Work Placement Portfolio

Data Integration and BI Automation Intern

Breda University of Applied Sciences

Vladislav Stoimenov (235030)

19 January 2026

Work Placement Portfolio

1. Introduction

This portfolio documents my 20-week work placement conducted as part of the Applied Data Science & Artificial Intelligence programme at Breda University of Applied Sciences. The placement took place at ‘Move Intermodal’, where I worked within the data and analytics domain on project related to data integration, data modelling, and analytical reporting for logistics operations. The primary objective of this portfolio is to provide a coherent and professional account of my activities, learning outcomes, and contributions during the placement period, demonstrating both technical competence and reflective professional development.

During the placement, I was assigned a role that combines elements of data engineering and data analytics, with a strong focus on building reliable data foundations to support business intelligence and decision-making. The central project revolved around the integration of operational logistics data from external source system into a structured Data Vault 2.0 architecture hosted on a cloud-based data platform. This involved translating raw transactional data into scalable, well-governed data models and making this data accessible through analytical views for reporting and further analysis.

The work was situated in a real-world business context, where technical decisions were influenced by organizational constraints such as data quality, legacy system structures, stakeholder requirements, and time limitations. Throughout the placement, I collaborated closely with colleagues from data engineering, analytics, and business teams, ensuring that technical implementations aligned with business needs.

This portfolio aims to demonstrate how academic knowledge from the Applied Data Science & Artificial Intelligence programme was applied and adapted in practice. Relevant theories, frameworks, and best practices are connected to the choices made during the project, and critical reflections are provided where discrepancies between theory and practice emerged. Ethical and legal considerations, such as data governance and privacy, are also addressed to illustrate responsible data handling within an organisational context.

The report is structured to guide the reader from an overview of the company and project context, through theoretical foundations and technical execution, to an evaluation outcomes and a reflection on personal development. As such, the portfolio is intended to be understandable to both academic assessors and professionals, without requiring additional explanation beyond what is presented in this document.

2. Project Description and Timeline/Plan

The work placement project was initiated following the acquisition of a new business unit, Van Dijk Logistics, by Move Intermodal. After the acquisition, reporting for Van Dijk Logistics was still largely performed manually. Data was extracted from operational systems and processed in spreadsheets to produce management reports. This approach required significant manual effort, increased the risk of errors, and made it difficult to scale reporting as the business grew. The aim of the placement project was to support this integration by creating an automated and structured reporting foundation.

The main objective of the project was to design and implement a Data Vault 2.0 data model that integrates data from Van Dijk Logistics into the existing data platform of Move Intermodal. This model was intended to replace manual reporting processes with an automated solution that supports consistent reporting, historical tracking of data, and long-term scalability. By doing so, the project contributed directly to the integration of the newly acquired business unit into the wider data landscape of Move Intermodal.

2.1 Initial Planning and Intended Scope

At the start of the placement, a structured plan was created to guide the internship over the 20-week period. The project was divided into several phases. The first phase focused on understanding the business context and data landscape of Van Dijk Logistics, including logistics processes, source systems, and existing reports. The second phase focused on designing and implementing the Data Vault 2.0 model, including the definition of business keys, relationships, and descriptive attributes.

Following the implementation of the core data model, the next planned phase was to rebuild the existing manual reports in an automated way using the new data foundation. After this, a separate phase was planned for mapping the data of Van Dijk Logistics to the existing Move Intermodal data structures, in order to align reporting across business units. The final phase of the project was reserved for validation, documentation, and handover.

The planning included a high-level timeline that reflected this phased approach. While clear milestones were defined, the plan was intentionally kept flexible. Given the complexity of working with a newly acquired business unit and unfamiliar data sources, it was expected that adjustments might be required during the placement.

2.2 Progress Up to the Halfway Assessment

During the first half of the placement, up to approximately week 10, the project progressed according to the original plan. In several areas, the work was even ahead of schedule. Significant progress was made in analysing the source data of Van Dijk Logistics and understanding how it could be integrated into the existing data platform. The Data Vault 2.0 model was designed and implemented as planned, including the creation of the core hubs, links, and satellites.

