March 2023

VENZO

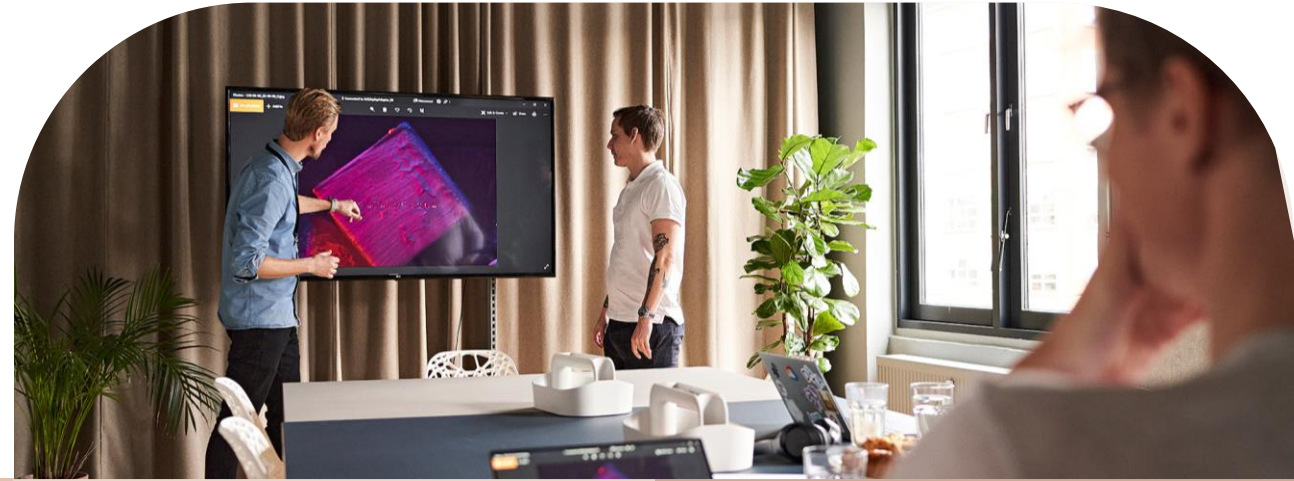Data Archiving — PoC Planning Workshop

VENZO_

Content

- The VENZO team & introduction

- Approach and Data Platform

- Standard data platform

- Archive Data Lake

- Data ingestion

- Governance & Security

- Q & A / Discussion

About VENZO

VENZO_

Different specialties.
One direction.

15

Years of customer success

240+

Full-time employees

340+

mDKK in revenues in 2020

110+

Customers across public and private sector

VENZO_

Our domains of expertise

VENZO_

Strategy & Execution

Automation & AI

Digital Life & Cloud

Trust & Security
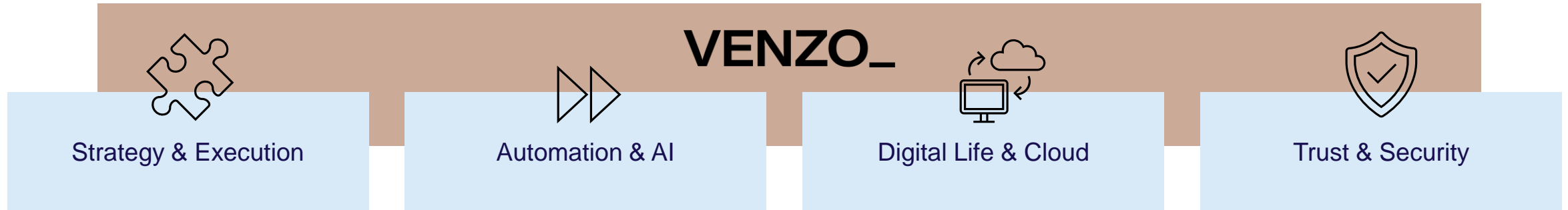
Focus on setting a direction for digital solutions and ensuring that the receiving organisation is well equipped and capable of driving value from digitisation.

This includes
- Formulating strategies
- Mapping and aligning people, processes and technology
- Creating implementation and adoption plans
- Anchoring the change rightfully in the organization

Focus on growing revenue and improving profitability through data insights and a more streamlined data estate supporting business operations.

This includes
- AI & Machine Learning
- Business intelligence including data automation and process mining
- Power Platform incl. Power Automate, Power Apps, Power BI & Virtual Agents
- Azure DevOps & ML OPS
- Integration Services & APIs
- Cloud Automation & Governance

Focus on delivering business critical services through a scalable and compliant application landscape in a modern digital work environment.

This includes
- Cloud Architectures & Applications
- 24/7/365 Cloud managed services & support
- Cloud Infrastructure Support
- Cloud Optimization, Automation & Governance
- Data Analytics and visualization
- Migration planning, execution & support

Focus on making cyber security tangible & actionable for the business through advisory on proactive security investments and monitoring of initiatives.

This includes
- Security and risk governance
- Managed detection and response
- Identity and access management
- Rent a Hacker
- Compliance & Information Protection

VENZO_

Bridging tech & advisory with a human perspective

We are pragmatic problem solvers who are passionate about people and organizations yet focused on reducing complexity and driving fast benefits.

- Digital Strategy
- Transformation roadmap
- Data strategy & qualification
- Business process excellence
- Program and project management
- Organizational Change management
- Technology Adoption
- ITSM & Life cycle management
- IT Regulatory & Compliance
- End-of-life assessment and execution

Business

VENZO

Technology
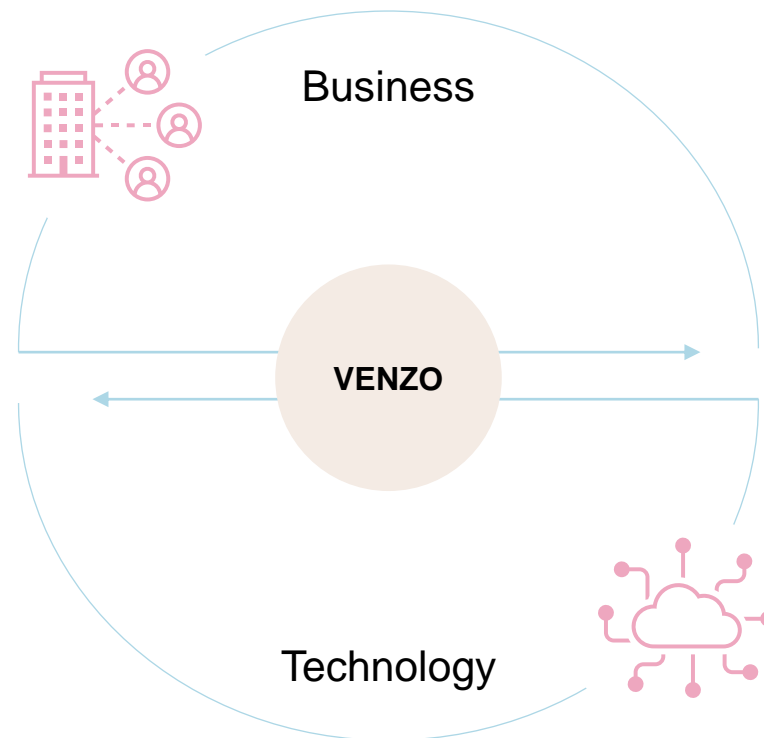
- Application re-factoring
- Agile and Azure DevOps
- Integration and automation
- AI & Machine Learning
- Data reporting & visualization
- Data warehousing & analytics
- Cloud infrastructure
- Incident and problem management
- Operations and monitoring
- Endpoint management

We are a Microsoft-centric technology powerhouse, who specializes within cloud platforms, digital products and modern work.

