



Success Story

Streamlining business analyses with an automated single source of truth

Boosting insights: How VirtusLab's platform streamlined analyses for a high-stakes insurer, cutting costs and enhancing accuracy.

Our client insures large entities, like ships, satellites, and factories, handling transactions worth hundreds of millions of pounds. They use data from 20+ sources to provide their services. Including all possible data is crucial, as any oversight could have serious consequences. The insurer aimed to build a modern, central analytical data platform for accurate reports and data-driven decision-making. VirtusLab helped create a scalable and error-free platform set up for machine learning, saving time and operational costs.



The Challenge

The insurer had to handle various data sources with different formats and integration methods. Each department downloaded and cleaned data separately to generate meaningful reports, leaving room for errors. The process often resulted in duplicate work on the same data, making maintaining high data quality difficult.

During the cleansing process, conflicting or problematic data was dropped, leading to inconsistencies. A data engineer and a business analyst had to carefully review the data set, engaging in close cooperation since both couldn't find errors individually. This process consumed time and resources, increasing operational costs due to expensive engineer engagement.

Managing intricate data transformations and cleaning procedures was at the forefront of their concerns. Our client aimed to avoid vendor lock-in, seeking to establish a custom and reliable single source of truth to streamline their operations. This was the point when the insurer sought the expertise of VirtusLab to tackle data quality challenges efficiently.



The solution

VirtusLab helped incorporate a medallion architecture as a central hub to collect and refine data. This cloud-based data lake solution was designed with three key layers:

1. **The Raw Layer houses data “as-is”** from diverse sources in a cloud-based
2. **The Curated Layer ensures** that the data is cleaned and undergoes simple transformations to enhance its quality
3. **The Business Aggregates Layer** stores combined and transformed data, offering a unified, single source of truth, and perspective for end users

The solution, built on Scala and Apache Spark, adopted a domain-driven design approach. Users could effortlessly access the data they sought by creating distinct aggregates for each domain. The implementation of the new solution also led to a reduction in data quality issues. This solution enabled Business Analysts to spend less time analysing data quality problems. The time savings came in two ways:

- **Firstly, troublesome data was stored separately** and shown in an easy-to-consume manner, removing the need for analysts to create complex queries.
- **Secondly, as it was a self-served process**, there was no waiting for data engineers to participate.

By keeping the platform simple and free from external dependencies, VirtusLab avoided unnecessary complexities. This approach evades vendor lock-in and simplifies the addition of new components, enabling scalability. VirtusLab fortified the code with the Scala-type system, elevating code correctness and minimising maintenance costs.





The results

The implementation of the custom analytical data platform resulted in a unified, error-free, and scalable solution, meeting the insurer's goal of becoming a modern and data-driven company. They:

1

Processed thousands of tables daily and provided data to various business users and executives

2

Reduced time and costs by eliminating duplicate coding and fixing data quality issues promptly

3

Avoided vendor lock-in and established a future-proof, maintainable data platform

4

Improved data quality and accuracy, leading to better business insights and decision-making

5

Obtained a sought-after solution that is highly desired within the industry but seldom attainable

6

Reduced time and operational costs

The tech stack

/ Languages:

- Scala

/ Infrastructure:

- AWS
- EMR
- S3
- CodeBuild
- EC2

/ Framework & Libraries:

- Spark
- Cats
- Circe



Scala

aws

amazon
EMR

amazon
S3

AWS CodeBuild

EC2

Spark



circe
RESEARCH CENTRE
FOR ENERGY RESOURCES
AND CONSUMPTION



About VirtusLab

At VirtusLab, we aim to lead in software technology, working consistently to enhance efficiency. Our profound commitment to research and development and a dedicated focus on emerging trends and inspirations fuels an innovative culture. This ethos precisely guides advancing our cutting-edge solutions, inviting collaboration to expand the boundaries of software technology collectively. We welcome you to be a part of this transformative journey.

[Let's connect](#)

Contact Details

info@virtuslab.com

POLAND

Kraków Headquarters

Virtus Lab Sp. z o.o.
ul. Szlak 49
31-153 Kraków

GERMANY

Berlin Office

+49 30 52014256
VirtusLab GmbH
Potsdamer Platz 10
10785 Berlin

UNITED KINGDOM

London Office

+44 (0)20 4577 1051
Virtuslab Ltd.
40 Bank Street HQ3
London E14 5NR