Regular meetings with the company supervisor ensured alignment and provided timely feedback. Design decisions were validated early, which reduced uncertainty and limited the need for rework. The halfway assessment confirmed that the project objectives were being met and that progress was in line with expectations. At the same time, the assessment highlighted the importance of stronger stakeholder involvement, both from mine and the Van Dijk Logistics side, in the upcoming phases of the project. This feedback was given in anticipation of the reporting phase, where detailed business knowledge and validation would be essential. Following this assessment, the focus shifted towards rebuilding the existing manual reports using the newly implemented data model.

2.3 Challenges During the Report Rebuilding Phase

The phase focused on rebuilding reports proved to be more challenging than expected. Recreating the existing reports required detailed knowledge of business logic, calculations, and exceptions that were not fully documented. This knowledge was mainly held by employees at the Van Dijk Logistics office.

During this period, the availability of relevant stakeholders from Van Dijk Logistics was limited due to operational priorities. As a result, it was difficult to obtain timely feedback and clarification on reporting requirements. Despite multiple attempts to request support and validation, progress in rebuilding the reports remained limited for a period of time.

Although I continued working independently by analysing existing reports and testing possible approaches, the lack of clear direction meant that this work involved a high degree of uncertainty and did not lead to substantial progress in finalising the reports.

2.4 Adjustment of the Project Plan: Data Mapping First

To avoid remaining blocked in the reporting phase, the project plan was adjusted. Instead of continuing to work on report rebuilding without sufficient input, the focus temporarily shifted to the data mapping phase, which was originally planned for a later stage of the project. This phase involved mapping the data of Van Dijk Logistics to the existing Move Intermodal data structures in order to ensure consistency across business units.

By moving this phase forward, progress could be made independently of immediate stakeholder availability. The data mapping work helped to clarify how the data from both organisations related to each other and revealed differences in definitions, structures, and reporting logic. This work also strengthened the overall data foundation and reduced uncertainty, which later supported the reporting phase.

This adjustment did not change the overall objectives of the project, but it did change the order in which certain tasks were executed. Prioritising data mapping first proved to be a practical decision, as it reduced dependency on unavailable stakeholders and ensured that the reporting logic could later be built on a clearer and more consistent data structure.

2.5 Final Phase and Planning Reflection

In the final phase of the placement (weeks 17–20), the focus was on consolidating the completed work. This included refining the data mappings, improving documentation, and preparing the project for handover. Although the rebuilding of reports took longer than originally planned, the strengthened data foundation provided a solid basis for completing this work and for future extensions by the organisation.

Overall, the project followed the original plan closely during the first half of the placement, with progress meeting or exceeding expectations. The challenges encountered during the reporting phase highlighted the importance of stakeholder availability and domain knowledge when automating reporting for a newly acquired business unit. By adjusting the plan and moving the data mapping phase forward, the project was able to continue delivering value while remaining aligned with its core objectives. This experience demonstrates the need for flexibility and proactive decision-making in real-world data integration projects.

3. Company Description

Move Intermodal is a European logistics company that specialises in intermodal transport solutions, with a strong focus on full truckload (FTL) and full container load (FCL) shipments. The company has more than 35 years of experience in the logistics sector and operates across major European corridors. Move Intermodal combines road, rail, and sea transport to offer efficient, reliable, and more sustainable logistics solutions for its customers.

The company was founded in 1986 and started as part of Ewals Cargo Care, where it became one of the early pioneers in intermodal transport in the Benelux region. Over the years, Move Intermodal continued to grow by developing its own intermodal concepts and expanding its network. A major milestone was reached in 1999, when Move Intermodal became the first Belgian transport company to operate its own rail shuttle between Genk (Belgium) and Novara (Italy). Since 2014, following a management buy-out initiated by the current CEO, Luc Driessen, the company has operated under the name Move Intermodal. Today, the company continues to grow by expanding its network to key economic regions across Europe, including Spain, the United Kingdom, Scandinavia, and Turkey.

