



Get your coffee ready – we'll start in a few minutes



Welcome to the webinar!



Today's host: Oliver Kaae Vinter



End-to-end Data Estate Lineage with Purview and CluedIn

- Welcome
- Unified Data Governance with Purview by Srikanth Venkat
- Moving along the Data Estate by Tim Ward
- Outro + Q&A

Incl. demo

• Unified Data Governance with Azure Purview

- **Srikanth Venkat**

- *Principal PM Manager*
- *Azure Purview Team*

Agenda

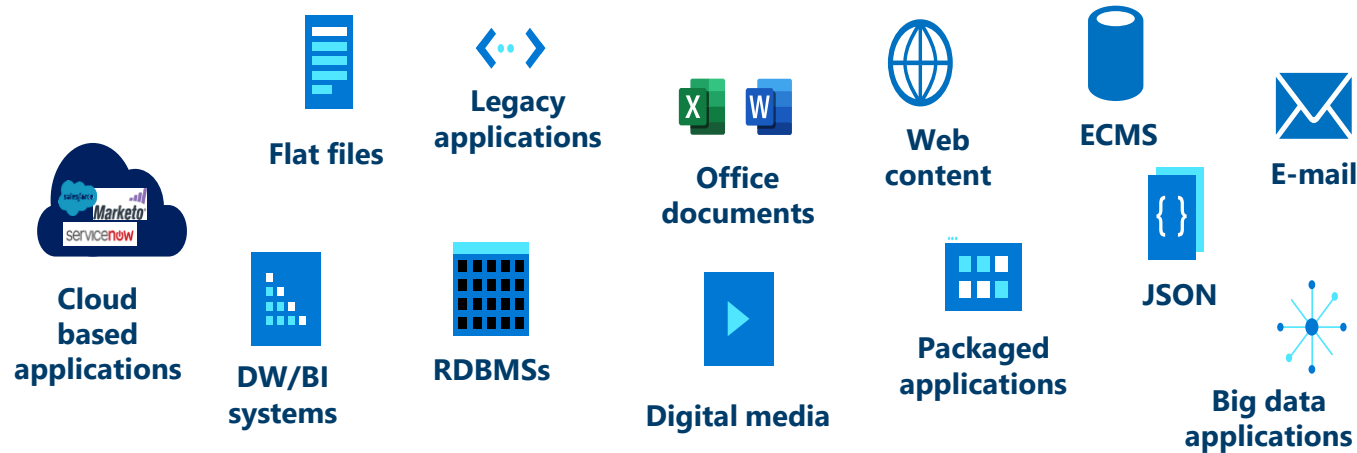
Data Governance Challenges & CDO

Azure Purview Overview

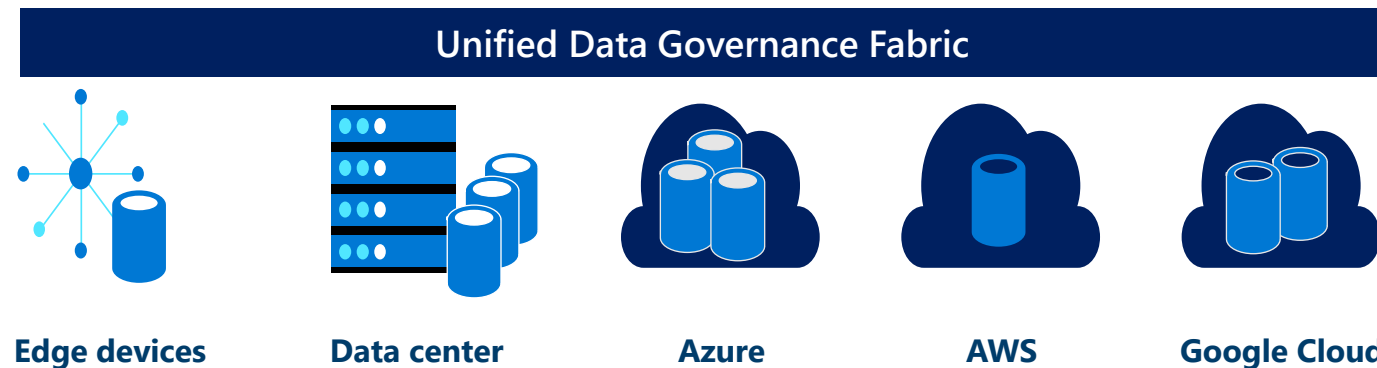
Demo

Data Governance Challenges Across Complex Data Estate

Our customers and Microsoft are managing data in increasingly complex data landscapes



- **92% of organisations have a multi-cloud and hybrid cloud strategy**
- **The number of solution choices is increasing with over 15000 SaaS products available in 2021**
- **Unstructured content is even more fractured, and 80-90% of organizations' data is estimated to be unstructured**

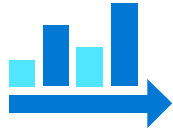


How to balance digital transformation needs and increasing demand with appropriate guardrails around data usage?

Data governance is becoming increasingly interdisciplinary



Elements of successful data governance



Manage growing
data landscape



Overcome
operational silos



Increase
data agility



