Data Engineering Project at ACA Group

Reflection

Bachelor's degree in ACS
Artificial Intelligence

Kyano Trevisan

Academic year 2024-2025

Campus Geel, Kleinhoefstraat 4, BE-2440 Geel

LID VAN ASSOCIATIE KU LEUVEN

THOMAS MORE

1 INTRODUCTION

This reflection document chronicles my transformative 13-week journey as a Data Engineer at ACA Group during the "Sustainathon" project - a mission-critical initiative to revolutionize Duvel Moortgat's sustainability reporting for CSRD compliance. From February 24th to May 28th, 2025, I experienced not just technical growth, but a complete transformation from student to professional consultant.

1.1 About ACA Group and the Datadots Team

ACA Group is a leading Belgian IT consulting firm with 253 employees, of which 152 are IT professionals. I joined their "Datadots" team - a collaborative group of data engineers, analysts, and consultants based across multiple offices in Hasselt, Antwerp, Leuven, and Gent. The team culture was exceptional; everyone was ready to answer questions and provide support, creating an environment where learning and growth flourished.

1.2 The Sustainathon Challenge

The Corporate Sustainability Reporting Directive (CSRD) is a European Union regulation requiring companies to report detailed ESG (Environmental, Social & Governance) information. For Duvel Moortgat, Belgium's premier brewery group with facilities across multiple countries, this meant transforming their manual, spreadsheet-based sustainability reporting into an automated, reliable system capable of handling complex data from 20+ global sites.

1.3 Building on Foundation

What made this project particularly interesting was that it built upon a proof of concept created the previous year by Kwinten Boes (now Project Manager/Data Analyst) and Thibo Vanderkam (Data Engineer), both former ACA interns. I joined alongside fellow intern Fabian Reyes, where Fabian focused on analysis and PowerBI dashboards while I concentrated on data engineering and pipeline development. This structure taught me the importance of specialization within collaborative teams.

This reflection explores both the technical achievements we delivered and my personal evolution from uncertain student to confident data engineering professional.

2 PART 1: SUBSTANTIVE PROJECT REFLECTION

2.1 The Starting Reality: A Complex Data Ecosystem

When I first encountered Duvel Moortgat's sustainability reporting challenge, I was struck by the sheer complexity. Data was scattered across dozens of Excel input sheets, multiple ERP systems (Salesforce, Workday), production databases (SCADA, WinCC), energy monitoring platforms, and manual input processes managed by stakeholders across the globe. Our primary contact, Dries from Duvel's sustainability team, had been maintaining this system through incredible manual effort - a process that was becoming unsustainable as CSRD requirements evolved.

The existing proof of concept had demonstrated potential, but it required significant engineering work to transform it from a working prototype into a production-ready system that could handle Duvel's global operations reliably.

2.2 Technical Foundation: Learning While Building

2.2.1 Weeks 1-4: Mastering New Technologies

My journey began with intensive learning of technologies I'd never used: Microsoft Fabric and dbt (data build tool). What struck me about Microsoft Fabric was that it's still actively being developed - many features were in preview, and I learned to work with its quirks, like how it doesn't override data by default, requiring me to manually remove and re-download Excel files in my scripts.

dbt, however, was a revelation. Coming from a background in Python and traditional SQL scripting, dbt's approach to modular, testable, version-controlled data transformations felt like a paradigm shift. The ability to write SQL that tells a story through documentation and lineage was something I'd never experienced.

2.2.2 The Microsoft Fabric Ecosystem

Working with Fabric taught me about enterprise-grade data platforms. I learned to implement the medallion architecture (Bronze, Silver, Gold layers) that ACA uses as their standard. Bronze for raw data, Silver for cleaned and enriched data, and Gold for business-ready analytics. This wasn't just a technical pattern - it was a way of thinking about data quality and governance that shaped my entire approach to the project.

2.3 Major Technical Breakthroughs

2.3.1 The Location Code Revolution (Week 8)

My first major confidence breakthrough came when I solved the location code assignment problem. Previously, each input sheet had to be loaded individually with hardcoded location mappings scattered throughout the transformation code. I created a dynamic system that automatically grabs location codes from a central reference sheet and assigns them to data rows during integration.

The impact was immediate: I transformed the water data loading process from over 300 lines of repetitive code to just 80 lines using Jinja templating in dbt - and this was just one model among many that benefited from this approach. More importantly, I created a system where adding a new Duvel facility required changes in only 2 places instead of 20+ locations throughout the codebase.