Move Intermodal's vision is to provide efficient, cost-effective, and sustainable transport solutions that integrate seamlessly into customers' supply chains. The company aims to reduce road congestion and environmental impact by shifting freight from road to rail wherever possible. Sustainability, safety, and reliability are central values within the organisation, supported by a strong belief that people play a key role in delivering high-quality service.

Currently, Move Intermodal employs around 300 people across 9 countries. The company manages approximately 2,800 load units, operates 170 own trucks, and runs around 60 company shuttle trains per week. With priority access to the European rail shuttle network, Move Intermodal organises more than 1,000 transport movements per day across Europe. The company is privately owned and has an annual turnover of over €100 million, with its main offices located in Genk (Belgium) and Novara (Italy).

As part of its growth strategy, Move Intermodal acquired Van Dijk Logistics, a company owned by the Move family that operates as a specialised business unit within the group. Van Dijk Logistics focuses on less-than-truckload (LTL) transport, meaning it combines smaller shipments from multiple customers into shared transports. The company mainly operates along key corridors between Belgium, the Netherlands, and Italy. LTL transport requires more complex planning, consolidation, and coordination compared to FTL, making accurate data and clear reporting especially important.

Together, Move Intermodal and Van Dijk Logistics allow the group to serve a wide range of customer needs, from large, direct shipments to smaller, consolidated goods. This combination strengthens the company's position in the European logistics market and increases the need for aligned processes, data, and reporting across business units. The project carried out during this placement directly supported this integration by working towards a unified and automated reporting foundation.

From an organisational perspective, Move Intermodal consists of several departments, including operations, planning, finance, sales, and data and IT-related functions. During the placement, I worked within the data and IT department, collaborating mainly with stakeholders involved in reporting and business engineering.

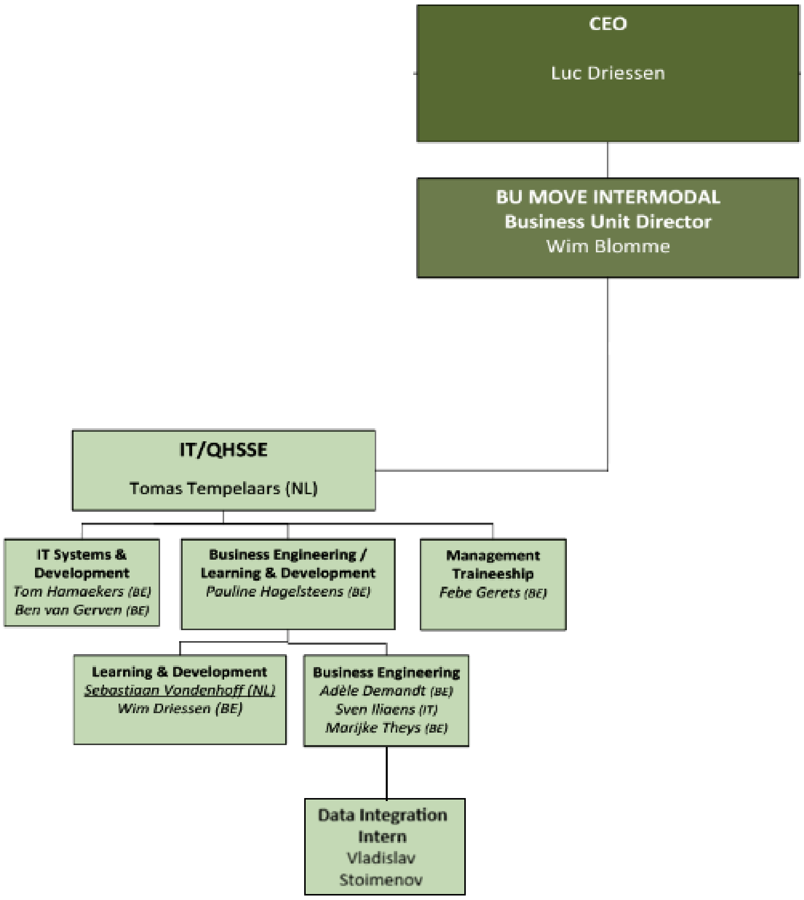

Figure 1. Organizational Chart (IT Department)

