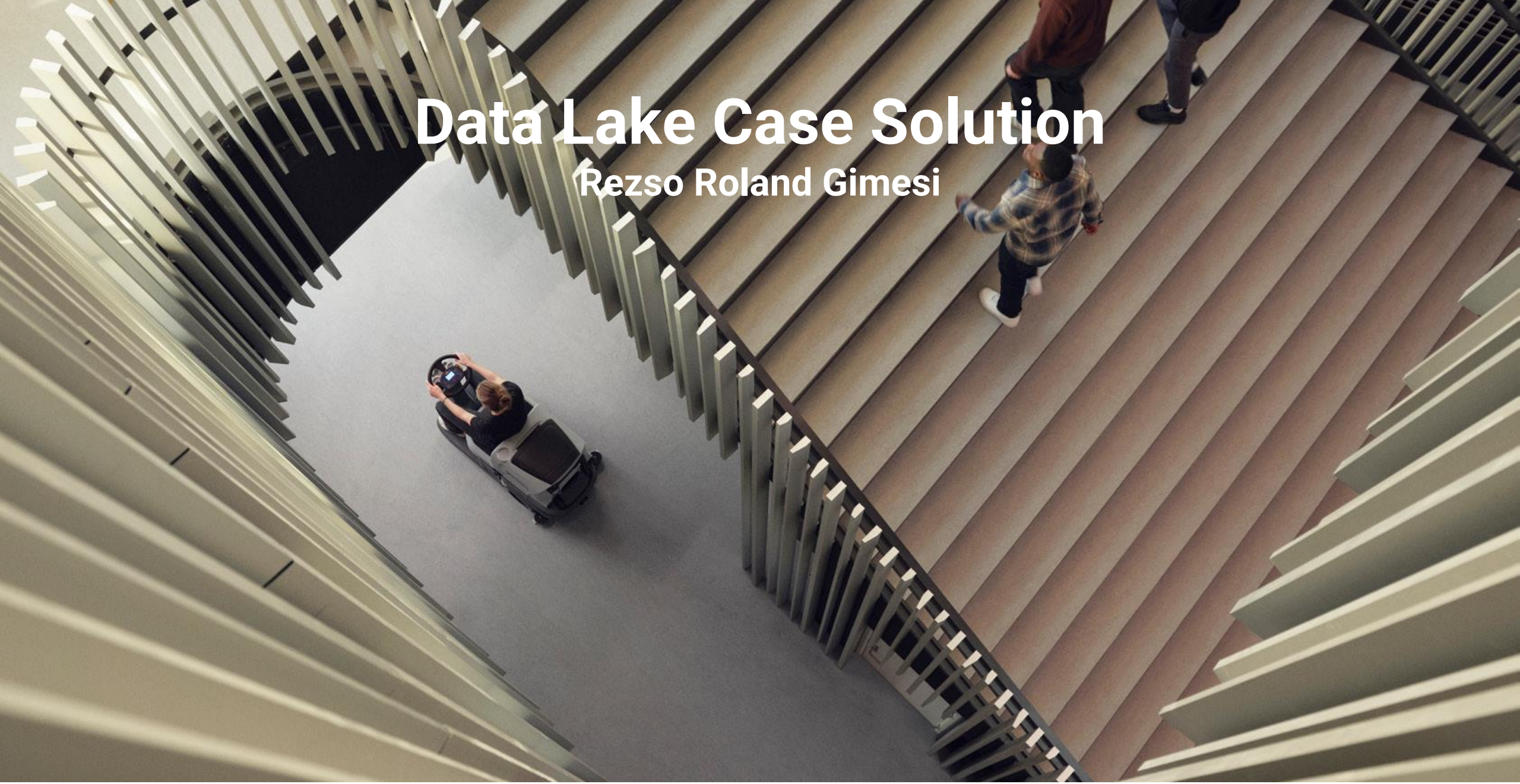


Data Lake Case Solution

Rezso Roland Gimesi

Table of contents

Personal introduction

- Bsc Marketing
- 1,5 years in Digital Marketing
- MSc Information Management, specialized in data analysis
- 1 year Student Worker in Data Analysis at SGRE

Rezso Roland Gimesi

The case

The company is one of the world's leading manufacturers of professional cleaning equipment.

