Development Log – MSc Research Poster Project  
Student: Efthymios Galanidis  
Co-authors: Ioannis Amourgianos, Dimitrios Xatzis  
Poster Title: Optimizing Compliance and Efficiency: Integrating Data Warehouses and Data Lakes under OECD Pillar Two Regulations

**Week 1: Topic Selection and Preliminary Research**  
This week was about aligning our interests as a team and settling on a focused topic. After some discussion, we agreed to explore how data architectures can support compliance with OECD Pillar Two regulations. Since I work in fields related to tax and financial reporting, the core ideas in the OECD (2023) document were already familiar to me, which helped me grasp the framework quickly. I proposed looking into Lakehouse architectures and sketched the first rough layout of the poster. We divided the reading material; I took OECD (2023) and Harby and Zulkernine (2022) to understand the regulatory context and technical foundations. We agreed to frame the architecture comparisons visually from the start, which helped define our direction early.

**Week 2: Research Division and Visual Planning**  
Each of us focused on different sources to avoid overlap. My task was to continue exploring the two references I started with, while the others looked into newer comparative studies. We worked together on drafting the first version of the compliance comparison table, and I suggested including dimensions like auditability, processing speed, and data structure. I remember one evening when we couldn’t agree on which chart format to use—a radar or a bar chart. Eventually, we sketched both and tested them side by side. John, with his visual design experience, had good instincts for layout balance and helped shape the early drafts. Translating dense articles into a clean chart felt like solving a puzzle.

**Week 3: Writing and Visual Execution**  
This week was focused on building the content. I worked mostly on drafting the methodology and refining the literature review based on OECD (2023) and Harby and Zulkernine (2022). Meanwhile, we all contributed to building the performance chart using benchmarking data from Salqvist (2023). We spent one evening testing chart colours and label placements—at one point, I accidentally erased the legend and had to reformat the entire chart! We also double-checked that our graph didn’t confuse users—e.g. using seconds instead of milliseconds to avoid ambiguity. I realised that visual clarity is just as important as academic accuracy. That small choice about time units ended up shaping our approach to consistency across the poster.

**Week 4: Presentation Prep and Peer Editing**  
This was a quieter week content-wise, but crucial for preparing for our final presentation. Dimitrios helped draft the speaker notes, and I rehearsed how to explain our visual comparisons. We fine-tuned transitions and ensured our visuals were legible at a distance. I remember revisiting the auditability row in our comparison table and wondering if it needed more context. In the end, we rehearsed how to explain it verbally rather than expanding the text. John continued helping refine visual details, especially around spacing and chart legibility. Reviewing sample posters gave us ideas for trimming unnecessary text and boosting clarity through visual hierarchy.

**Week 5: Presentation and Final Polish**  
I presented the poster in class, while Ioannis and Dimitrios were present to support in case questions came up. Our tutor commented that the poster was well-designed and clearly presented. Although there was no formal Q&A, John had reviewed technical terms in advance to be ready if clarification was needed. As a group, we reviewed the layout one last time, fixed minor spacing issues, and simplified some text in the Future Work section. This project was a real lesson in interdisciplinary teamwork—blending technical research, design logic, and concise writing. It made me more confident in visual storytelling and group collaboration.

References  
Armbrust, M. et al. (2020) 'Delta Lake: High-Performance ACID Table Storage over Cloud Object Stores', *Proceedings of the VLDB Endowment*, 13(12), pp. 3411–3424. Available at: https://www.vldb.org/pvldb/vol13/p3411-armbrust.pdf

Garg, V. (2023) 'Evaluating Data Lakes, Data Warehouses, and NoSQL Databases as Foundations for Modern Analytics Platforms', *Journal of Mathematical & Computer Applications*, 2(4), pp. 1–3. Available at: https://www.onlinescientificresearch.com/articles/evaluating-data-lakes-data-warehouses-and-nosql-databases-as-foundations-for-modern-analytics-platforms.pdf

Harby, A.A. and Zulkernine, F. (2022) 'From Data Warehouse to Lakehouse: A Comparative Review', *2022 IEEE International Conference on Big Data*. IEEE. Available at: https://ieeexplore.ieee.org/document/10020719

Nambiar, A. and Mundra, D. (2022) 'An Overview of Data Warehouse and Data Lake in Modern Enterprise Data Management', *Big Data and Cognitive Computing*, 6(4), p.132. Available at: https://www.mdpi.com/2504-2289/6/4/132

OECD (2023) *The Pillar Two Rules in a Nutshell*. Paris: OECD Publishing. Available at: https://www.oecd.org/content/dam/oecd/en/topics/policy-sub-issues/global-minimum-tax/pillar-two-model-rules-in-a-nutshell.pdf

Salqvist, P. (2023) *A Comparative Study of the Data Warehouse and Data Lakehouse Architectures*. MSc thesis. KTH Royal Institute of Technology. Available at: https://kth.diva-portal.org/smash/get/diva2:1851711/FULLTEXT01.pdf