2.3.2 Dynamic Currency Conversion (Week 9)

Building on this success, I developed a year-based currency conversion system that automatically processes conversion rates for any year from 2022 to the previous year. The system dynamically generates queries for all available years, but importantly, Duvel still needs to manually add the new year's conversion rates to their reference sheet before year-end.

This system automatically processes whatever historical conversion rates are available in the input sheet, but requires Duvel to manually update the conversion rates annually. The innovation was making the processing dynamic - instead of hardcoding years 2022, 2023, 2024, the system automatically adapts to whatever years are present in the data, reducing maintenance overhead while still requiring annual rate updates.

2.4 The ClimateCamp Integration Challenge

2.4.1 Weeks 8-10: Meeting External System Requirements

One of the most complex aspects of the project was integrating with ClimateCamp, Duvel's sustainability calculation platform. ClimateCamp required data in extremely specific formats for environmental impact calculations. Every data point needed precise categorization, validation, and formatting according to complex environmental accounting standards.

This challenge taught me that data engineering isn't just about moving data between systems - it's about understanding the business logic and regulatory requirements that govern how that data must be structured. Working with ClimateCamp forced me to think like a sustainability analyst while engineering like a developer.

2.5 Architecture and Code Quality Revolution

2.5.1 Week 10-11: The Great Refactoring

Perhaps my most significant technical achievement was restructuring the entire data pipeline to implement a proper medallion architecture. This wasn't just about reorganizing code - it was about establishing data quality patterns that would scale with Duvel's growing sustainability reporting needs.

I eliminated 253 lines of hardcoded elements from a single dbt model alone, replacing them with dynamic, parameterized solutions. Since this pattern was applied across multiple models throughout the project, the total code reduction was substantial. The result was a system that was not only more maintainable but also more reliable. Adding new locations went from a 20-step process prone to human error to a simple 2-step configuration change.

2.5.2 Performance and Efficiency Gains

The architectural improvements delivered measurable results:

- Reduction in Microsoft Fabric Compute Unit (CU) usage
- Faster data integrations trough optimizations in the integration notebook
- Faster data transformations through optimized query patterns
- Reduced maintenance overhead through parameterized pipelines
- Enhanced error detection and monitoring capabilities

2.6 Business Impact and Production Deployment

2.6.1 What We Delivered

By the end of 13 weeks, we had transformed Duvel Moortgat's sustainability reporting from a manual, error-prone process into an automated, scalable system:

- 41 completed project tasks covering all major sustainability reporting requirements
- Data integration from 15+ sources including production systems, HR platforms, and external APIs
- Automated ClimateCamp integration reducing manual data preparation from days to hours
- Comprehensive monitoring system with real-time alerts for data quality issues
- Test/production environments enabling safe development and deployment cycles

2.6.2 Current Operational Status

The system is now in production and actively used by Duvel's sustainability team for CSRD compliance reporting. Our dashboard processes data from facilities across Belgium, Netherlands, France, the United Kingdom, Spain, Italy, Czech Republic, China and the United States, providing real-time visibility into sustainability metrics that previously required weeks of manual compilation.

2.6.3 Recommendations for Future Development

Based on our implementation experience, I recommend several next steps:

- API Integration Expansion: Prioritize the Fluvius API integration for real-time energy data, eliminating the last major manual input requirement
- Enhanced Data Validation: Implement additional automated quality checks to further reduce manual oversight needs
- User Training Program: Develop structured training to ensure stakeholders can maximize the system's capabilities

3 PART 2: PERSONAL REFLECTION

3.1 Professional Identity Formation

When I started this internship, I thought data engineering was primarily about writing SQL queries and building ETL pipelines. I discovered it's actually about solving business problems, managing complexity across organizational boundaries, and building systems that people can trust with critical decisions.

3.1.1 From Student to Consultant

The transition from academic projects to professional consulting was profound. In university, requirements are clearly defined and success is measured by grades. In consulting, requirements evolve constantly, stakeholders have competing priorities, and success is measured by business value and client satisfaction.

The ability to work closely together with Thibo, my primary mentor, was instrumental in this transformation. Rather than just teaching me technical skills, Thibo helped me understand the consulting mindset: always consider the business context, communicate clearly about risks and trade-offs, and build relationships alongside technical solutions.

3.2 Technical Growth Beyond Tools

3.2.1 Microsoft Fabric Mastery

Learning Microsoft Fabric taught me more than just another platform - it taught me how to work with bleeding-edge technology in production environments. Since Fabric is still actively being developed with many features in preview, I learned to adapt to platform limitations, work around quirks, and build robust solutions despite technological uncertainty.