About VENZO

VENZO_

Trusted Microsoft Specialists for 15 years

Microsoft Partner — 2020 Partner of the Year Winner — Denmark — Microsoft

Microsoft Partner — 2017 Partner of the Year Winner — Denmark — Microsoft

Microsoft Solutions Partner — Data & AI — Azure

Microsoft Solutions Partner — Digital & App Innovation — Azure
Specialist — Low Code Application Development

Microsoft Solutions Partner — Infrastructure — Azure
Specialist — Microsoft Windows Virtual Desktop

Microsoft Solutions Partner — Modern Work
Specialist — Adoption and Change Management — Calling for Microsoft Teams — Modernize Endpoints

Microsoft Solutions Partner — Security
Specialist — Cloud Security — Identity and Access Management — Information Protection and Governance — Threat Protection

Partner Awards

2022
- Security Partner

2021
- Security Partner
- Customer Success Partner

2020
- Modern Workplace Partner

2017
- Modern Desktop Partner
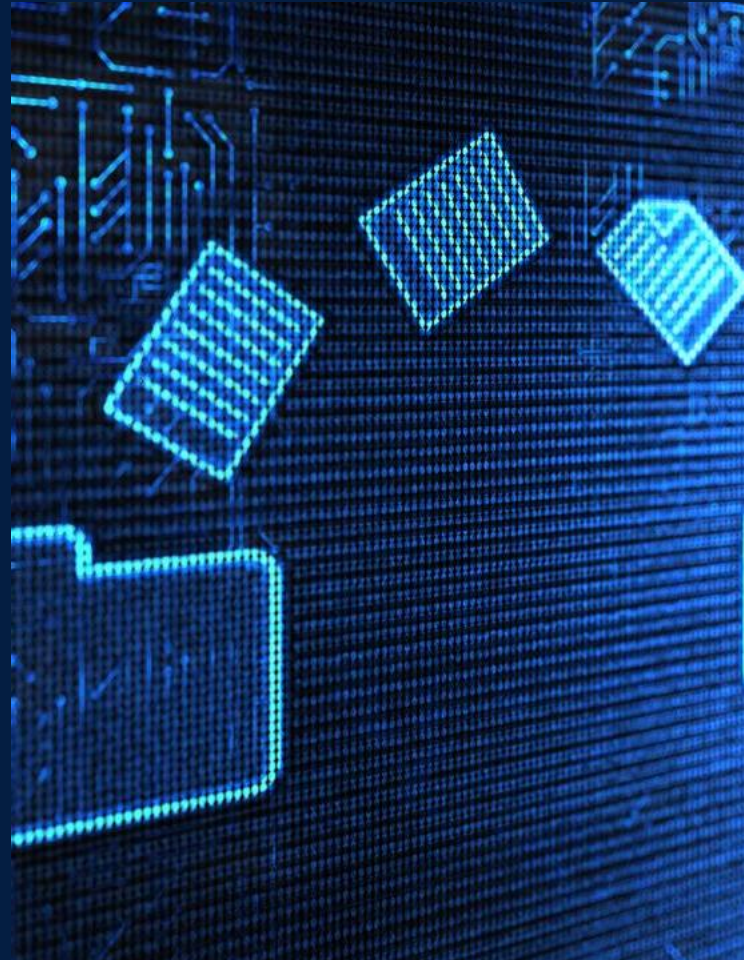
2016
- Modern Platform

Areas of discussion > Outcome

Ingestion & Indexing	Storage & resillience	Seach & Engage

Governance & Tooling

Proof of Concept
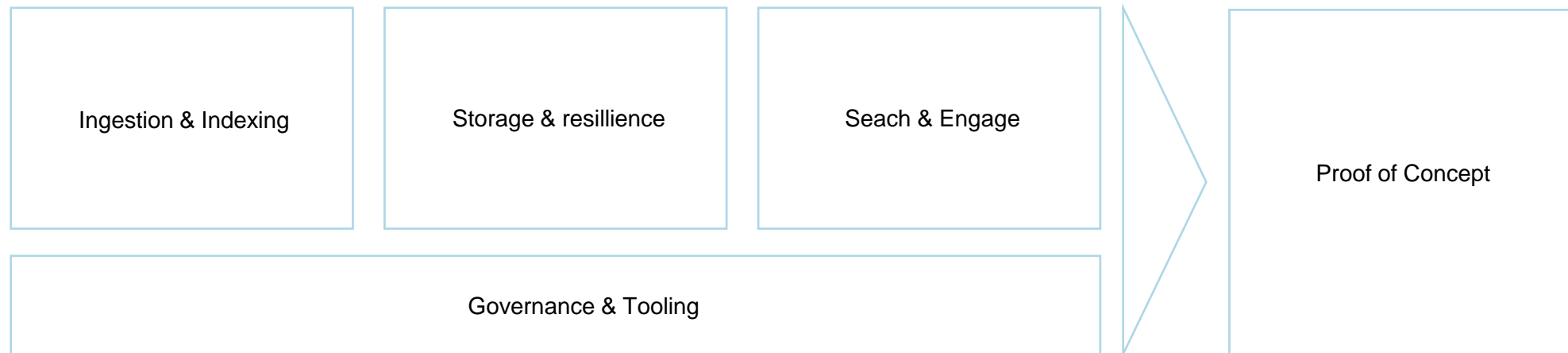
01

Approach and Archive Data Platform

We don't give answers.

Unless we explore questions first.

VENZO_

Example: Data Platform

Analytics Platform

Synapse

Delta Lake

Enrichment of data in each stage for actionable insights, dashboards, and reports of business metrics

Raw → Cleaned → Enterprise

1.1 Data Lake

Analytics
Power BI — Synapse — SQL Database

Data Science / ML Ops
Azure ML — Cognitive Services

1.2 Data Products

Line of Business

CoE

a. Access & Consumption

b. Logging, Governance, Data Catalog & Operations
Log Analytics — Monitor — Purview — Key Vault

1. Analytics Platform

1 Source data moves its raw data layer into the Delta Lake's raw layer on a fixed time interval.

2 The Delta Lake is a parallel computing and storing unit, which will clean and transform the data using Spark Continuous to prepare it for analytical work.

3 Data products are created for the business units. A data product can be anything from a Power-BI report to a machine learning model, performing continuous predictions.