The work culture at Move Intermodal can be described as professional, open, and responsibility-driven. Employees are given a high level of ownership over their tasks, and communication is generally direct and informal. While the company does not strictly follow formal frameworks such as SCRUM or Kanban, work is organised through regular meetings and weekly check-ins. This flexible way of working allows teams to adapt quickly to operational priorities, which is common in the logistics sector.

Understanding this organisational, operational, and cultural context was essential for the placement project. It helped shape realistic expectations, influenced planning decisions, and provided insight into how data solutions can effectively support business processes within a complex logistics environment.

4. Theoretical Background and Literature Study

The work carried out during the placement was informed by established concepts from data engineering, data modelling, and business intelligence. In order to justify the design choices made during the project, it is important to relate the practical work to existing theory and industry best practices. This section discusses the main frameworks and concepts that guided the project and reflects on how they were applied in a real-world organisational context.

A central theoretical foundation for the project was the Data Vault 2.0 modelling methodology. Data Vault was developed as an approach for building scalable, flexible, and auditable data warehouses that can adapt to changing business requirements (Linstedt & Olschimke, 2015). The methodology structures data into three main components: hubs, which store business keys; links, which represent relationships between business entities; and satellites, which contain descriptive attributes and historical information. This structure supports clear data lineage and makes it possible to track changes over time.

The use of Data Vault 2.0 is particularly well suited for situations where data from multiple source systems needs to be integrated, or where new systems are added to an existing data landscape. This was directly relevant in the context of the project, as data from Van Dijk Logistics needed to be integrated into the existing reporting environment of Move Intermodal. From a theoretical perspective, Data Vault supports this type of integration by separating stable business concepts from changing descriptive data, which reduces the impact of future changes to source systems (Linstedt & Olschimke, 2015).

In practice, applying the Data Vault methodology required careful interpretation of business keys and relationships. Although the theory provides clear guidelines, business definitions were not always explicitly documented in the source system. As a result, close collaboration with stakeholders and iterative refinement were necessary to ensure that the model correctly reflected the business reality. This highlights a common gap between theory and practice, where frameworks offer structure but still rely on domain knowledge to be applied effectively.

In addition to data modelling theory, the project was guided by best practices related to data integration and automation. Industry literature emphasises the importance of replacing manual data handling with automated and reproducible data pipelines in order to improve reliability and reduce operational risk (Stonebraker et al., 2007). These principles informed the decision to focus on structured data transformations, consistent data mappings, and automated processes rather than manual report preparation. Automation was a key requirement of the project, as the goal was to eliminate manual reporting steps and ensure that reports are updated automatically.

The overall structure of the project can also be related to general frameworks used in data and analytics projects, such as CRISP-DM (Chapman et al., 2000). Although CRISP-DM was originally developed for data mining, its phases provide a useful way to describe the progression of the placement. The project started with business and data understanding, followed by data preparation and modelling through the implementation of the Data Vault model, and continued with evaluation through validation of reporting outputs. In practice, these phases were not followed in a strict sequence but often overlapped, reflecting the iterative nature of real-world projects.

Taken together, the literature provided a strong foundation for the decisions made during the placement, particularly with regard to data modelling and automation. At the same time, the project demonstrated that theoretical frameworks must often be adapted to fit organisational constraints, such as stakeholder availability, and existing system limitations. This reinforces the idea that successful data integration projects require not only sound technical frameworks, but also flexibility, communication, and practical judgement when applying theory in a real-world environment.

5. Data Description

The data used during the placement project originated from internal operational systems of Van Dijk Logistics and represented historical and ongoing logistics activities. The dataset covered a long time period, with records dating back to 2005, and reflected the operational growth of the company over time.

