Deloitte.
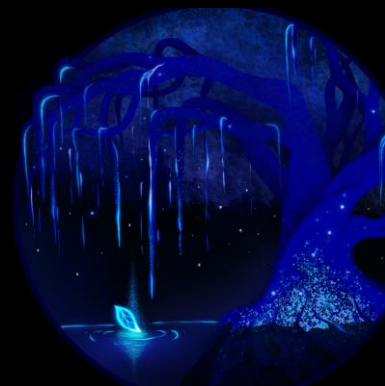
Eywa Unilever Data Lake & BI tool
Project description

Consulting, June 2020

Table of contents

Client issues

Client issues
Client issues

Non-standardized approach to form sales reports
- Number of unconnected IT-systems
- Number manual reports
- Number of unconnected reporting systems

Strong dependency on external data-source
- Low quality data from external provider
- Number of external data sources providing system data in XLSX files

Lack of technical documentation
- IT-systems are not described in technical documentation
- Large amount of outdated documentation

Universal Data Lake / Storage System absence
- Unilever RUS LLC does not have any special data warehouses, which are consolidating whole data

Project description

Project description
What we should do

Jobs to be done

- Create end to end reporting for Retail Execution, Ice Cream, Key Accounts, Distributors business lines

- Make first steps for creating Sales ecosystem

- Deploy centralized data storage and prepare to connect to external vendors to the Company's system

Russia, Belarus - 2020
Ukraine - 2021

Project participants

Client side

Deloitte side

Sales
- Defining a strategy based on data, indicators, and creating a concept for storing and visualizing data

IT
- Coordinating the approach with the global teams
- Technical expertise and provision of Azure infrastructure

Shopper Marketing
- Content of reporting advisory

Deloitte team
- Implementation, technology expertise
- Creation of a scalable solution for subsequent integration to global Data Lake

Project description
Project timeline

February-March

Getting an Azure subscription, interviewing experts, and CD Excellence team

April-May

Creating Azure storage, integration with systems, incremental data transfer
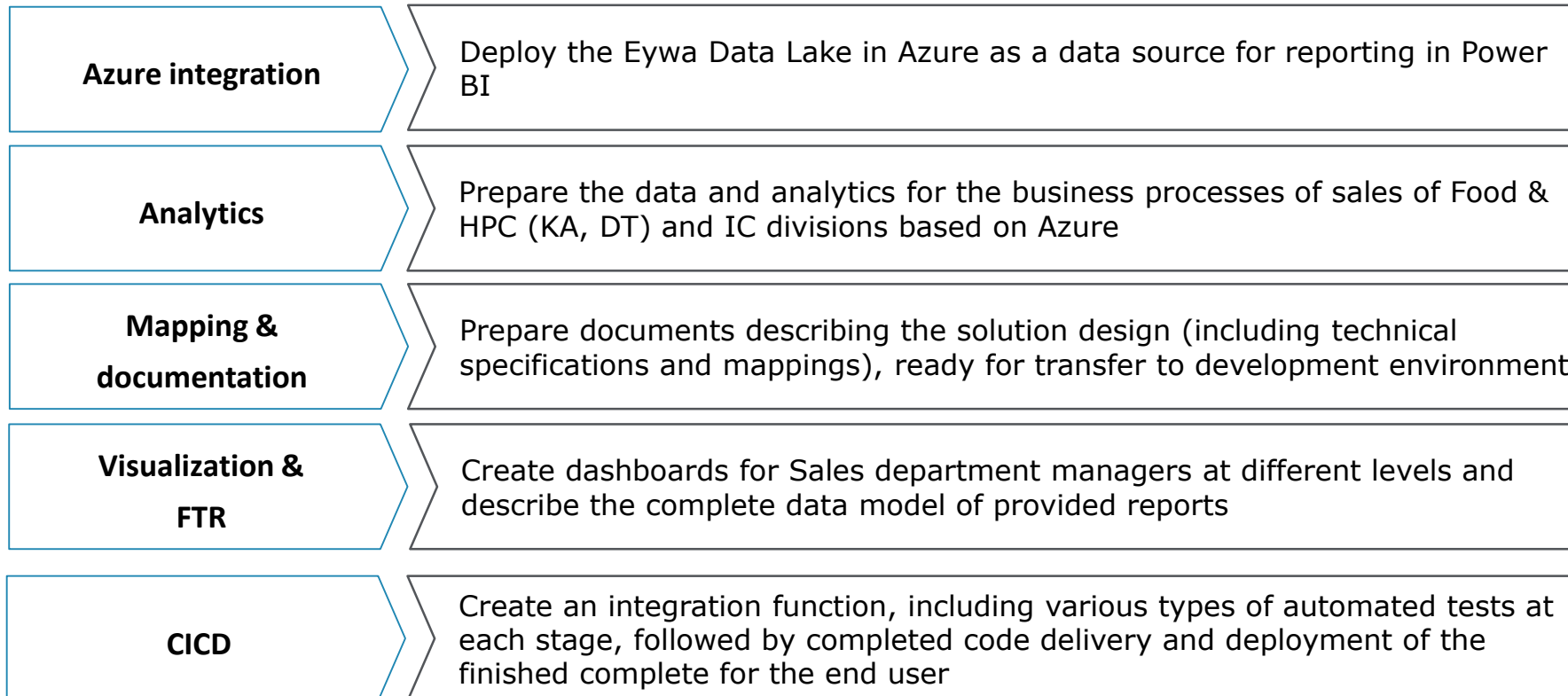
June-July