Comply with
industry regulations



Cloud native, hyperscale, hybrid cloud aware data governance fabric



Extensible platform for data governance with rich open ecosystem of partners and community



Self-service data discovery for knowledge workers & strategic insights for CDO, CPO, & other
Security & Compliance officer personas



Current Status

Partner

Preview

GA

AZURE PURVIEW – UNIFIED DATA GOVERNANCE

APPS

Data Producers, Consumers, Data Engineers, and SMEs

Chief Data Officer

Data Catalog

Business glossary
Data Catalog

Data Sharing

Intra-
Inter-

 CluedIn
Master Data

Ref. data management
Master record management

Data Use Governance

Data policy, Insights
Access governance
Loss prevention

Publish, discover and curate

 Apache Atlas

PLATFORM

Data Map = Data assets | Lineage | Classifications | Business context | ...

Automated Scanning & Classification

Database Platforms (Operational & Analytical)



Files/Object Stores/Data Lakes



Business Apps



Multi-cloud



Data Catalog for Self-Service by Knowledge Workers

Self-service Search & Browse

- Search for data using business terms, technical terms or data classification
- Browse assets by business LOB hierarchy and more
- Embedded search within Microsoft data systems (e.g. Synapse workspace)
- Open-in experience for catalog assets (e.g. Open In Power BI Desktop)