4 To ensure compliance, we implement proper logging and monitoring of all the systems above.
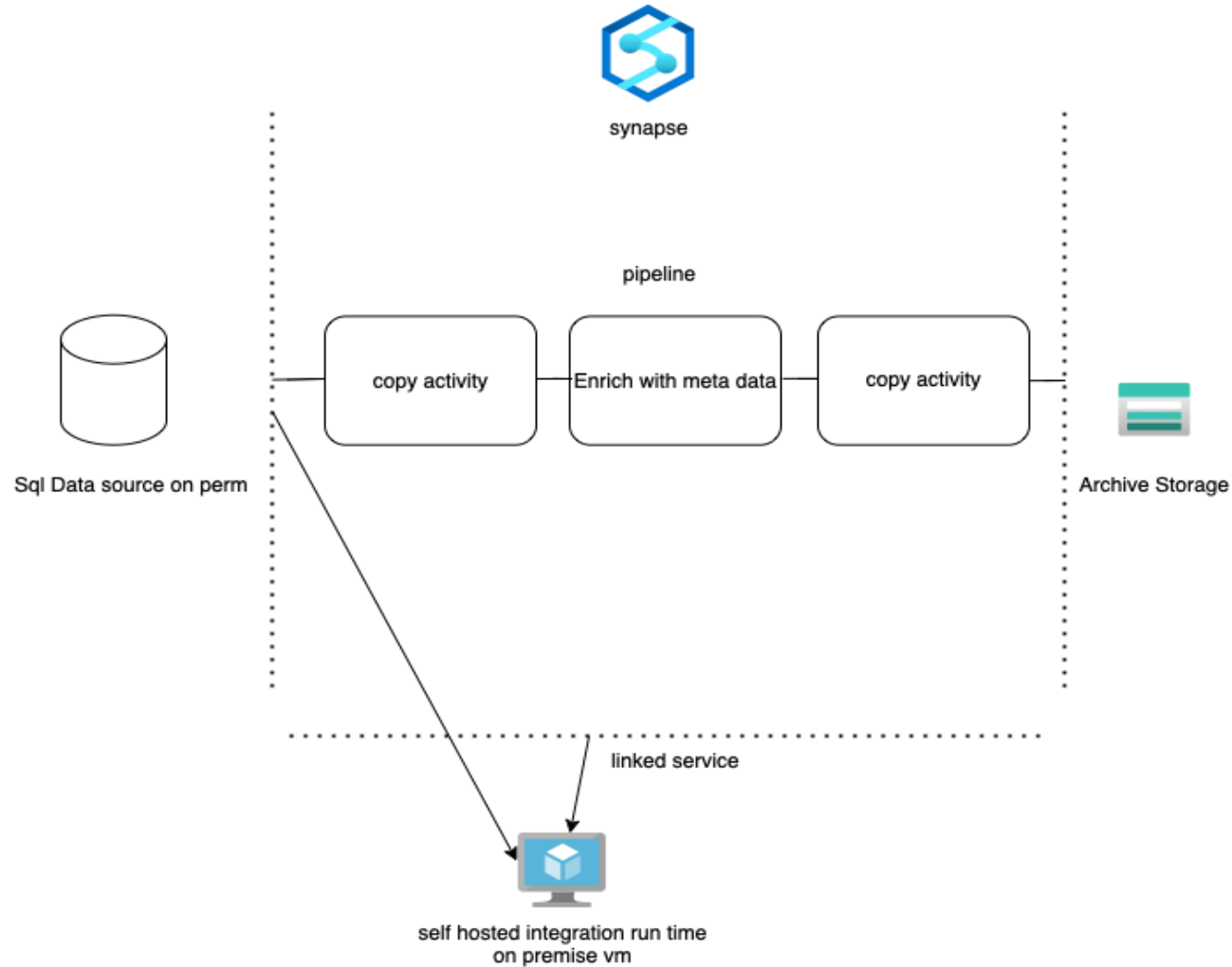Sensitive password and tokens that our services need are handled in the Key Vault.

9

VENZO_

Data Archive: Data Platform

Archive Data Lake

Storage

Compute Power, Data Ingestion

SQL

Data Sources

Logging, Monitoring, Operations
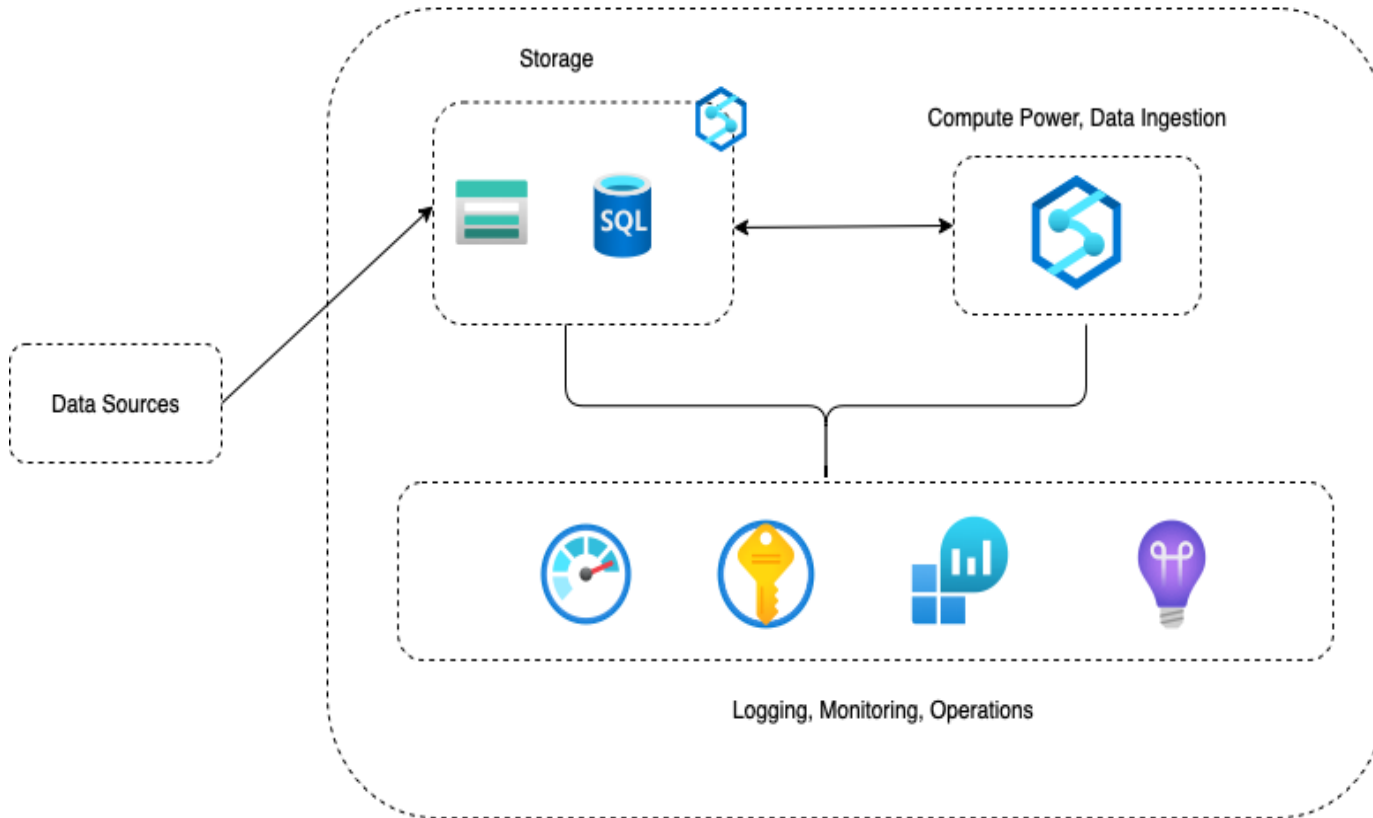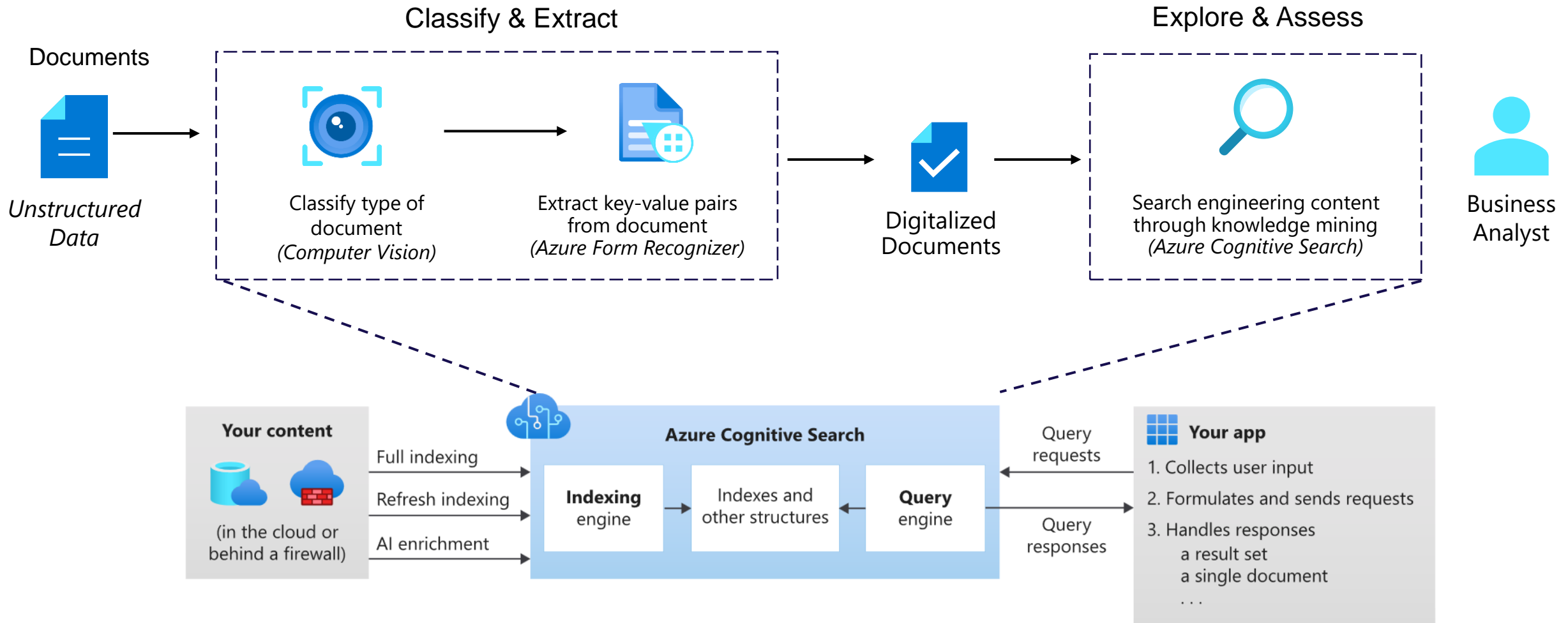
1 Data Immutable to ensure data quality and protection.

