

BUSINESS CASE

One Pager

Time: 5 minutes

Introduction (30 seconds)

Good morning/afternoon. This one-pager provides an overview of a strategic initiative titled Data as a Strategic Asset – Transforming Architecture. Simplifying Decisions. It outlines how Allianz can modernize its data architecture to accelerate growth, enhance operational excellence, and foster innovation, aligned with the company's 2025 goals.

Executive Summary (45 seconds)

Over the Executive Summary, it's possible to see the already mentioned objective, along with a roadmap and resources needed. The plan is structured over less than three years (6 months of planning, 1.5 years of execution, and a final year for system maturation). It is estimated a need for 6 FTEs and a €3 million budget (€1 million for platform investment and €2 million for consultancy support).

Opportunity (1 minute)

In terms of opportunities, today the companies rely on on-premises architecture, which is a major bottleneck. It is costly, inflexible, and difficult to scale. Legacy systems and an outdated data warehouse hinder agility, while reliance on old platforms limits the ability to adopt modern analytics and integration tools. This severely impacts the speed of decision-making and innovation potential.

Solution (1 minute)

However, if the problem cannot exist without its counterpart, a solution – in this case, that comes as cloud-based architecture. By investing €3 million and mobilizing 6 FTEs, within 3 years, we can modernize, boost scalability, agility, and collaboration across the entire organization. This transformation will enable fast, trusted, and intelligent decision-making through AI, automation, and insights.

Strategy (1 minute)

Ensuring the solution also follows in line with the company's strategy for 2025:

- It will drive above-market growth.
 - It will automate processes and enhance technical excellence.
 - It will improve margins and service levels by making the organization more efficient, connected, and data-driven.
-

Closing (30 seconds)

In short, this is not just a technological upgrade — it's a business enabler. Modern data architecture will simplify decisions, fuel innovation, and position Allianz for the best.

BUSINESS CASE

Slide 1 - Introduction

Time: 1 minute

With this presentation, I would like to introduce a prototype that demonstrates how Allianz can transform its data architecture to simplify decision-making and unlock strategic value. In addition, I will address the four other challenges outlined in the assessment document.

Slide 2 – Agenda

Time: 1 minute

This is today's Agenda, and it is composed by main 2 topics:

1. New Architecture Overview
 2. E2E Data Engineering Projects:
- Architecture Diagram, Layers & Functionalities, Artifacts and Modules
 - Sales Dashboard and Churn Prediction Model

Slide 3 - New Architecture Overview

Time: 15 seconds

The first topic, New Architecture Overview

Slide 4 - Architecture Diagram

Time: 5 minutes

For the diagram, I bring in an architecture with two different streams.

- Main Stream: Represents the data flow through the architecture. Data originates from the source, IBM DB2, within the Operational Stack. It is ingested using Azure Synapse Analytics Pipelines and Virtual Machines (for triggers), then stored in Azure Data Lake Store Gen2 across the Bronze and Silver layers. Processing occurs through Azure Synapse Serverless and Dedicated Pools, particularly in the **Gold layer***. Data Stewards, typically associated with Data Engineers, overseeing quality throughout this flow. Enrichment and serving are handled by **Azure AI Services****, Machine Learning, and Power BI, delivering outputs and insights that support business processes and functional areas.

* Gold Layer in Process ; ** Azure AI Foundry icon;

- Support Stream: Covers the essential elements that enable the architecture to operate securely and efficiently. Access management is handled through Active Directory (AD) Groups, ensuring the right people have the right permissions. Governance is supported by Microsoft Purview, with Governance Stewards, typically Data Architects, overseeing compliance and standards. Maintenance relies on seven key applications to monitor, automate, and sustain the architecture. Finally, collaboration and documentation are managed through Jira and Confluence, providing structure and visibility for ongoing activities.
- Notes: This architecture is primarily focused on structured data and batch ingestion. However, it is designed to be extensible, supporting semi-structured, unstructured, and streaming data if needed. While the initial focus is on serving internal stakeholders, the setup can also be extended to support third-party access when required.

Slide 5 - Architecture Layers and Functionalities

Time: 5 minutes

In this slide, we can see what each item does across streams, layers, and roles. I'll focus on the items that are less intuitive to help us stay on time. Specifically, I'll explain the data layers and the key governance and maintenance aspects.

- Layers: In a modern data architecture, data is organized into layers to ensure quality, reliability, and performance as it moves through the system.

Bronze refers to raw data captured directly from source systems with minimal processing.	Silver represents cleaned, filtered, and structured data, ready for basic analytics and reporting.	Gold is the refined, highly curated data that is optimized for business use, advanced analytics, and decision-making.
--	--	---
- Governance: Governance over data architecture ensures that data is accurate, secure, and properly managed across the organization. It defines clear roles, responsibilities, and processes for how data is created, maintained, accessed, and protected. Good governance enforces data quality, compliance with regulations, security standards, and controlled access, helping build trust in the data. Ultimately, it allows the architecture to scale effectively while enabling faster, smarter, and safer decision-making.