Curated & Standardized Business Glossaries

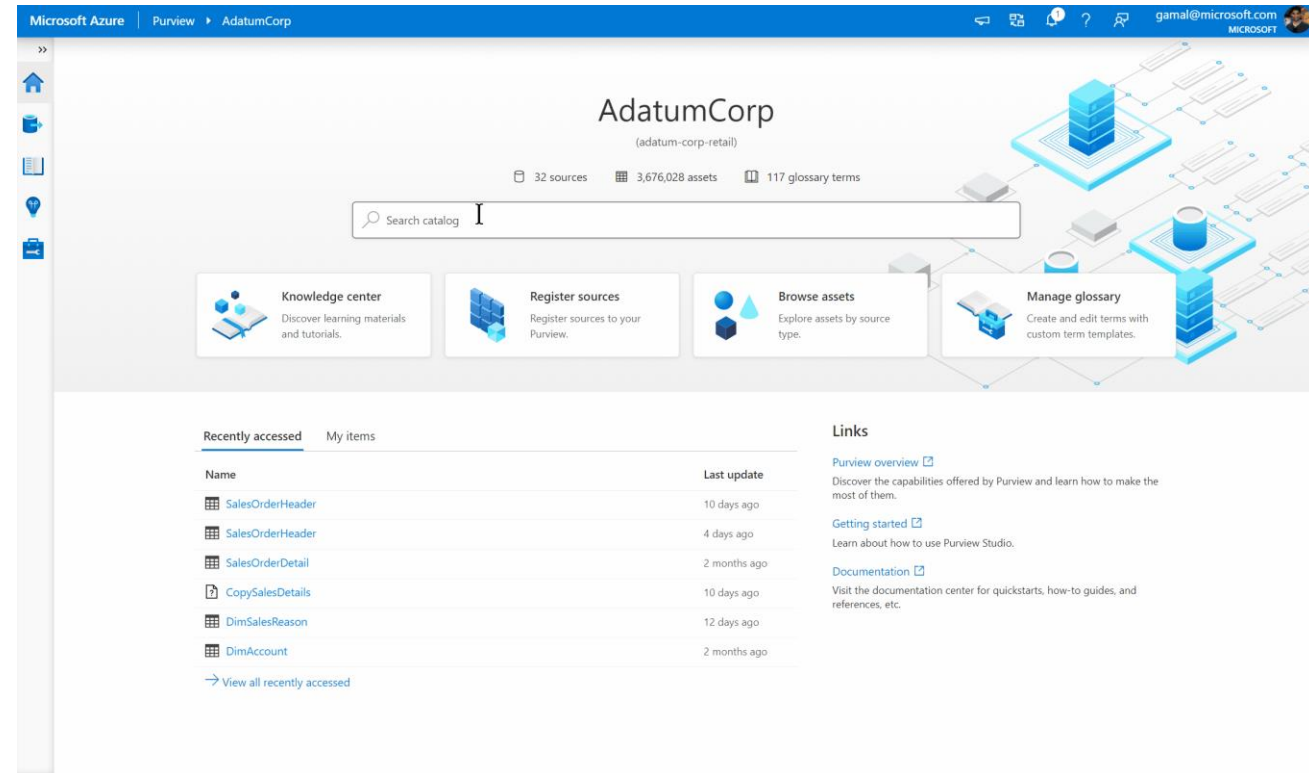
- Define business terms, their definitions, related terms, owners & stewards
- Bulk import existing data dictionaries
- Bulk assign glossary terms to assets
- Hierarchical business glossaries

Interactive Lineage Visualization

- Interactive lineage visualization from raw data through ETL processing to BI
- Object and column level lineage to do root cause analysis, impact analysis

Simplified Data Curation and Stewardship

- Interactive bulk curation to assign ownership, add terms and more
- Approval workflows & business rules (e.g., apply business term based on data classification) coming soon