2 Parallel computing for data ingestion. Easy ingestion to data sources.

3 Automatic life cycle management of data.

4 Data can be made quarriable with SQL through Azure Synapse

VENZO_

Example: Data Ingestion from Source to Platform

synapse

pipeline

| copy activity | Enrich with meta data | copy activity |

Sql Data source on perm

Archive Storage

linked service

self hosted integration run time
on premise vm

1 On-premises VM to integrate to the on-premises data source.

2 Linked services to conncet to each data source and archive storage.

3 Pipeline made from drag and drop interface. Fast and easy to develop.

4 Data is saved in archive mode for minimal cost.

VENZO_

Automatic indexing of unstructured data

Classify & Extract

Explore & Assess

Documents

Unstructured Data

Classify type of document
(Computer Vision)

Extract key-value pairs from document
(Azure Form Recognizer)

Digitalized Documents

Search engineering content through knowledge mining
(Azure Cognitive Search)

Business Analyst

Your content

(in the cloud or behind a firewall)

Full indexing

Refresh indexing

AI enrichment

Azure Cognitive Search

Indexing engine

Indexes and other structures

Query engine

Query requests

Query responses

Your app

1. Collects user input
2. Formulates and sends requests
3. Handles responses
   a result set
   a single document
   . . .

VENZO_

# Azure Form Recognizer

02

Data governance and security

Secure data, ensure trust - no shortcuts, no compromises.

What is Microsoft Purview Studio?

Microsoft Purview is an Azure resource for admins to gain an overview of data within an organization and for end-users to use this overview for business value purposes. Built on Apache Atlas.