Building visualization

August

Hypercare

Goals

Goals
Project goals description

The project goal is to provide structured information based on cloud solutions to sales management and field employees of the Unilever sales Department. The data contains current and historical values of key performance indicators of sales business processes which must be visualized n PowerBI.

Azure integration	Deploy the Eywa Data Lake in Azure as a data source for reporting in Power BI
Analytics	Prepare the data and analytics for the business processes of sales of Food & HPC (KA, DT) and IC divisions based on Azure
Mapping & documentation	Prepare documents describing the solution design (including technical specifications and mappings), ready for transfer to development environment
Visualization & FTR	Create dashboards for Sales department managers at different levels and describe the complete data model of provided reports
CICD	Create an integration function, including various types of automated tests at each stage, followed by completed code delivery and deployment of the finished complete for the end user

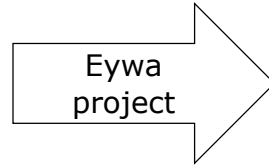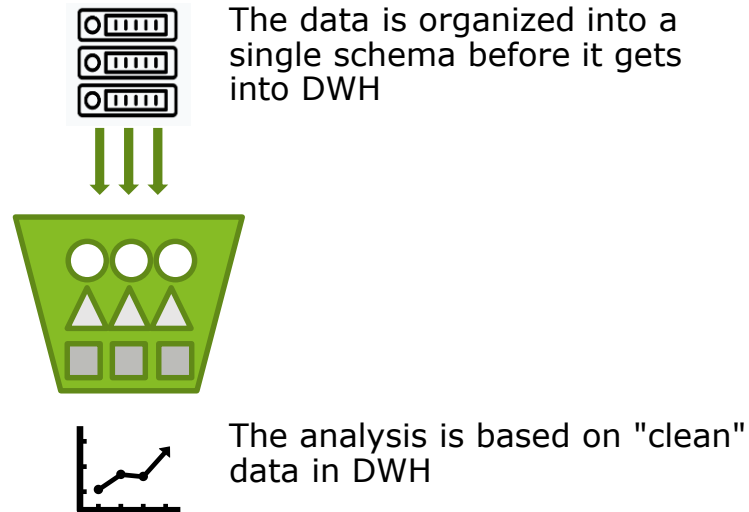# Our approach

Our approach
Common approach description

Data Warehouse is an information Database that is processed and stored in a hardware and software complex that provides access to information and multidimensional data analysis.
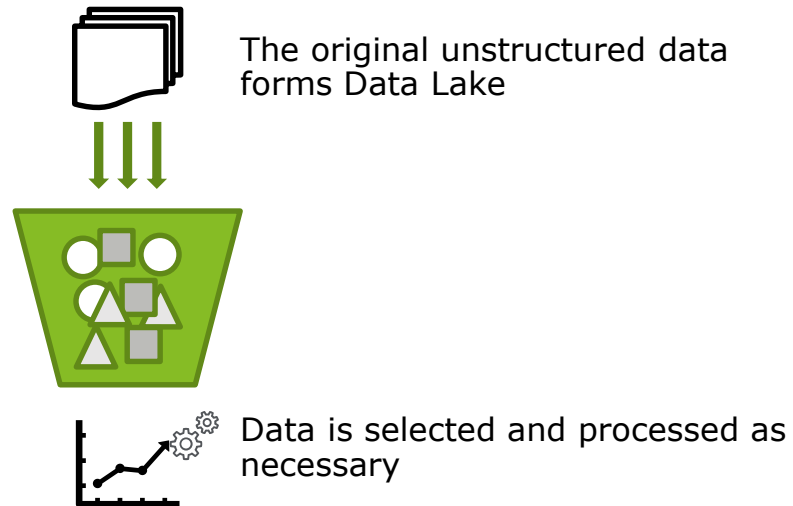
Eywa project

Data Lake is a cloud computing model that provides cloud data storage and data storage and management is provided as a service.