BUSINESS CASE

- **Maintenance:** Maintenance of the data architecture ensures that the environment remains healthy, secure, and cost-effective over time. It involves actively monitoring system performance, managing cloud costs through Cost Management tools, and using continuous integration/continuous deployment (CI/CD) pipelines to keep updates safe, automated, and efficient. Maintenance also includes enforcing Policies to control data usage and security standards, and managing identities and access through Entra ID (Microsoft's modern identity platform). Together, these activities keep the architecture optimized, resilient, and aligned with business and regulatory needs.

Slide 6 – Data Ops Artifacts and Modules

Time: 3 minutes

Here, given the target architecture, I emphasize potential enhancement opportunities if we aim to reduce maintenance and operating costs.

- **Synapse → Azure Data Factory** for better scalability and cost-efficiency.
- **VMs → Azure Logic Apps** for simpler, serverless orchestration.
- **Fuse Silver + Gold Layer** to reduces redundancy and minimizes maintenance efforts.
- **Spark Pools** for heavy transformations.

To identify these opportunities, I would rely on tools like Policy, Monitor, and especially Cost Management, where I can input budget constraints and guide optimization efforts.

Slide 7 - E2E Data Engineering Projects

Time: 15 seconds

The second topic, E2E Data Engineering Project:

- Sales Dashboard
- Churn Prediction Model

Slide 8 - Sales Dashboard — Data Model Design

Time: 2 minutes

For the sales dashboard, I have decided to proceed with a star schema, featuring a fact table connected to five dimensions: client, product, channel, region, and time, all within Power BI. This structure enables insights into timely performance, demographic analysis, policy retention, and identification of at-risk customers. To illustrate the data flow, I will also provide examples of how to establish connections with the database and navigate its tables and views.

Slide 9 – High Level Project Plan: Sales Dashboard

Time: 3 minutes

In this slide, I present how the dashboard can be built. However, before explaining the project plan, it is important to clarify a few assumptions that help narrow the scope and avoid making the approach too generic.

Assumptions:

Business requirements will be finalized during the Design phase without major changes; all necessary data sources will be accessible; stakeholders will be available for UAT and feedback sessions as planned; and budget approval for the external consultant will be secured prior to starting the Build phase.

Roadmap:

The project will be executed in four phases over 4–6 weeks: **Design** will gather business requirements, define the data model, and create dashboard wireframes; **Build** will handle data ingestion, star schema modeling, and dashboard development in Power BI; **Test** will focus on data validation, stakeholder review, and User Acceptance Testing (UAT); and **Launch** will deliver user training, production deployment, and hypercare support to ensure smooth adoption.

Resources:

The team will consist of four FTEs: a Data Engineer responsible for ETL development, a Data Architect to design scalable and high-performing data structures, a Data Analyst to translate business needs into dashboard outputs, and a Process Owner to ensure alignment with operational requirements.

Costs:

Monthly operational costs (OPEX) will include software licensing and Azure services, depending on the final deployment architecture. If any initial infrastructure investment (CAPEX) is needed, it will be estimated and allocated separately. This is an estimation, and the values provided were created using Microsoft's cost simulator, which may contain significant discrepancies when applied to corporate environments.

Slide 10 - Churn Prediction — Data Model Design

Time: 2 minutes

BUSINESS CASE

For the churn prediction model, I am applying a similar approach; however, this process involves completing the data science development lifecycle, and in parallel, using the star schema to support feature selection during data-centric steps. The final output is a churn probability score ranging from 0 to 1.

Slide 11 – High Level Project Plan: Churn Prediction

Time: 2 minutes

Once again, following the same logic as for the Sales Dashboard.

Assumptions:

The assumptions for this project will remain the same as for the Sales Dashboard, with the addition that historical data on customer behavior is clean and reliable enough to build a solid predictive model.

Roadmap:

The churn prediction model will be developed over a 6–8 week period. By nature, data science initiatives are slightly longer due to the additional steps involved. The first phase, **Problem Understanding**, will define the scope of churn prediction and establish the business objectives. In the **Data Mining** phase, the focus will be on identifying and gathering relevant data sources. The **Preparation** phase will involve cleaning and structuring the data for analysis. During **Feature Engineering**, meaningful variables will be created to capture patterns in customer behavior. The **Modeling** phase will apply machine learning algorithms to predict churn probability. In the **Evaluation** phase, the model's performance will be assessed using metrics such as precision and recall. Finally, during **Deployment**, the model will be integrated into production to enable real-time churn predictions.

Resources:

The team for this project will consist of five FTEs, each with a specific role. A Data Engineer will be responsible for data ingestion and preparation. A Data Architect will design the data architecture and ensure its scalability. A Data Analyst will perform exploratory data analysis and prepare the datasets for modeling. A Data Scientist will develop and train the churn prediction model, including feature selection and tuning the machine learning algorithms. Lastly, a Process Owner will oversee the project execution, ensuring it aligns with the operational objectives.

Costs:

The cost structure will remain the same, with adjustments made for the specific functionalities and tools required.

Slide 12 - Thank You / Questions

Time: 30 seconds

That concludes my prototype demonstration. Thank you for your time.

You can access the one-pager, this presentation, and my speaking notes on my GitHub. I'll share the link in the chat. Now, I'm happy to open the floor for any questions.