Industrial Revolution for the Circular Economy," World Economic Forum Centre for the Fourth Industrial Revolution, nd nd 2024. [Online]. Available: <https://www.weforum.org/publications/harnessing-the-fourth-industrial-revolution-for-the-circular-economy/>. [Accessed 09 07 2026].
- [12] Towards AI, "The Data Engineering: How Modern Data Infrastructure is Powering the AI Revolution," Towards AI Media, 18 04 2026. [Online]. Available: <https://pub.towardsai.net/the-data-engineering-how-modern-data-infrastructure-is-powering-the-ai-revolution-279d1af04635>. [Accessed 10 07 2026].

- [13] RAC Drive, "Fleets could cut fuel costs by up to 15% through smarter driver data," RAC Motoring Services, 25 06 2026. [Online]. Available: <https://www.rac.co.uk/drive/news/motoring-news/fleets-could-cut-fuel-costs-through-smarter-driver-data/>. [Accessed 10 07 2026].
- [14] Dataintelo Group, "AI-Powered Recycling Robot Market Research Report 2033," Dataintelo Market Research, nd nd 2025. [Online]. Available: <https://dataintelo.com/report/global-ai-powered-recycling-robot-market>. [Accessed 10 07 2026].
- [15] SAP SE, "What is CSRD (Corporate Sustainability Reporting Directive)?," SAP Global Regulatory Compliance and Sustainability Solutions, nd 02 2025. [Online]. Available: <https://www.sap.com/sea/resources/csrd-guide>. [Accessed 10 07 2026].
- [16] The Hong Kong Polytechnic University, "Secure environmental, social, and governance (ESG) data management framework using blockchain technology," PolyU Research Output Repository, nd nd 2024. [Online]. Available: <https://research.polyu.edu.hk/en/publications/secure-environmental-social-and-governance-esg-data-management-fo>. [Accessed 10 07 2026].
- [17] MDPI Group, "Digital Enablers of the Circular Economy: A Systematic Review of Applications, Barriers, and Future Directions," <https://www.mdpi.com/2504-4494/10/4/112>, nd nd 2025. [Online]. Available: <https://www.mdpi.com/2504-4494/10/4/112>. [Accessed 10 07 2026].

Appendices

Research Ethics Self-Assessment Form

Section	Response
Project Title	AI Driven ESG Data Arrangement: MSc Research Proposal for Automated Compliance Monitoring and Circular Intelligence within SUEZ Operations
Research Supervisor	Mohammed Ben Farah
Human Participation (Yes/ No)	Yes
Recruitment Method	University Email Contact
Consent Process	Custom Consent Form and written informed consent forms provided digitally and signed by participants prior to scheduling any interview sessions.
Right to Withdraw	Yes, without any problems.
Personal Data Collected? (Yes/ No)	No
Data Storage	University outlook onedrive storage
Datasets Used for Modelling	Public SUEZ UK environmental reports, public carbon disclosure databases, EU corporate sustainability portals, and anonymized interview.
Foreseeable Risks to Participants	Minimal to None
Risk Mitigation	Enforcing strict data minimization protocols alongside complete identity masking and total anonymization of text files before processing
Researcher Declaration	I declare that this project will be conducted in strict alignment with institutional ethical standards, upholding total participant confidentiality