The database consisted of approximately 60 relational tables, each serving a specific purpose within the logistics domain. These tables stored information related to customer orders, transport executions, shipment parts, costs, train operations, documents, clients, and transport equipment. Together, they provided a comprehensive view of intermodal logistics processes from order creation to execution and completion.

Across the database, the tables contained structured attributes such as identifiers, dates, locations, routes, addresses, goods information, client-related data, equipment details, and financial values. The level of detail varied between tables, with some containing high-level reference information and others storing detailed transactional records.

The dataset was also characterised by its large volume. The orders table (Dossier) contained more than 330,000 records, while the shipment-related table (Vracht) contained over 960,000 records, reflecting the fact that single orders can be split into multiple shipment parts. This highlights that the project involved working with a large-scale operational dataset rather than a limited sample.

Overall, the data represented a realistic and complex logistics environment. Its size, historical depth, and variety of information required careful consideration when designing a solution for automated reporting, which is addressed in the subsequent sections of this portfolio.

6. Methodology and Implementation

This section describes in detail how the project was executed during the placement period. It explains the tools, methods, and processes that were used to move from manual reporting to an automated reporting solution. The work was carried out in several phases, which followed a logical order but were adapted when needed to deal with practical constraints such as data complexity and stakeholder availability.

6.1 Initial Data Exploration Using Transpas

During the first two weeks of the placement, the main focus was on understanding the source data. For this purpose, Transpas, the transport management system (TMS) used by Van Dijk Logistics, was explored in detail. Transpas contains operational data related to orders, transport actions, shipments, and costs.

At this stage, the goal was not yet to build reports, but to understand how the system works and how data is entered in practice. This included analysing the structure of the database, understanding the relationships between tables, and reviewing example cases where orders were created in the system. By following how an order is entered step by step, it became clearer which data fields are mandatory, which are optional, and how business processes are reflected in the data. This initial understanding formed the foundation for all later implementation steps.

6.2 Snowflake as Central Data Warehouse

Snowflake was used as the central data warehouse for both Move Intermodal and Van Dijk Logistics. All raw data, transformed data, and analytical views were stored in Snowflake. Using Snowflake allowed the project to separate data storage from reporting and ensured that all data was available in a single, consistent environment.

SQL was used throughout the project to query the data, develop the data model, validate results, and perform data transformations. SQL was essential not only for exploring the data, but also for building views, applying filters, and supporting the data mapping and reporting phases.

6.3 Data Vault 2.0 Implementation Using Matillion

The core of the project was the creation of a Data Vault 2.0 model, which was implemented using Matillion. It was used as the ETL tool to connect Snowflake tables, apply transformations, and load data into the Data Vault structure.

The first step in Matillion was setting up the staging layer, where raw tables from Snowflake were connected and made available for further processing. After this, the tables were explored in more detail to understand their features, the business concepts they represent, and the values they contain. Special attention was given to identifying business keys, as these are essential for Data Vault modelling.

Based on this analysis, the Data Vault components were created:

- Hubs, which store unique business keys (for example, orders or shipments)
- Links, which represent relationships between business entities
- Satellites, which store descriptive attributes and historical data
- Link satellites (LSATs), which store descriptive information about relationships

Using these components, multiple Data Vault views were created. These views combine data from different hubs, links, and satellites through their business keys. The views act as structured tables that bring together all relevant information needed for reporting.

During this phase, it became clear that three tables were central to the data model:

- Dossier, which contains information about orders
- Rit, which represents transport actions (each shipment can consist of multiple transport actions such as truck–rail–truck)
- Vracht, which represents the concrete part of an order within a transport action

Because orders can be split across multiple transport actions, and transport actions can include parts of multiple orders, the Vracht level was chosen as the lowest and most detailed level of analysis. Around 60 tables were available in the database, which required extensive exploration at the beginning. By analysing the existing manual reports of Van Dijk, only the relevant tables were selected, while outdated or unused tables were excluded.

Using the selected tables, more than 10 Data Vault views were created and linked together, with Vracht-based views serving as the main join level.