2019

2020

The data is organized into a single schema before it gets into DWH

The original unstructured data forms Data Lake

The analysis is based on "clean" data in DWH

Data is selected and processed as necessary

- Does not require to buy physical hardware

- Pay to go model-you pay exactly for what you use, and you scale the solution faster and cheaper

- Execution of complex analytical queries using massive parallel processing

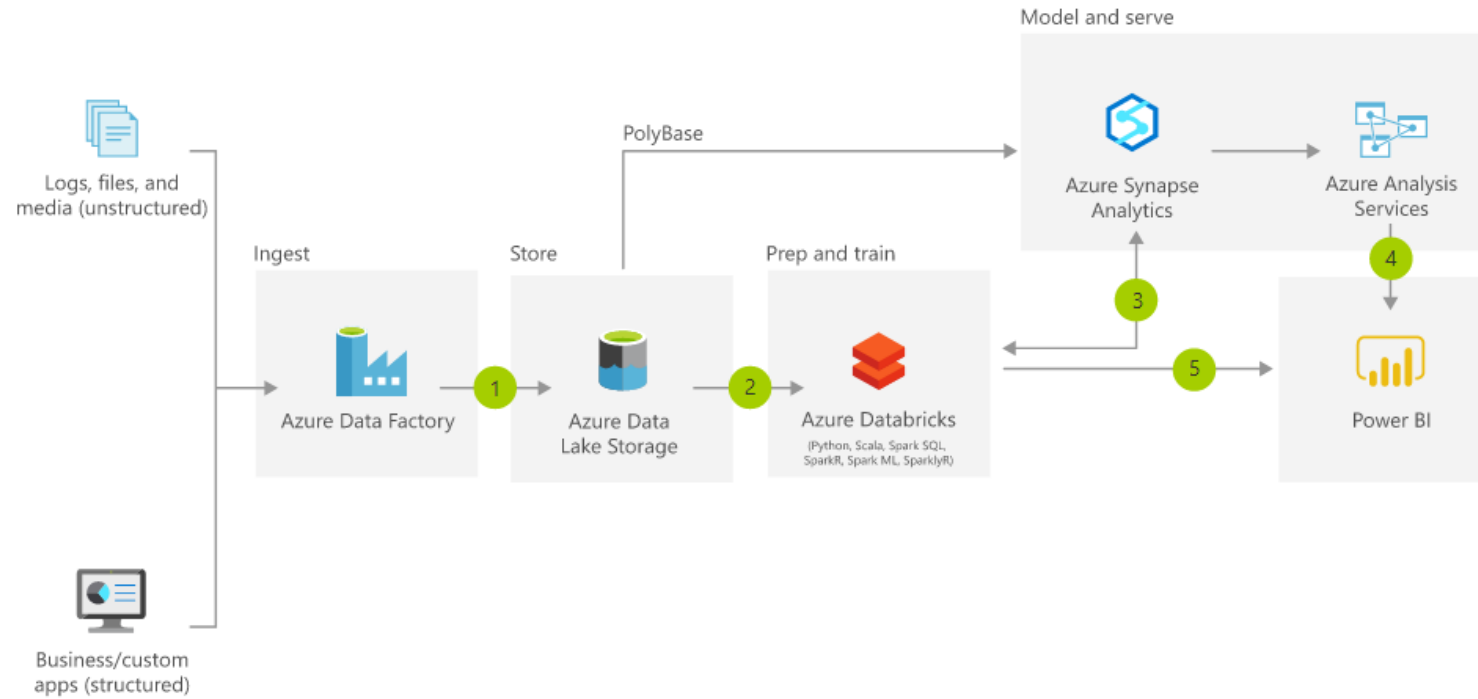- Ability to use machine learning to support strategic decisions using large Big Data

Our approach
Selected Azure architecture

- **Azure Synapse Analytics** is a fast, flexible, and reliable data warehouse that allows you to scale computing and storage resources independently, based on a massively parallel computing architecture

- **Azure Data Factory** is a hybrid data integration service that allows you to create, plan, and organize ETL/ELT workflows

- **Azure large binary object storage** is a convenient and cost-effective massively scalable object storage for any type of unstructured data

- **Azure Databricks** is a fast and convenient Analytics platform based on Apache Spark