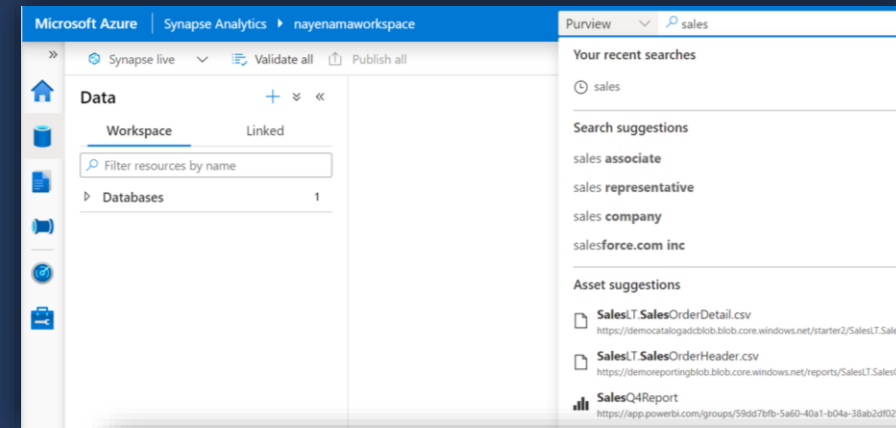
Azure Purview + Synapse

Data Discovery & Governance

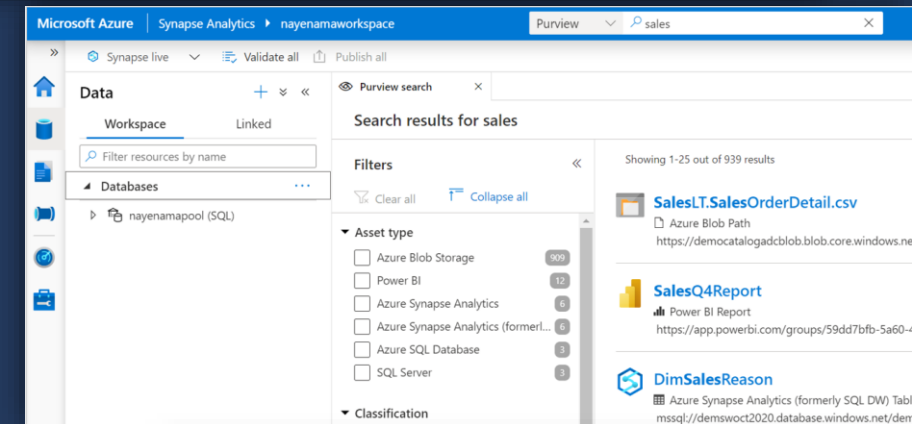
Data Engineers and Analysts can search the Purview data catalog **directly** from Synapse Studio and act on new data immediately

- Automatically discover data across multi-cloud, SaaS, and on-premises, classify and catalog for easy discoverability
- Within Synapse Studio, query data using SQL or Spark, or launch the code-free data flow designer to integrate the data from disparate sources
- Automated lineage extraction from Azure Data Factory pipelines and data flows (including lineage from SSIS packages)