This constitutes **Data Governance** and **Data Adoption**

VENZO_
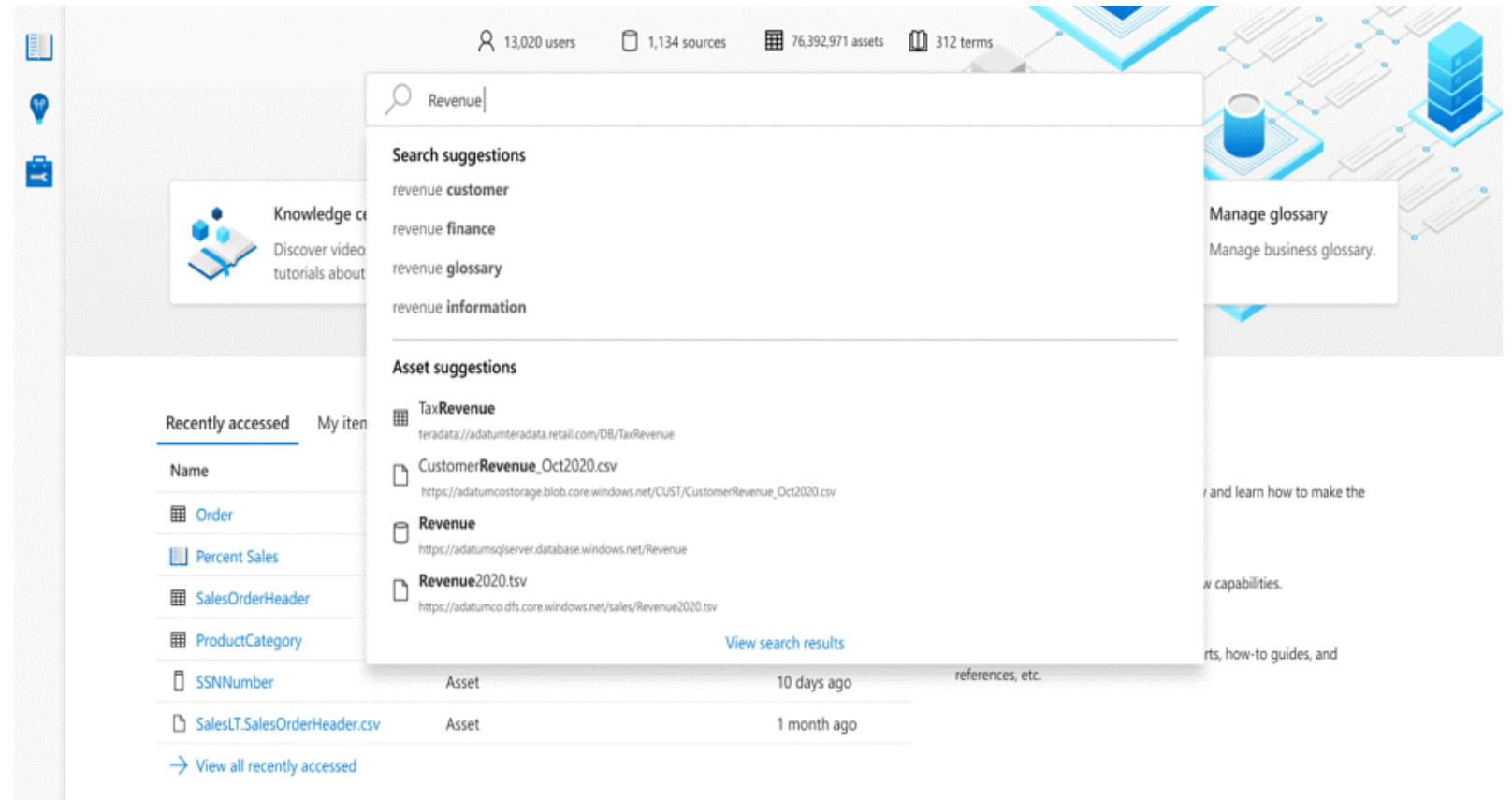
Reference Architecture

Microsoft Purview Studio is hierarchically situated on the top of infrastructure zones or directly peered to them in order to retrieve metadata. A common scenario for organizations in the pharmaceutical sector is to have a non-production and production environment (sometimes with associated GxP / Non-GxP environments). Segregation is possible and multi- or cross tenant functionality is similarly available.

Again, emphasis on the fact that Purview is an overlay to data already existing within the organization.
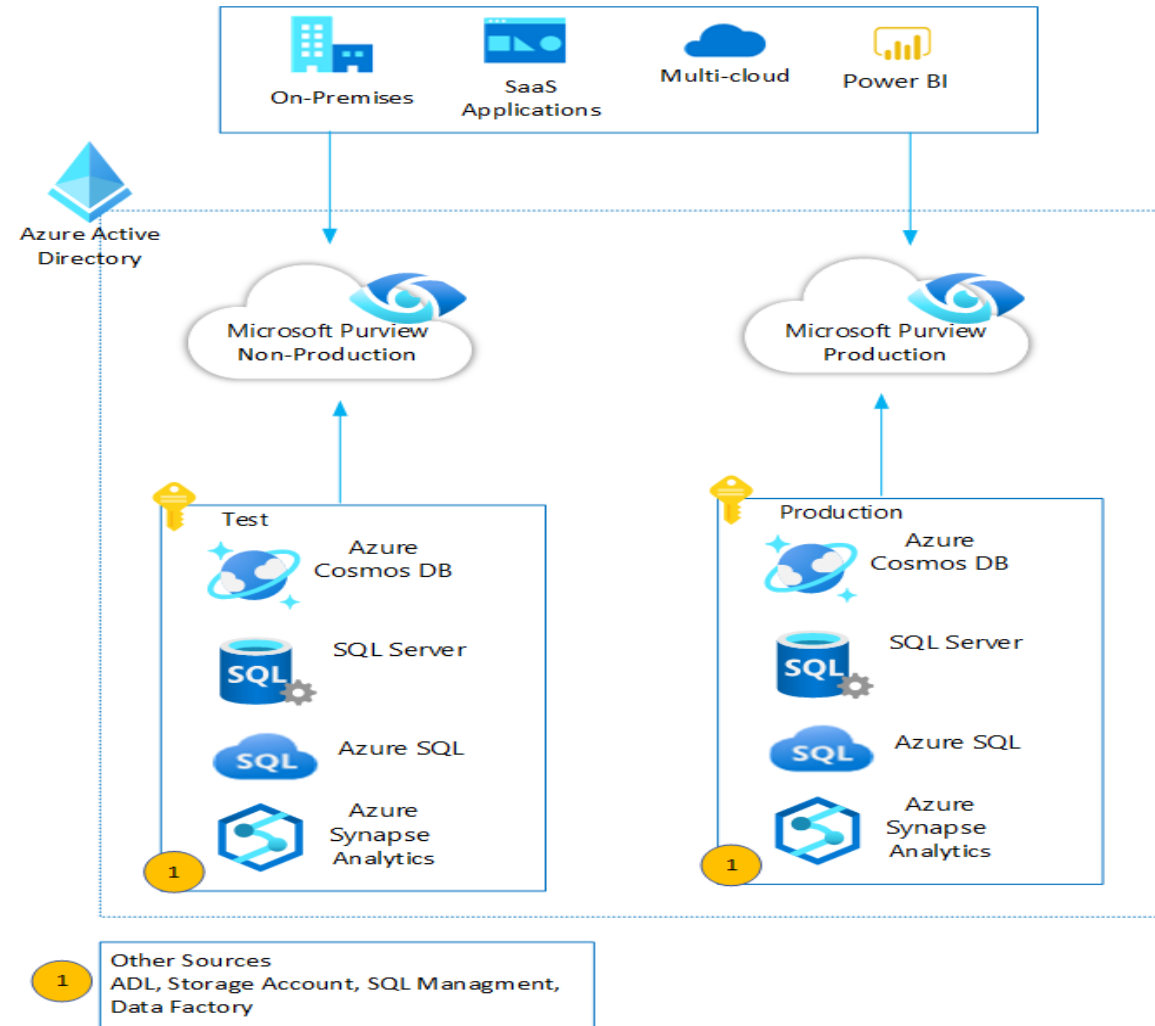
What can we do with Purview Studio?

Data Governance

Automatically apply labels to data
Create custom labels, that are applied upon scan to resources – tagging sensitive and secret data.

Glossary
Create common definitions, synonyms and thesaurus for descriptions surrounding data. Ensures common data understanding.

Data Lineage
See where data lives, where it is transformed and what it is turned into

Data Adoption

Data Catalog and Insights
Discover data across the entire data estate depending on end-user's exact use case.

Self-service access policies
Ask for viewing access to data automatically, with minimal time-to-decision.