The Global Digital Marketing team aims to create a Data Lake using Google Cloud Platform, as a first step towards their ‘Single Source of Truth’ vision.

The aim is to establishing a data infrastructure and creating a single place for shared knowledge.

The case questions

Considerations and pitfalls

Which considerations and potential pitfalls must the project team on the Marketing Data Lake project be aware of?

Data visualization

You are to provide the leadership team with recommendations on visual design and relevant Marketing KPIs to report on an Executive Summary dashboard in Power BI.

Assumptions

Certain assumptions were made while working on the case.

- The project is already approved by management
- There is sufficient budget available for it
- The company possesses a level of BI maturity that enables the implementation of a project this scale.

Pitfalls

~70% of all projects fail for various reasons.

1. “If we build it, they will come”
    - Every system is only as successful as it is used. → Implementation is key.
2. Resistance among users
    - They don’t know it can help them
    - They have there own ‘Shadow systems’

Solution:

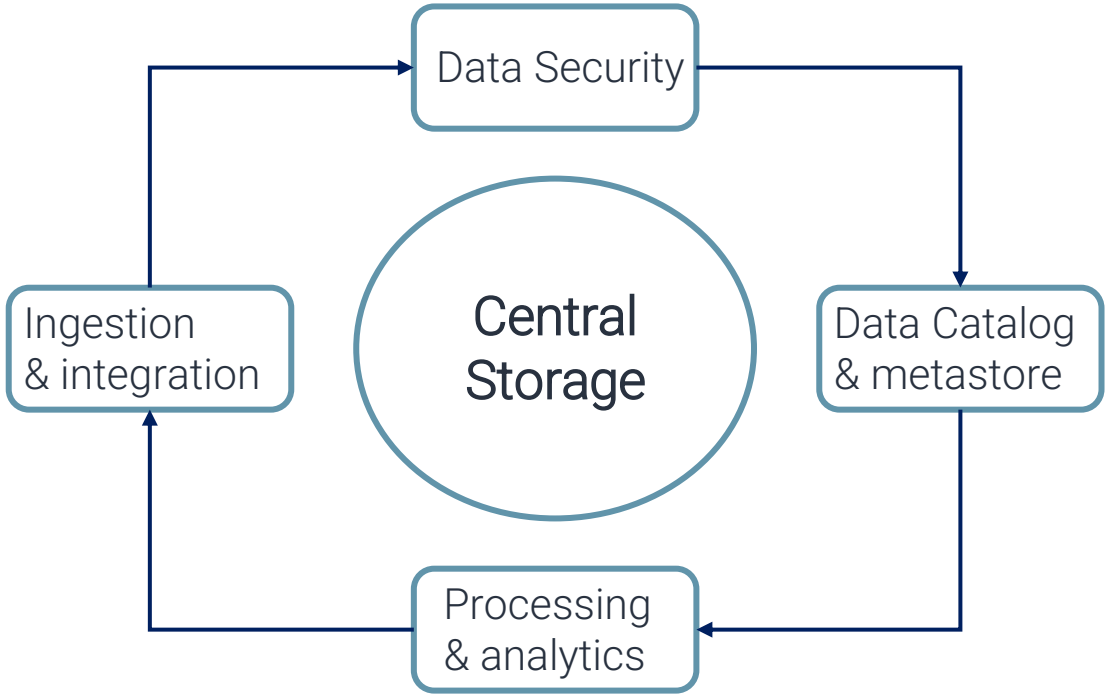
- ✓ Proper change management: Ensure that everyone is on board!
    - Sense of urgency → Vision → Communicate it extensively!
- ✓ Interview key users
    - How do they use data and how do they want to use it?
    - What are their requirements for the system?
    - Ensure that solution satisfies their requirements.