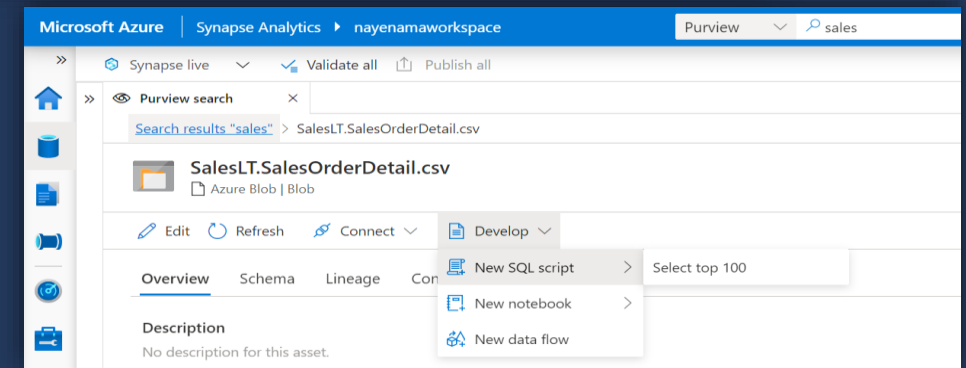
Embedded
Search



Search
Results



Analyze



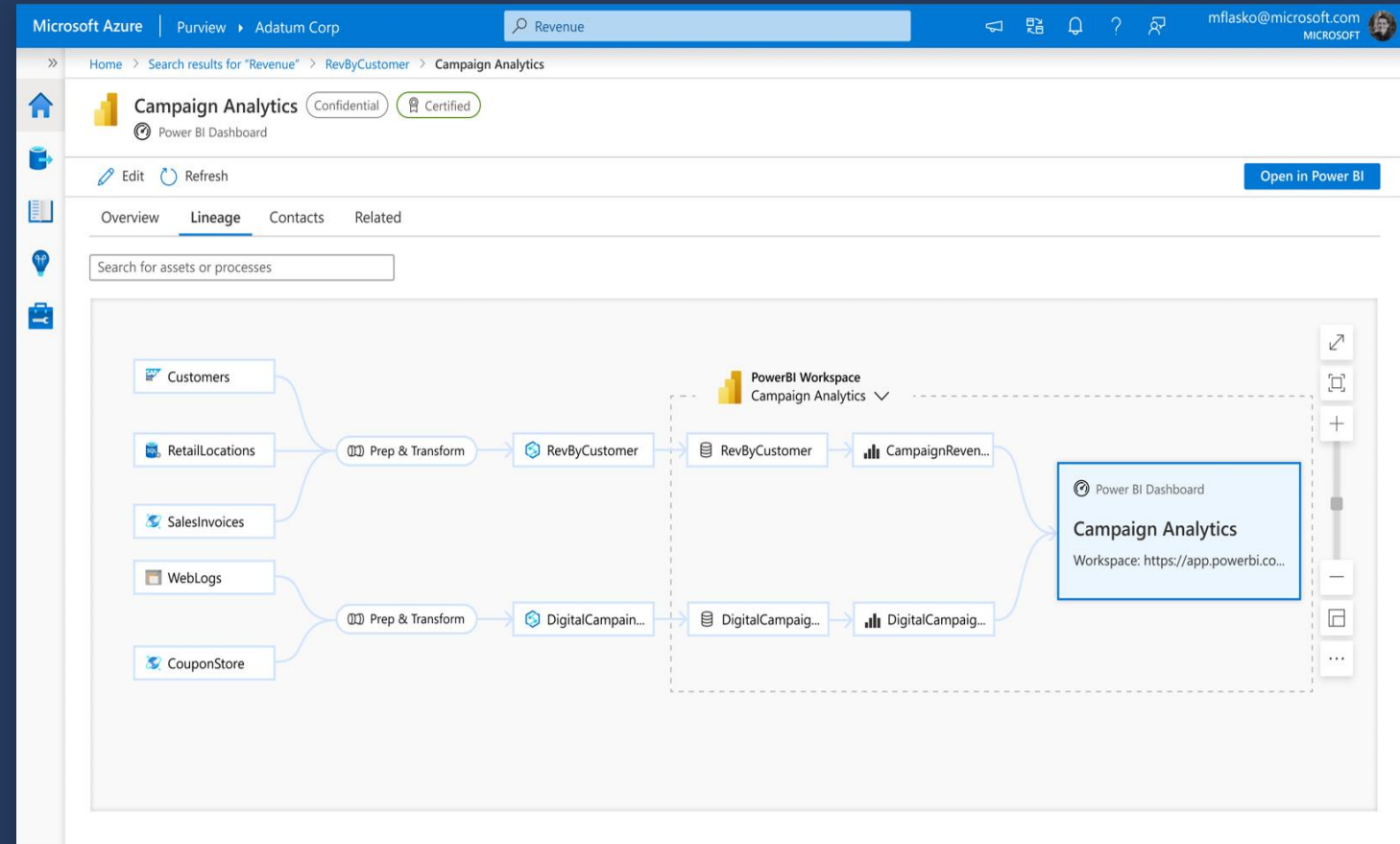
Azure Purview + Power BI

Data Discovery & Governance

Discover Power BI assets in Purview Data Map