6.4 Incremental Logic and Historical Tracking

After the Data Vault views were created, incremental logic was implemented to support historical tracking. The purpose of this logic was to ensure that changes in the data are preserved over time.

This was achieved by adding LOAD_DATE and LOAD_END_DATE fields to satellites and links. When a new record with the same business key appears, the previous record is closed by setting its LOAD_END_DATE, and the new record is marked as the current one. This approach allows the system to keep a full history of changes.

To ensure that reports always show the most recent data, filters were applied in the views so that only records with the active LOAD_DATE were selected. In this way, historical data is preserved in the database, while reports remain up to date.

6.5 Creation of Data Mart Views

Once the Data Vault views were stable, Data Mart views were created. Data Mart views are simplified views that are specifically designed for reporting. They use the Data Vault views as input and apply additional filters, calculations, and selections based on reporting needs.

For each report or dashboard view, a separate Data Mart view was created containing only the required data. This made reporting easier, improved performance, and reduced complexity in the BI layer.

6.6 Data Mapping Between Van Dijk and Move Intermodal

During the reporting phase, progress was temporarily blocked due to limited stakeholder availability. To avoid losing time, the focus shifted to data mapping, which could be done more independently.

The data mapping started by analysing the existing data model of Move Intermodal. An Excel file containing sample data and column descriptions was used to map Van Dijk columns to their corresponding Move columns. It became clear that Move's data model was based on order level, while Van Dijk's final view was at Vracht level.

To resolve this, a separate view was created to make orders (Dossier) unique before mapping. The mapping process started with simple column renaming, followed by more complex transformations where values needed to be adjusted. Additional calculated columns were created, including financial metrics such as margins and percentages, as well as time-based metrics such as loading and unloading differences.

An important challenge during this phase was handling duplicate order numbers caused by cross-docking, where one order is split across multiple transports. This was resolved by aggregating records so that all information was preserved in a single order-level record. Data types were aligned, and the mapped views were connected to the existing Move data model. Schedulers were then set to update the data automatically.

6.7 Report Rebuilding and Validation in Bold BI

After completing the data mapping and once stakeholder availability improved, the focus returned to rebuilding the reports. Bold BI was used as the BI platform for Move Intermodal. Data sources were created in Bold BI by connecting directly to the Data Mart views in Snowflake.

The reports were rebuilt to match the existing manual reports as closely as possible. Through meetings with stakeholders from Van Dijk and the team responsible for the manual reports, detailed clarification was obtained on required calculations and filters. Based on this input, remaining Data Mart views were finalised.

The final step involved validating the dashboards by comparing results with the manual reports and reviewing them together with stakeholders. This ensured that the automated reports were correct, complete, and ready for operational use.

7. Evaluation and Testing

This section evaluates the outcomes of the project and explains how the implemented solution was tested and validated. The main goal of the project was to replace manual reporting for Van Dijk Logistics with an automated and reliable reporting solution that integrates with the existing data platform of Move Intermodal. Evaluation therefore focused on correctness, consistency, usability, and alignment with business expectations.

The first level of evaluation was data correctness. Outputs from the automated data views and dashboards were compared with the existing manual reports that were previously created using spreadsheets. Key figures such as shipment counts, costs, and operational metrics were checked in detail. Differences between the automated results and the manual reports were analysed to understand whether they were caused by data quality issues, differences in business logic, or limitations in the original reporting process. This comparison helped identify and resolve inconsistencies and increased confidence in the new solution.

The second level of evaluation focused on logic and transformation validation. SQL queries and transformation logic were tested by applying filters, checking edge cases, and validating intermediate results. Special attention was given to complex scenarios such as cross-docking, split shipments, and duplicate records. By testing these scenarios explicitly, it was ensured that the automated logic handled real-world situations correctly and consistently.

