Our poster explores how businesses can adapt their **data architectures** to comply with the **OECD’s Pillar Two regulations**, which introduce a **global minimum tax framework**. These rules demand a high level of **financial transparency, risk reporting**, and **data governance**—and this is especially challenging for large, multinational organizations managing complex data systems.

**🔍 Why this topic?**

We chose this topic because Pillar Two marks a major shift in international taxation policy. It forces companies to rethink how they manage financial and compliance data. Traditional **data warehouses** and newer **data lakes** each have strengths—but neither is perfect on its own. That’s where the **Lakehouse architecture** comes in: it combines the structured, fast-query strengths of warehouses with the flexibility and scalability of data lakes.

**🎯 Our Research Objective**

We wanted to find out how companies can adapt their architecture using this hybrid approach—**the Lakehouse**—to meet compliance goals **efficiently and accurately**. That includes reducing costs, improving auditability, and enabling real-time reporting.

**🧪 Our Methodology**

We applied a mix of methods:

* **Qualitative**: We reviewed OECD regulations and data management literature.
* **Quantitative**: We used **TPC-DS benchmarks** to assess system performance (Salqvist, 2023).
* **Case studies**: We looked at how actual companies handle compliance and data integration.

**📚 What the Literature Says**

According to **Garg (2023)**, data warehouses are strong in performance and integration, while data lakes offer better scalability and cost-efficiency. **Lakehouses**, as discussed by **Harby & Zulkernine (2022)** and **Nambiar & Mundra (2022)**, blend these strengths and are ideal for meeting dynamic compliance requirements like those in Pillar Two.

## Diagram 1: Architecture Comparison

From left to right, this schematic visualizes the evolution of enterprise data architecture — starting with Data Warehouses, transitioning to Data Lakes, and culminating in the Data Lakehouse. Each architecture is shown with its data flow and key capabilities, including support for BI, reports, data science, and machine learning. The diagram emphasizes that Lakehouses unify the structure and control of Data Warehouses with the **flexibility and scalability** of Data Lakes, offering a single platform for diverse, compliance-critical use cases.

## Diagram 2: Compliance Capability Comparison Table

This table compares how Data Warehouses, Data Lakes, and Lakehouses perform across key compliance dimensions such as **data structure, processing speed, scalability, and auditability**. Lakehouses demonstrate the most balanced performance, combining strong real-time reporting, hybrid data accuracy, and flexible scalability. Warehouses excel in speed and precision but are limited in flexibility and cost-efficiency, while Data Lakes provide scalability and low cost at the expense of real-time processing and data accuracy.

## Diagram 3: Performance Benchmarking for Compliance Queries

This bar chart compares average response times for compliance-related queries across three architectures. Data Warehouses perform best (~2s), followed by Lakehouses (~4s), while Data Lakes are the slowest (~8s). The graph shows how Lakehouses provide a practical balance — fast enough for real-time compliance, while supporting flexible data use.

## Diagram 4: Use Case Suitability for Pillar Two Compliance

This radar chart evaluates how well each architecture supports three key use cases: Compliance Reporting, Risk Assessment, and Advanced Analytics. Lakehouses demonstrate strong, balanced performance across all areas. Data Warehouses perform well in reporting but fall short in analytics and flexibility, while Data Lakes excel in analytics but lack precision in reporting and risk evaluation.

**📊 Our Key Findings**

* **Data Warehouses** are great for structured financial data but can be expensive and less scalable.
* **Data Lakes** are cheaper and scalable but less precise.
* **Lakehouses** give you the best of both—making them the most suitable for real-time compliance reporting, analytics, and auditability (Armbrust et al., 2020).

**✅ Conclusion & Recommendations**

We recommend that organizations **adopt a Lakehouse architecture** as a foundation for Pillar Two compliance. But—**they also need a clear strategic plan** to align technology with regulatory requirements. It's not just about the tools, it's about governance and planning.

**🔮 Future Work**

We’re proposing several next steps:

* **Pilot studies** to validate these findings in real companies,
* **Advanced analytics** to predict compliance risks,
* And even the potential use of **blockchain** for audit trails.

**Thank you for your time!**

**MEANINGS**

* **OECD Pillar Two:**A global initiative by the **Organisation for Economic Co-operation and Development** (OECD) that enforces a **minimum effective tax rate (15%)** for multinational enterprises. It demands enhanced **financial transparency**, **compliance reporting**, and **cross-border auditability**.
* **Lakehouse Architecture:** A **hybrid data platform** that combines the **structured control** of Data Warehouses with the **scalability and flexibility** of Data Lakes. It supports **real-time analytics**, **machine learning**, and **audit-compliant reporting** in a single system.
* **Data Warehouse:** A **centralized repository** for structured data optimized for **fast queries** and **business intelligence** (BI). Ideal for compliance reports and financial analysis but limited in handling unstructured data.
* **Data Lake:** A **scalable storage system** for raw, semi-structured, and unstructured data.  
  Useful for flexibility and machine learning, but **less governed** and **slower for compliance reporting**.
* **TPC-DS(Transaction Processing Performance Council – Decision Support):** simulates real-world data warehousing environments, making it ideal for testing **query performance**, **scalability**, and **compliance-related reporting capabilities** — especially relevant when comparing architectures like **Data Warehouses, Data Lakes, and Lakehouses**.

**Real-time Compliance Reporting:** This refers to a system’s ability to **generate accurate compliance reports immediately or within very short timeframes**, as data is being ingested or updated. It enables organizations to:

* Detect and report compliance breaches **as they happen**
* Respond quickly to regulatory requests
* Maintain **up-to-date tax and audit records** without manual batch processing

Under regulations like **OECD Pillar Two**, tax calculations and reporting must often happen **within strict timelines**. Real-time capabilities reduce delays and human error.

**Scalability (for future regulations):** This refers to how easily a data system can **grow or adapt** in response to **new regulatory requirements**, larger data volumes, or more complex compliance logic.  
Scalability includes:

* **Horizontal scalability** (adding more machines or storage)
* **Architectural flexibility** (supporting new data types, rules, formats)

Tax and compliance regulations evolve. A scalable system ensures an organization can remain compliant **without needing to rebuild or overhaul** their data infrastructure.