3.2.2 The Power of dbt

dbt fundamentally changed how I think about data transformation. The combination of SQL, version control, testing, and documentation created a development experience that felt both powerful and maintainable. More importantly, dbt's emphasis on clear naming and documentation taught me that code should tell a story that future maintainers can follow.

3.2.3 Security and Enterprise Thinking

Moving credentials from hardcoded values to Azure KeyVault wasn't just a security improvement - it taught me that professional development requires thinking about operations, maintenance, and security from day one. This experience made me appreciate that secure, maintainable systems require additional complexity, but that complexity pays dividends in reliability and peace of mind.

3.3 Problem-Solving Evolution

3.3.1 Early Uncertainty to Systematic Confidence

My approach to complex problems evolved dramatically over the 13 weeks. Initially, when faced with issues in weeks 2-3, I tried to solve everything myself, spending hours debugging configuration files. I learned that in professional environments, knowing when to ask for help and how to communicate problems clearly is just as important as technical troubleshooting skills.

By the time I encountered the ClimateCamp integration challenge in weeks 8-10, I had developed a systematic approach: break complex problems into components, research the business context, design step-by-step solutions, and communicate progress clearly to stakeholders.

3.3.2 The Data Quality Mindset

Working with Duvel's input sheets taught me that data engineering is as much about human psychology as it is about technology. People move entire data tables three rows down for no apparent reason, leave rows half-filled, hide problematic data instead of fixing it, and add comments in unexpected places.

Learning to anticipate and handle these human factors taught me that robust data systems must be designed for human unpredictability. The monitoring and validation systems I built weren't just technical solutions - they were responses to the reality of how people interact with data in high-pressure environments.

3.4 Communication and Stakeholder Management

3.4.1 Learning Multiple Languages

One of my most significant growth areas was learning to communicate the same technical concept to different audiences. For example, when explaining our location code solution:

- To Thibo (Technical Mentor): "I implemented dynamic location mapping using Jinja macros in dbt, which eliminated hardcoded location references and reduced the codebase by 253 lines."
- To Dries (Sustainability Manager): "We automated the process of linking data to facility locations, so adding a new site now takes 5 minutes instead of 2 hours and eliminates the risk of mapping errors."
- To End Users: "The system now automatically knows which data belongs to which brewery, so you don't need to worry about manually tagging everything."

This experience taught me that technical expertise without communication skills severely limits your impact as a consultant.

3.5 Personal Insights and Growth

3.5.1 Discovering My Professional Strengths

This internship helped me identify what energizes me professionally. I discovered that I thrive at the intersection of technical complexity and business impact. Building elegant technical solutions is satisfying, but seeing how those solutions enable better decision-making and regulatory compliance is what truly motivates me.

3.5.2 The Value of Collaborative Expertise

Working with Fabian taught me the power of complementary skills. While I focused on data engineering and pipeline development, Fabian managed client communication and dashboard development. This specialization allowed us to deliver higher quality work than either of us could have achieved alone.

3.5.3 Building Confidence Through Contribution

Perhaps the most significant personal growth was developing genuine confidence in my ability to contribute meaningfully to complex professional projects. Starting as someone who had never used the primary technologies, and ending as someone who could implement effective solutions and explain these technologies to others, was profoundly empowering.

3.6 Technical Passion Project: The Home Lab

3.6.1 Open Source Alternative Development

In my free time, I built a complete open-source alternative to Microsoft Fabric on my home server. Using Jupyter notebooks, Airflow, MinIO, PostgreSQL, dbt, and Superset, I recreated our entire data pipeline architecture at smaller scale. This wasn't just a technical exercise - it reflected my deep interest in privacy, security, and understanding technology at a fundamental level.

The home lab project taught me that the tools matter less than understanding the underlying patterns. Whether using Microsoft's enterprise stack or assembling open-source components, the core challenges of data quality, pipeline reliability, and user experience remain the same.

3.7 Career Clarity and Future Direction

3.7.1 The Cybersecurity Pivot

While this internship confirmed my passion for data engineering, it also reminded me of my broader interests across technology fields. My decision to pursue a master's degree in cybersecurity abroad represents my curiosity and passion for exploring different areas of technology - whether that's AI, programming, data engineering, or cybersecurity.

Cybersecurity has always been one of my interests alongside data engineering, not necessarily as a combination but as separate fields that both fascinate me. The internship experience gave me confidence that I can adapt to and excel in different technology domains, making me excited to explore cybersecurity as its own field of study.