Stakeholder feedback played an important role in the evaluation process. Once initial versions of the automated reports were available, results were reviewed together with stakeholders from the Van Dijk office and the team responsible for the manual reports. During these sessions, discrepancies were discussed, assumptions were clarified, and adjustments were made where needed. This collaborative validation helped align the automated reports with business expectations and increased acceptance of the solution.

Another evaluation aspect was usability and maintainability. The use of Data Mart views ensured that reporting logic was separated from the underlying Data Vault structure, making the dashboards easier to understand and maintain. Automated data refreshes reduced manual effort and lowered the risk of human error. Compared to the previous manual reporting process, the new solution required significantly less time to update and was less dependent on individual knowledge.

As a result, the evaluation showed that the implemented solution met the main project objectives. While some limitations remained due to data quality and stakeholder availability, the automated reporting foundation provided a reliable and scalable basis for future development. The testing and validation process demonstrated that the solution was technically sound, aligned with business needs, and suitable for operational use within the organisation.

8. Ethical and Legal Considerations

During the placement, ethical and legal aspects were an important consideration, mainly because the project involved operational and financial data from logistics processes. Although the data did not include sensitive personal information, it still required careful handling to ensure compliance with company policies and legal regulations. The data used in the project was collected through existing operational systems and processed only for internal reporting purposes. No additional data was gathered outside the established business processes, and anonymization was therefore not required for the scope of this project.

One of the main legal considerations was data protection and privacy, particularly in relation to the General Data Protection Regulation (GDPR). The data used during the project mainly consisted of business-related information, such as shipment details, routes, and costs. Personal data was limited and, where present, was not exposed in reporting outputs. Access to systems such as Transpas, Snowflake, and reporting tools was restricted based on user roles, ensuring that only authorised people could view or modify the data. This approach supported responsible data governance in terms of data collection, storage, and access control.

From an ethical perspective, data accuracy and transparency were key concerns. Automated reports can strongly influence business decisions, so it was important to ensure that calculations, filters, and assumptions were correct and clearly documented. Where differences between automated and manual reports were identified, these were openly communicated to stakeholders instead of being adjusted without explanation. This helped maintain trust and avoided misleading conclusions.

Another ethical aspect related to fair and consistent reporting across business units. Since the project involved integrating data from a newly acquired company, care was taken to align definitions and metrics in a way that did not disadvantage one business unit compared to another. Differences in data structure and historical reporting practices were addressed through transparent mapping and documentation, reducing the risk of bias in reporting outcomes.

Finally, intellectual property and confidentiality were considered throughout the project. All work was performed within the company's systems, and sensitive business logic and data were not shared outside the organisation. Code, queries, and documentation included in the portfolio were selected carefully to ensure that no confidential information was disclosed.

With regard to formal ethical frameworks, the company does not apply specific AI or data ethics guidelines for reporting and data integration projects. Ethical considerations were therefore guided by general company policies, GDPR requirements, and professional judgement. Based on this experience, introducing more explicit internal guidelines on data governance and reporting transparency could further support consistent and responsible data use in future projects.

9. Discussion

The project demonstrated that manual reporting processes can be effectively replaced by an automated data solution when a well-structured data foundation is established. By integrating data from a newly acquired business unit and implementing a Data Vault 2.0 model, reporting became more consistent, reliable, and easier to maintain across the organisation.

A key challenge during the project was limited stakeholder availability, particularly during the reporting phase. This affected the timing of validation and required careful communication once alignment was possible. When stakeholder input became available, differences between automated and manual reports were analysed and discussed in detail. These discussions helped clarify business logic, identify inconsistencies in historical reporting, and improve the accuracy of the automated outputs.

The project also highlighted the relationship between theoretical frameworks and practical implementation. While Data Vault 2.0 provides a clear structure for handling complex and changing data, its successful application depends on domain knowledge and collaboration with business stakeholders. Data quality issues, legacy practices, and differences between business units required adaptations beyond what is described in theory.

In summary, the discussion shows that successful data integration and reporting projects require more than technical implementation alone. Clear communication, transparency in decision-making, and alignment with business understanding were essential factors in delivering a solution that is both technically sound and accepted by stakeholders.