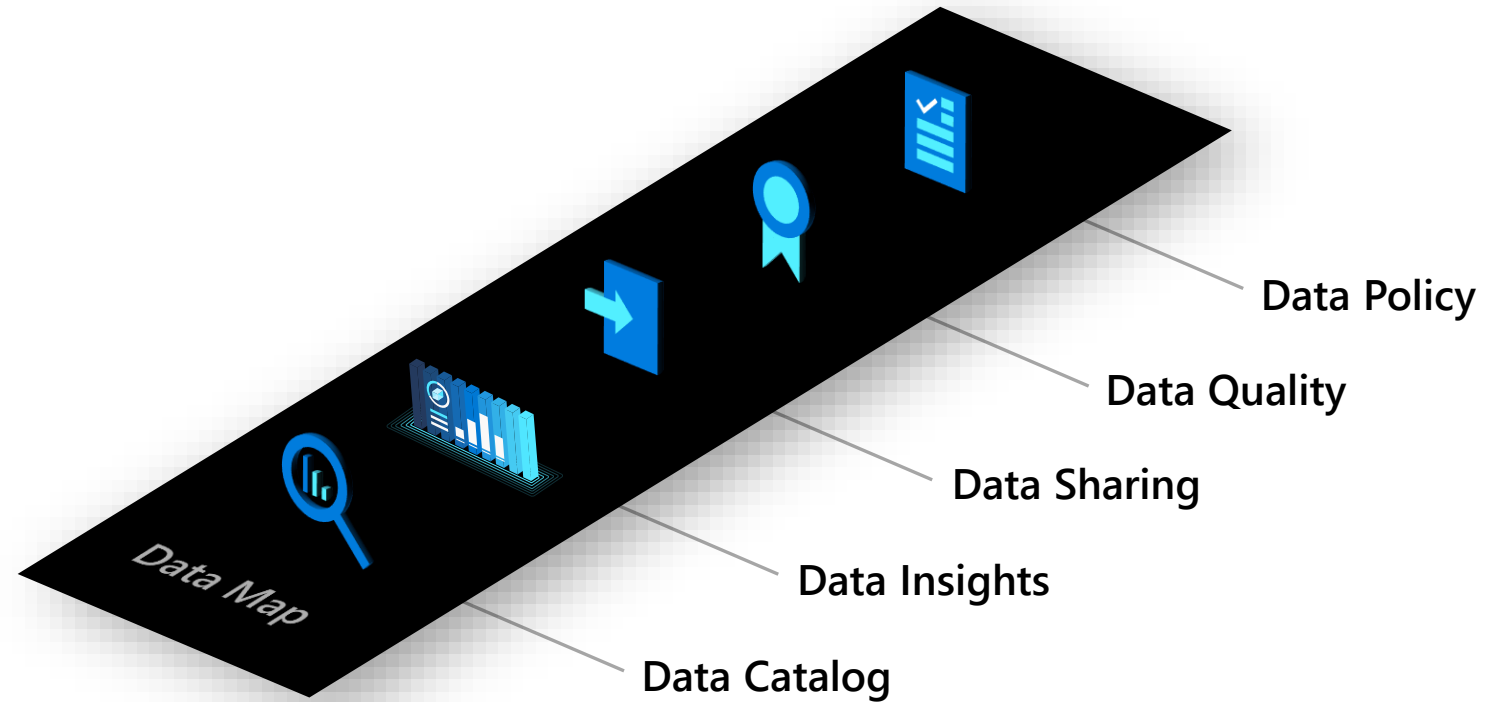
Effortless discovery of data and PBI assets in Purview Data Catalog

View enhanced metadata, lineage, sensitivity labels, endorsement labels for Power BI assets