Common Data Understanding
Ensure that your datasets follow the agreed upon data schematics

Example: Labeling in the Microsoft Purview Data Map

Automatically apply sensitivity labels to data in the Microsoft Purview Data Map

Create labels → Register asset → Scan asset → Classifications found → Labels applied

Example: Labeling in the Microsoft Purview Data Map

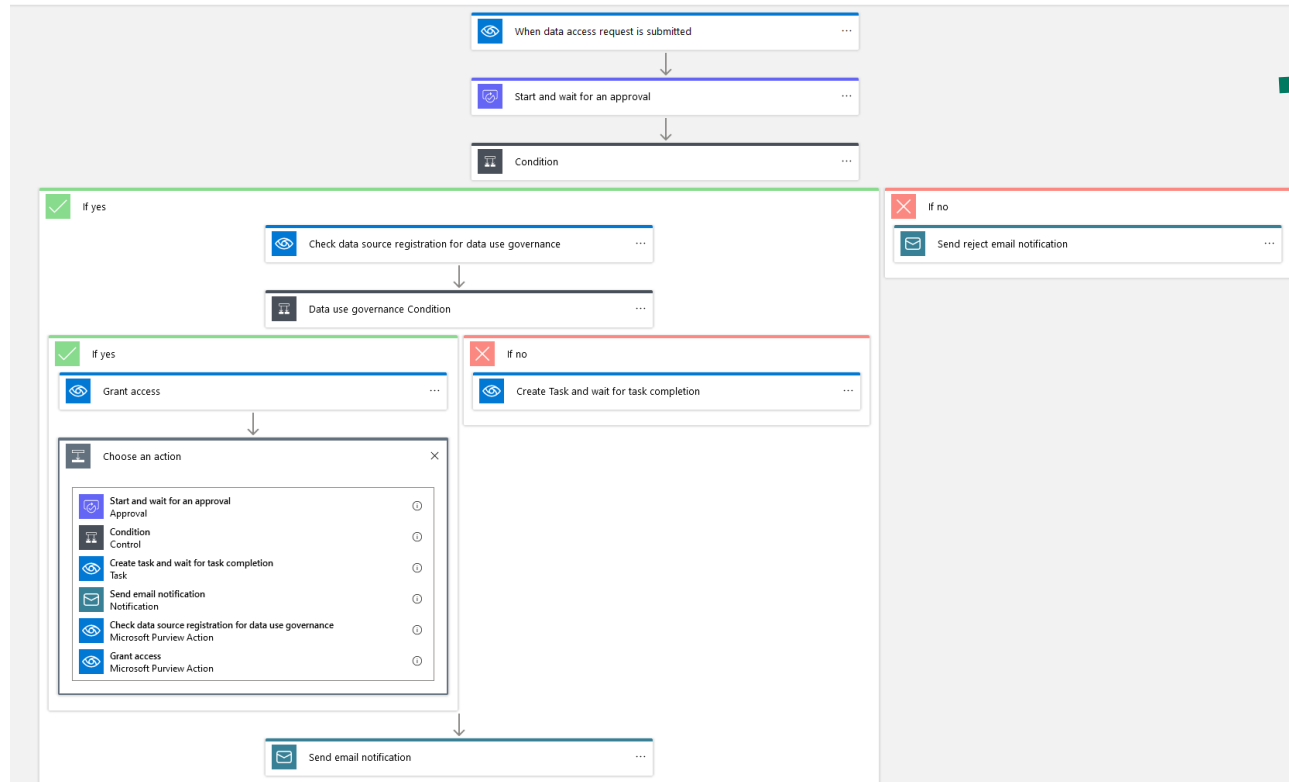Automatically apply sensitivity labels to data in the Microsoft Purview Data Map

Example: Labeling in the Microsoft Purview Data Map

Automatically apply sensitivity labels to data in the Microsoft Purview Data Map

Example: Request access to data sets

Create custom access grant scripts, with minimal time-to-decision.

VENZO_

Securing
Infrastructure

VENZO recommends that organization use the Microsoft best practice framework (Cloud-Adoption Framework) as the guiding principle for cloud-based infrastructure. This best practice inherently focuses on segregation of duties, accesses and infrastructure. Security is thus a built-in principle of the architecture.

Policies

Strict deployment policies should be applied on the supporting infrastructure. This ensures that resources/infrastructure is not able to be deployed without applying to the secure principles that the organization has agreed upon.

Least Privilege Principle

When resources are deployed, they are deployed with the minimal required privilege needed to fulfill the task at hand. Additionally, managed identities that support resource should be assigned directly to resources in a limited privilege scope.

Policies that are non-compliant can be audited, reworked and exempt if necessary. *However, nothing is left to guessing.*

VENZO_
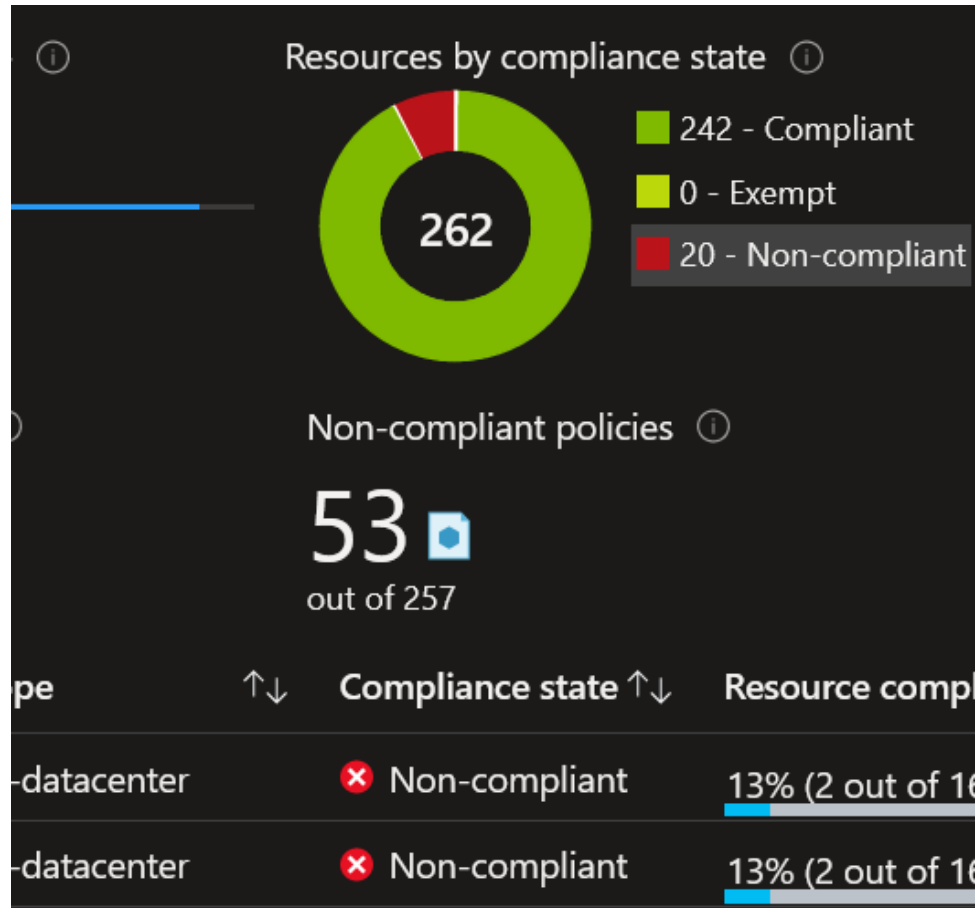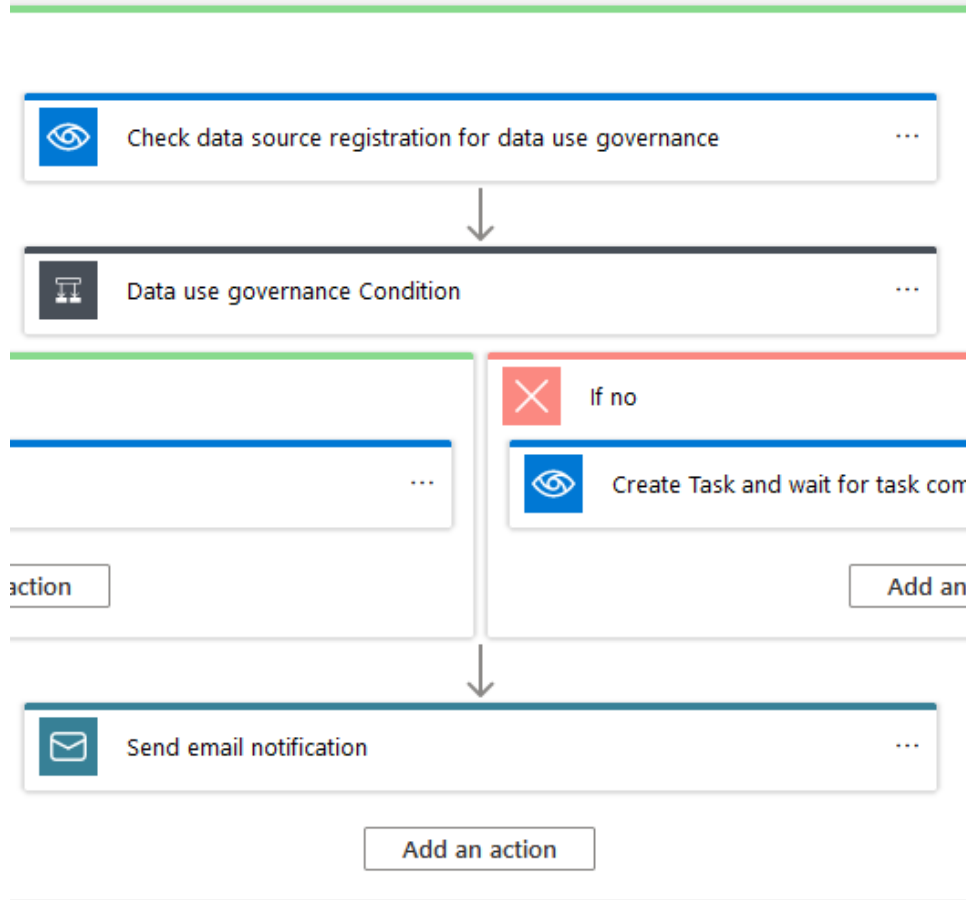
Securing Access