Considerations - Technical

Data lake is not a single product but a network of interconnected systems.

Data Lake architecture

- **Central storage** – where to store the data and how?
- **Data security** – how to protect, who will have access?
- **Data catalog** – what is in the data?
- **Processing & analytics** – how will data be used?
- **Ingestion & Integration** – how to extract the data, clean/transform it and load it into the lake?



Considerations – Marketing/Non-technical

1. Who will use the system?
   - Identify key user groups (include representatives in project team)
   - KPI framework needs to be done in collaboration

2. How should it help the user groups?
   - Create use cases
   - Communicate extensively

3. Who will help the users after the launch
   - Comprehensive, accessible documentation
   - Super user

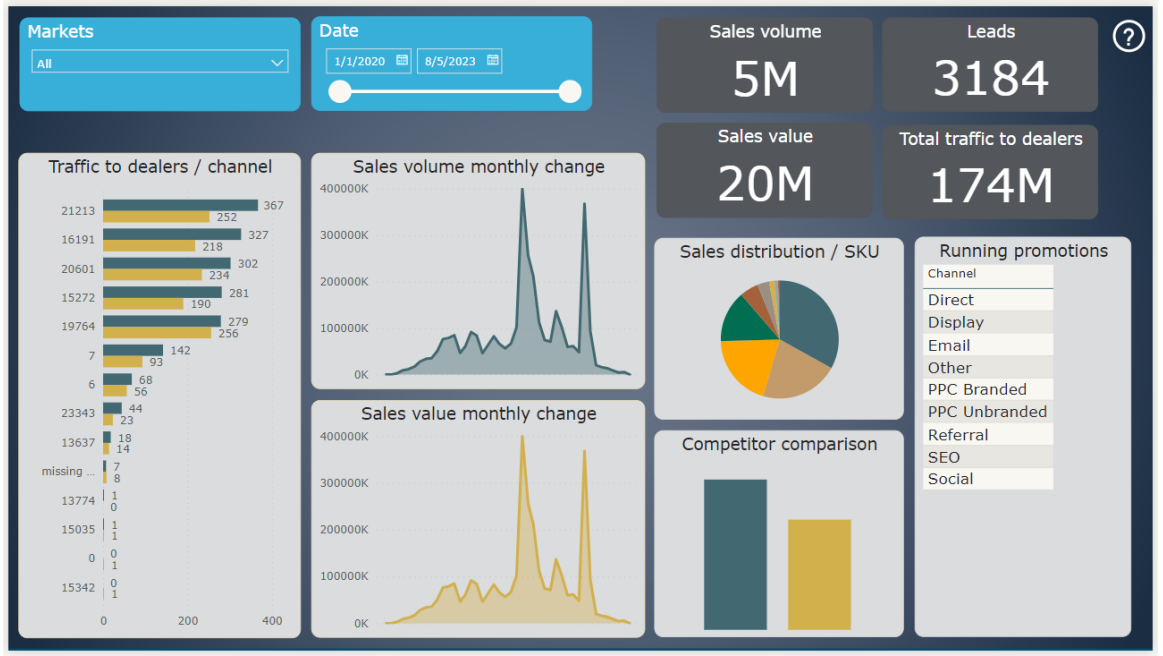
Power BI dashboard design

KPIs – Should be established in collaboration with users

- Sales value, volume (total, monthly change)
- Leads
- Total traffic to dealers
- Sales distribution/SKU
- Traffic to dealers/channel
- Competitor comparison (if data is available)
- Promotions, etc.