- **Azure Analysis Services** is an enterprise-level Analytics service that allows you to confidently manage, deploy, test, and deliver BI solutions

- **Power BI** is a set of business intelligence tools that provide a complete view of your entire organization
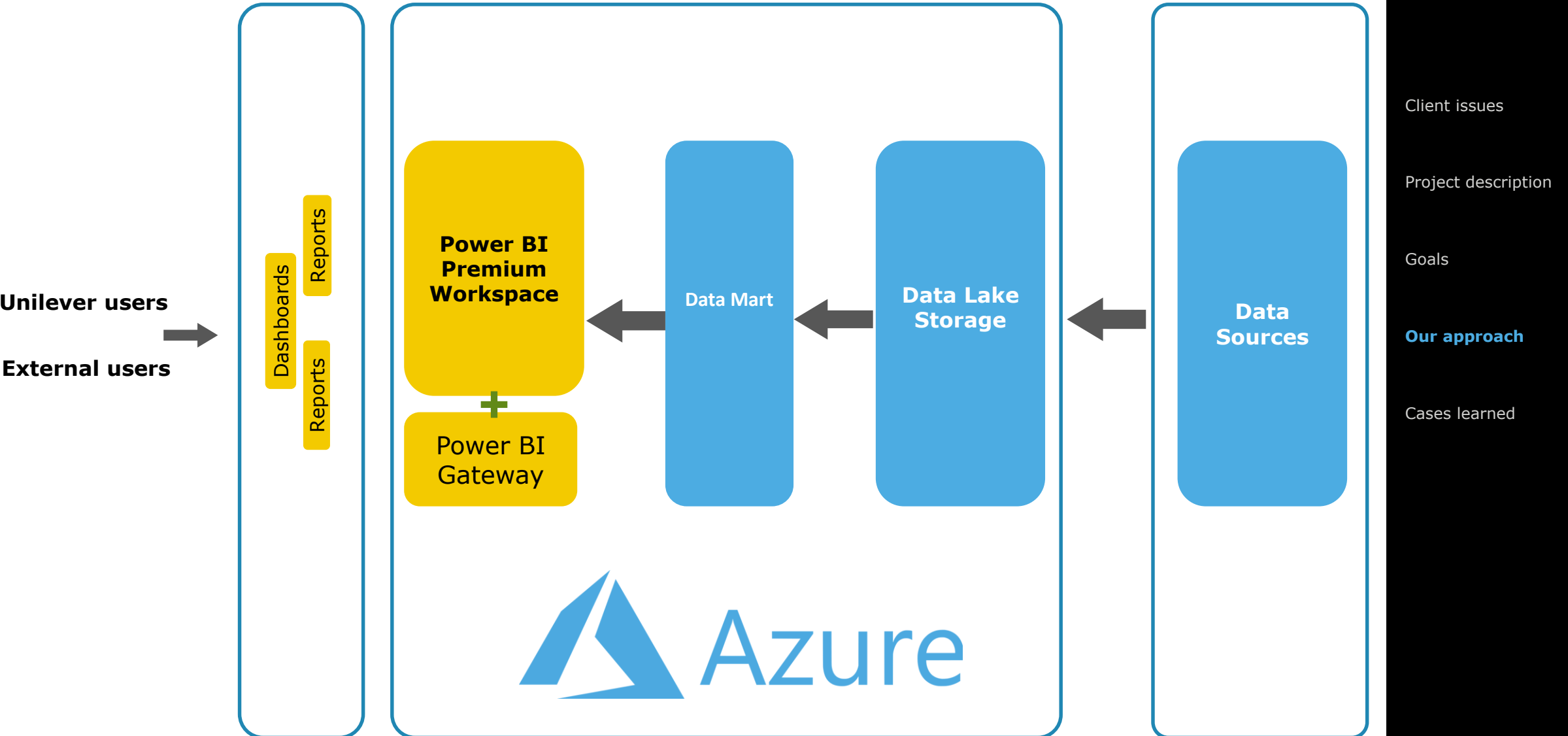
Our approach
Eywa business architecture

Unilever users

External users

Dashboards

Reports

Reports

Power BI Premium Workspace

+

Power BI Gateway

Data Mart

Data Lake Storage

Data Sources

Azure

Cases learned

Cases learned
Important mistakes

Eywa takeaways

Source To Target Mapping
- Prepare highly detailed STTM documents in order to make developers life easier
- In almost all cases analyze client even bad scripts

Bring more prepared cases
- To collect more cases ready to implement through projects

External data providers
- Pay more attention to external data providers and their data transfer method

Collect methodologies
- This project has many interesting cases and methodologies, which can be implemented on other projects