Data access in Purview Studio seeks to combine easy-of-use and reduced time-to-resolve, whilst ensuring that authorization and authentication processes are not left behind.

Authorization (Granting access)

Smart Authorization is enabled by creating automatic grant workflows on data and infrastructure (as shown in Azure Purview demo). Authorization structure follows native Azure Role-Based Access Controls (RBAC).

Authentication (Securing the access grant)

Validate users' access. When a user tries to log into a system, the system needs to verify that the person is who he or she claims to be.

All transmissions during authentication are encrypted to prevent theft of the authenticating information. MFA is likewise enforced upon authentication attempts.

VENZO_

ProductSalesReport (Secret)
🗋 Power BI Dataset
https://app.powerbi.com/groups/00000000-1234-1234-1234-000000000000/datasets/123456

ProductSalesReport (Secret)
🗋 Power BI Dataset
https://app.powerbi.com/groups/00000000-1234-1234-1234-000000000000/datasets/123456

ProductSalesReport (Secret)
📊 Power BI Report
https://app.powerbi.com/groups/00000000-1234-1234-1234-000000000000/datasets/123456

Users.csv (Secret)
🗋 Amazon S3 File
s3://hr-department/Users.csv

ProductSalesReport (Secret)
📊 Power BI Report
https://app.powerbi.com/groups/00000000-1234-1234-1234-000000000000/datasets/123456

NewHire_DATA.json (Secret)
🗋 Azure Blob Path
https://purviewportosblob.blob.core.windows.net/files/userdata/NewHire_DATA.json

Securing Data

Data Security within Microsoft Purview Studio seeks to ensure data encryption becomes a standard and that data accessibility is limited using said encryption. Sensitive data should be treated as such and by combining capabilities, we ensure that this is possible.

Labelling

Automatically labelling based on custom rules allow for any data types (either by pattern or dictionary) to become marked as "Secret", "Sensitive" etc. This automation assists in automatically assessing one's data estate – even at scale.

Data Masking and Data Encryption

Transparent data encryption encrypts databases, backups, and logs at rest without any changes to your application. Without access to the databases, data cannot be read unencrypted.

Data Masking offers the possibility of segregating specific tables, rows or columns of data to select users and/or roles. This can become necessary to enforce sensitivity labels.

Monitoring & auditing

We create infrastructure to host data products and data models. However, we also want to ensure that when data is extracted, applied, transferred or changed, that we know what happened, and where it moved to. Additionally, we want to ensure that if unwanted data sharing occurs, that it be blocked, and alerts be raised.

Where is our data? Where did it move?

Using Data Lineage, we are able to see where Data moves (within scope of scanned infrastructure), when it leaves the realm of Azure into excel reports and the likes, we want to extend the functionality by securing it similarly with retention and sensitivity labels.

Who accessed what, when and why?

With active monitoring and sensitivity labels, we can create a direct correlation between logs and actions.
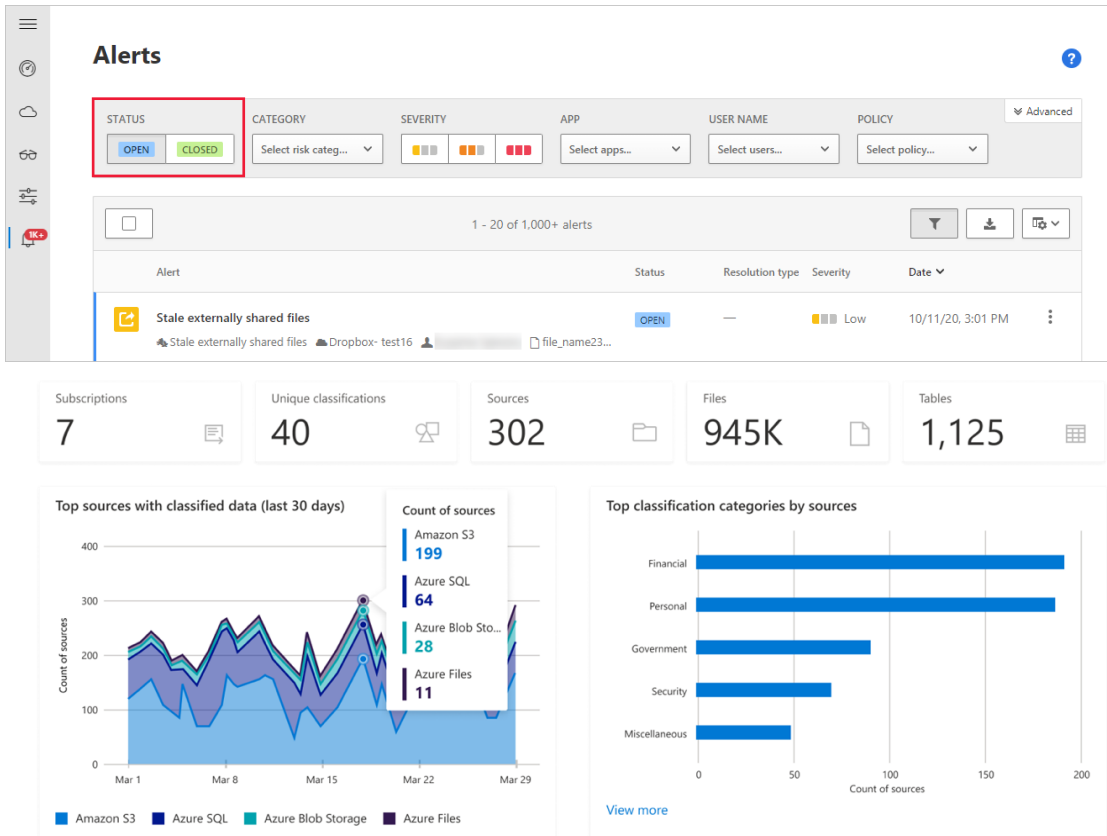
Wrap up & Next steps