3.7.2 Long-term Professional Vision

This internship taught me that I can thrive in different technology domains and that I'm genuinely excited by multiple fields - data engineering, AI, cybersecurity, and programming. Whether I continue in data engineering, pursue cybersecurity specialization, or explore other technology areas, the foundation of problem-solving, stakeholder communication, and professional growth I built during this internship will be invaluable.

The experience showed me that technical skills transfer across domains, but more importantly, that the ability to learn, adapt, and deliver business value is what matters most in any technology career.

3.8 Competencies Developed

3.8.1 Technical Competencies

- Enterprise data platform development and architecture design
- Microsoft Fabric ecosystem expertise
- dbt for data transformation and modeling
- Security best practices for cloud environments
- Data quality management and validation strategies
- Pipeline orchestration and monitoring

3.8.2 Professional Competencies

- Client relationship management and stakeholder communication
- Project execution under evolving requirements
- Technical documentation and knowledge transfer
- Cross-functional collaboration in consulting environments
- Problem-solving under uncertainty and time pressure

3.8.3 Personal Competencies

- Professional communication across technical skill levels
- Adaptability and continuous learning in fast-changing technological environments
- Systematic thinking and architectural decision-making
- Confidence in technical leadership and mentoring

4 CONCLUSION

My internship at ACA Group's Sustainathon project was far more than a capstone to my academic program - it was a complete transformation from student to professional, from individual contributor to collaborative team member, and from someone who writes code to someone who solves business problems using technology.

The technical achievements were substantial: we delivered a production-ready sustainability reporting platform that processes data from 15+ sources across multiple countries, automates complex regulatory calculations, and provides real-time visibility into Duvel Moortgat's environmental impact. But the personal growth was even more valuable.

I developed not just technical skills, but professional competencies: the ability to communicate complex ideas clearly, manage stakeholder relationships effectively, and deliver business value under uncertainty. Working with experienced professionals like Thibo and Kwinten, collaborating with talented peers like Fabian, and contributing to a meaningful project with real regulatory and business impact provided a foundation that could never be gained through academic study alone.

4.1 Looking Forward

As I transition to studying cybersecurity, I carry with me not just technical knowledge and project experience, but a deeper understanding of what it means to be a professional who can bridge the gap between complex technology and business value. The problems I learned to solve at ACA - data quality, system reliability, stakeholder communication, and value delivery under pressure - are universal challenges in technology consulting.

This internship didn't just prepare me for my career - it launched it. The confidence, skills, and professional network I built during these 13 weeks have fundamentally changed my trajectory and capabilities as a technology professional.

4.2 Final Reflection

The most rewarding aspect of this experience was seeing our technical work translate directly into business value. Duvel Moortgat now has a system that not only meets their CSRD compliance requirements but provides the foundation for data-driven sustainability decisions. Knowing that our engineering work contributes to environmental accountability and transparency makes every line of code and every solved problem feel meaningful.

This internship confirmed that technology work is most fulfilling when it serves a purpose larger than the technology itself. As I continue my career, I will seek opportunities to apply technical skills to challenges that matter - whether in cybersecurity, data engineering, or the intersection of both.

This reflection represents my honest assessment of both the technical project we delivered and the personal growth I experienced during my internship at ACA Group. The challenges were real, the learning was intense, and the impact - both on Duvel Moortgat's operations and on my professional development - was transformative.

AI Tool Usage Disclosure

This reflection document was enhanced using Claude (Anthropic's AI assistant) to improve structure, clarity, and academic writing quality while maintaining complete authenticity of content and experiences.

What AI was used for:

- Transforming technical bullet points into engaging narrative storytelling
- Improving document structure and flow according to academic reflection standards
- Enhancing clarity and readability for both technical and non-technical audiences
- Organizing complex technical achievements into coherent sections
- Refining language and professional writing style

Input provided to AI:

- A draft of the initial reflection document, which included the main structure and content filled in up to this point.
- Complete weekly status reports (12 weeks of detailed logs)
- Technical code examples and architectural decisions
- Project presentation materials and documentation
- Specific feedback from supervising professor
- A comprehensive document exceeding 12,000 characters that provided context for my experiences, challenges, and growth moments.
- Clarifications on technical accuracy and project specifics

Authenticity statement:

All experiences, technical achievements, challenges, code examples, team relationships, and personal growth insights described in this document are genuinely from my internship experience at ACA Group. The AI assisted only with organization, clarity, and academic writing structure - not with creating or fabricating content.

Limitations acknowledged:

While AI helped improve the presentation of my experiences, this reflection represents my personal assessment of the internship and may contain subjective interpretations of events and growth that are unique to my perspective.