Design principles

- Main metrics on one page (possibility to drill down)
- Easily understandable structure (filters, totals, etc.)
- Automatic updates
- Available documentation
- Clean, structured data



Summary

Pitfalls:

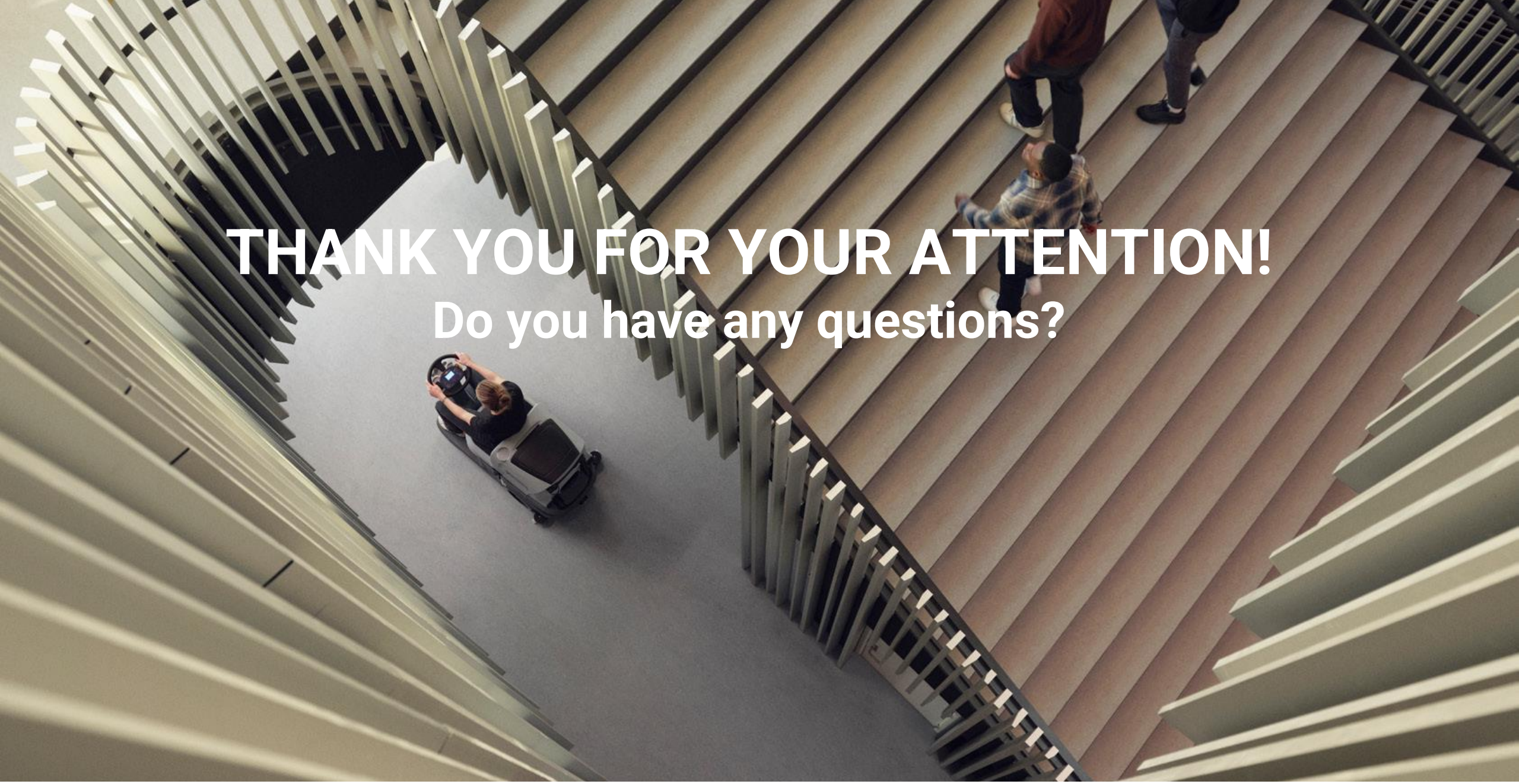
- Lack of focus on implementation & Resistance

Considerations:

- Data lake architecture
- Key users, Use cases, Super user

Dashboard design:

- KPI recommendations
- Design principles


THANK YOU FOR YOUR ATTENTION!
Do you have any questions?