- **Q & A / Discussion**

- **Definition & scope of Pilot/PoC**

- **Dependencies**

- **Timelines**

VENZO assisted accelerating the data-journey of a transforming pharmaceutical company

Background

The fully integrated bio-technology company was on a journey from going from a limited number of customers products to multiple of each.

Hence, this resulted in increased need for data, more structure, governance and transparency in order to fully utilize data & insights as a strong value-driver. The company was looking for a partner who could strategize, plan and deliver on a data & architecture roadmap, while ensuring data literacy and adoption in Line of Business.

VENZO's role

VENZO designed a roadmap for the customer, with clear direction in a formalized and adopted data strategy. To this, VENZO rolled out a plan for implementing a data quality & governance structure and proposed a future data architecture design, mapped to selected use cases.

With a three-step plan, VENZO discovered, designed and developed a plan for how to realize and embark on the data journey.

Output

✓ A data strategy, including a governance framework and operating model with roles & responsibilities to be aligned with DnA teams and Line of Business (with a hybrid model suggested)

✓ Target architecture and co-creation of fully GxP qualified data platform

✓ MVP designed and delivered within first month of partnership

✓ First data products incl. BI reports for selected line of business and GxP data product

VENZO_

Eating M&A's for breakfast: VENZO getting data under control 1 month of newly acquired companies

Need

Our Logistics client executed on a M&A-strategy to expand their business geographically. With a new company acquired on a monthly basis, The client needed to manage the data complexity and report on financial and logistical data across acquired companies.

The client wanted to build an efficient model for integrating data into a data warehouse, so management quickly could compare performance across all companies in their portfolio.

Process

1. Clarification of required insights

2. Goal-setting for the Data-model

3. Design of delivery model and deployment pipelines

4. Implementation of delivery model an and hand-over to daily operations

Deliverables

VENZO delivered

- Team of data engineers and analysts to handle integration of data
- Agile delivery designed for accelerated execution
- Standardized end-to-end integration process; from the deal is signed to the
- Continuous development of data warehouse to meet changing reporting demand

Results and value add

✓ Integration of new source systems shortened from 6 to 1 month

✓ Data model optimized to report more accurately on financial results across acquired companies

✓ Several integration activities automated

VENZO assisted global infrastructure company in optimizing their logistic forecasting

Need

The CFO of the global infrastructure company struggled to get a full overview of train utilization including kilometers driven, vehicle capacity, and goods transported. The overview is instrumental for the financial forecast at the company and hence used on both an operational and strategic level.

Process

1. Prototype development

2. Scoping workshop

3. System integration.

4. Agile development practice with 3 iterations in 6 weeks with prototyping and redefinition.
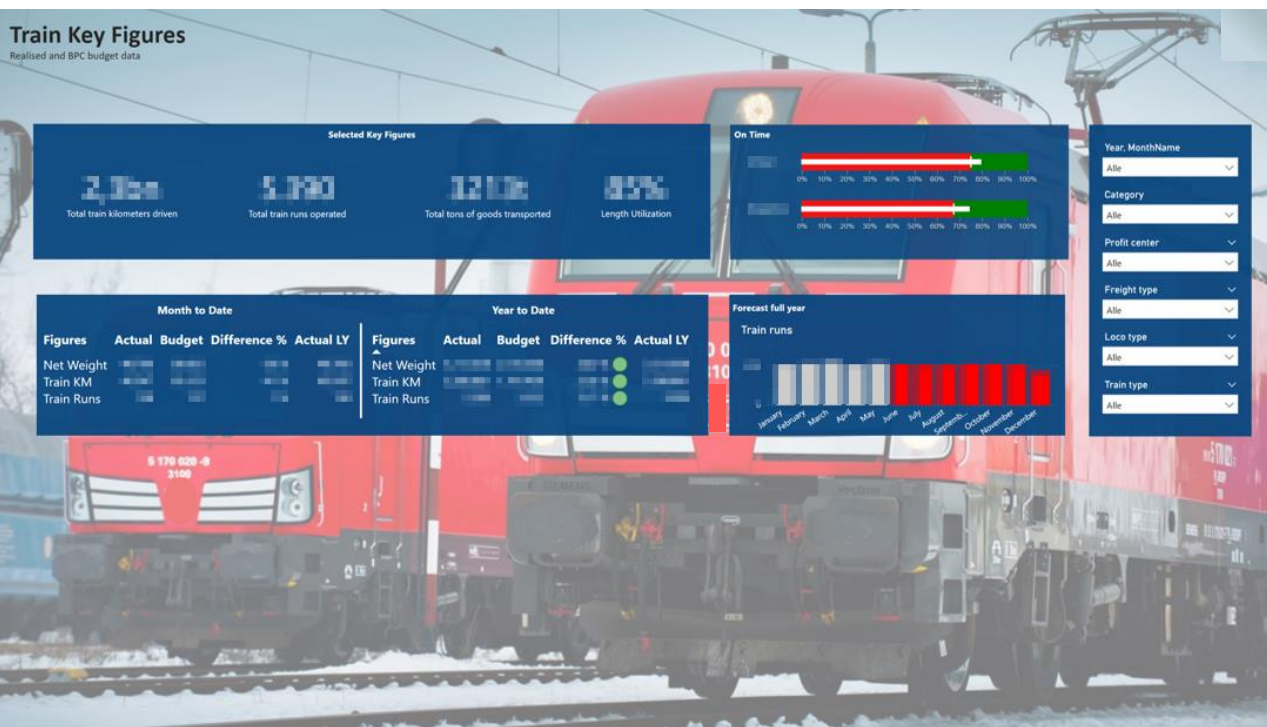
Deliverables

VENZO delivered

- Data-integration from 5 systems to one source of truth.

- Dashboard views for daily operation and strategic optimization.

Results and value add

✓ Better forecasting: Complete data driven overview.

✓ Better variance analysis: Ability to find root-causes through data.

✓ Optimize train utilization