10. Conclusion

This portfolio described the work carried out during the placement at Move Intermodal, with a focus on integrating data from the newly acquired business unit, Van Dijk Logistics, and transforming manual reporting processes into an automated and structured solution. The main objective of the project was to create a reliable data foundation that supports consistent and up-to-date reporting across business units.

Through the implementation of a Data Vault 2.0 model, the project provided a scalable and flexible structure for integrating complex logistics data. By using tools such as Transpas, Snowflake, Matillion, and Bold BI, operational data was transformed into automated reports that reduce manual effort and improve reliability. The project demonstrated that a well-designed data architecture can significantly improve reporting quality and transparency within a logistics organisation.

While the project achieved its main goals, some limitations remained. Data quality issues and differences in historical reporting practices required additional validation and stakeholder involvement. Limited availability of business stakeholders during certain phases slowed the reporting validation process. These challenges highlight the importance of early and continuous alignment between technical and business teams in data-driven projects.

Based on the outcomes of the placement, several recommendations can be made. First, continued refinement of the data model and reporting logic would further improve consistency and accuracy. Second, establishing clearer documentation and shared definitions for key metrics across business units could reduce ambiguity in future reporting. Finally, increasing stakeholder involvement earlier in the reporting design phase would support faster validation and stronger alignment with business needs.

Overall, the project delivered a solid foundation for automated reporting and demonstrated clear value for Move Intermodal. The implemented solution provides a strong basis for future development and supports more data-driven decision-making within the organisation.

11. Reflection on learning experience

The placement at Move Intermodal provided valuable learning experiences in both technical and professional areas. From a technical perspective, I significantly improved my skills in data modelling, SQL, and data integration. Working with Data Vault 2.0 in a real business context helped me understand not only the theory behind the methodology, but also the practical decisions required when dealing with complex and imperfect data.

One of the main challenges during the placement was working with limited stakeholder availability, especially during the reporting phase. This required me to become more independent in analysing data, validating assumptions, and documenting decisions. By adapting my approach and focusing on tasks that could be progressed independently, I was able to continue delivering value despite these constraints.

The placement also strengthened my communication skills. Explaining technical issues, discussing data differences, and aligning expectations with business stakeholders required clear and structured communication. I learned the importance of transparency and patience when working in a data-driven environment.

Based on this experience, my future learning goals include further developing my data engineering skills, particularly in scalable data architectures and automation, as well as improving my ability to translate business requirements into technical solutions. These goals build directly on the foundation developed during this placement.

References

[1] Data Vault 2.0 (Linstedt & Olschimke, 2015)

Linstedt, D., & Olschimke, M. (2015). Building a scalable data warehouse with Data Vault 2.0. Morgan Kaufmann. Available:https://www.google.nl/books/edition/Building_a_Scalable_Data_Warehouse_with/lgDJBAAAQBAJ?hl=en&gbpv=1&dq=Linstedt+%26+Olschimke,+2015&printsec=frontcover

[2] Conference Paper – Architecture Rewrite

Stonebraker, M., Madden, S., Abadi, D. J., Harizopoulos, S., Hachem, N., & Helland, P. (2007). The end of an architectural era (It's time for a complete rewrite). Proceedings of the 33rd International Conference on Very Large Data Bases. Retrieved from https://www.vldb.org/conf/2007/papers/industrial/p1150-stonebraker.pdf

[3] CRISP-DM (Chapman et al., 2000)

Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (2000). CRISP-DM 1.0: Step-by-step data mining guide (CRISP-WP-0800). SPSS Inc. Available: https://mineracaodedados.wordpress.com/wp-content/uploads/2012/12/crisp-dm-1-0.pdf

[4] Move Intermodal

Move Intermodal. (n.d.). *Move Intermodal – Intermodal logistics solutions across Europe*. Available: https://move-intermodal.com/

[5] ChatGPT

OpenAI. (2025). *ChatGPT* (GPT-5.2) [Large language model]. Available: https://chat.openai.com/