Unified Data Governance with Azure Purview

Roadmap



Demo



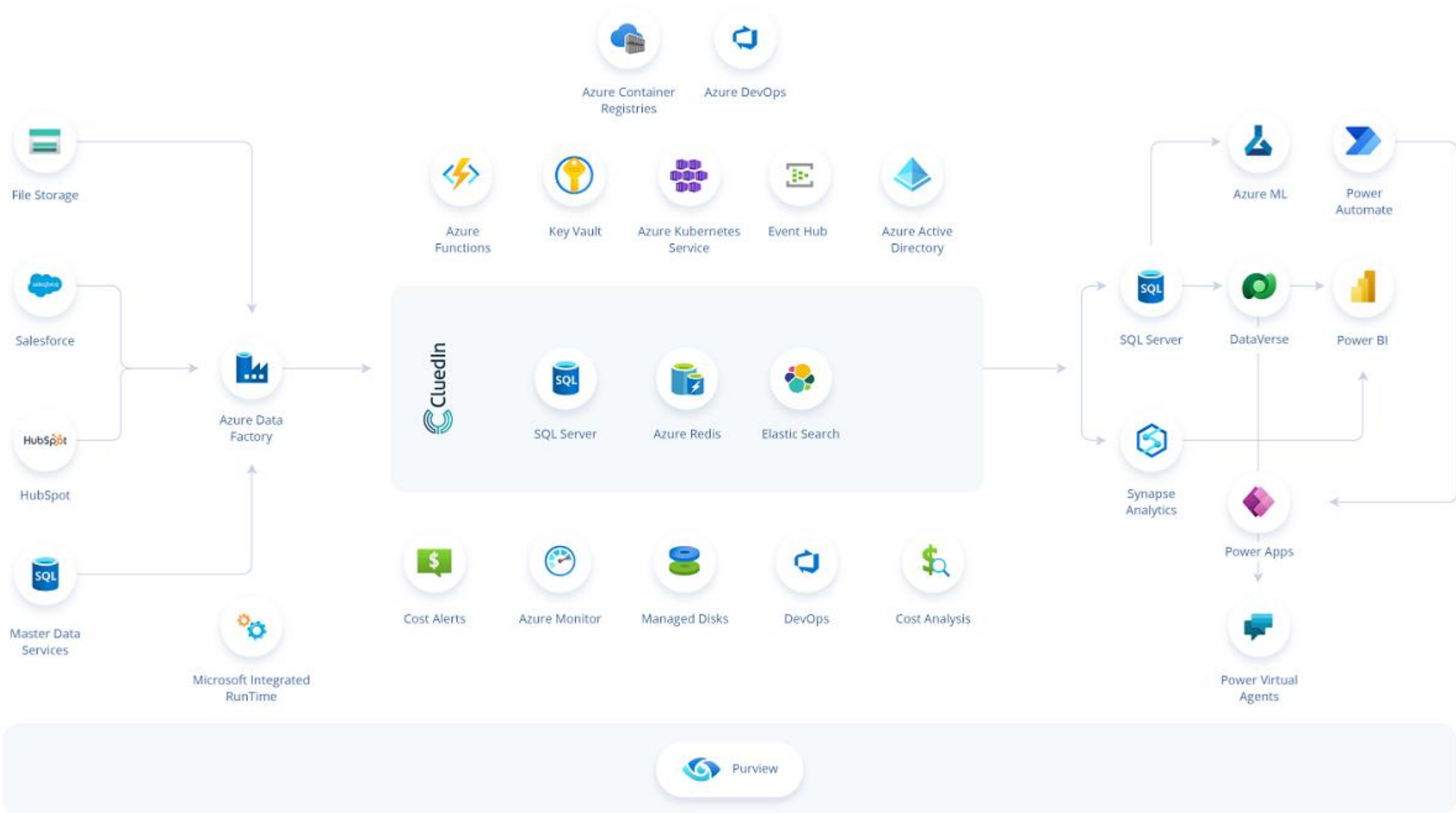
Moving along the Data Estate

By Tim

Moving along the Data Estate

Let's set the scene

- We now have a central repository of all our registered systems, datasets, glossary terms.
- We have Azure Data Factory as a delivering mechanism of data.
- The data that we want to provide insights from is not in a state to be used, even after it has gone through our Azure Synapse / Azure Databricks Pipelines.
- Let's make this data ready for insight.



End-to-End Data Lineage

- Azure Purview provides Data Asset Level Governance
- CluedIn provides record level Governance
- Your cloud environment is bigger than just Azure. Something needs to provide a mechanism for visibility across cloud providers and on-premises.

Use case demonstration

- How can we bring data from Purview together from across multiple sources, combine them, clean them, enrich them and share out to the business, all tracked and governed with Azure Purview.
- On its journey, we want full data lineage and explain-ability so that any consumer knows exactly where the data came from and how it was transformed on its journey.

Outro

Create CluedIn Master Data Management ...

Event Hub - Please select the options for Event Hub integration.

Click this box to learn more.

Event Hub Deployment * ⓘ

Do Nothing

Purview - Please select the options for Purview integration.

Click this box to learn more.

Purview Deployment * ⓘ

Use Existing Instance

Purview - Base URL * ⓘ

Purview - Catalog Base URL * ⓘ

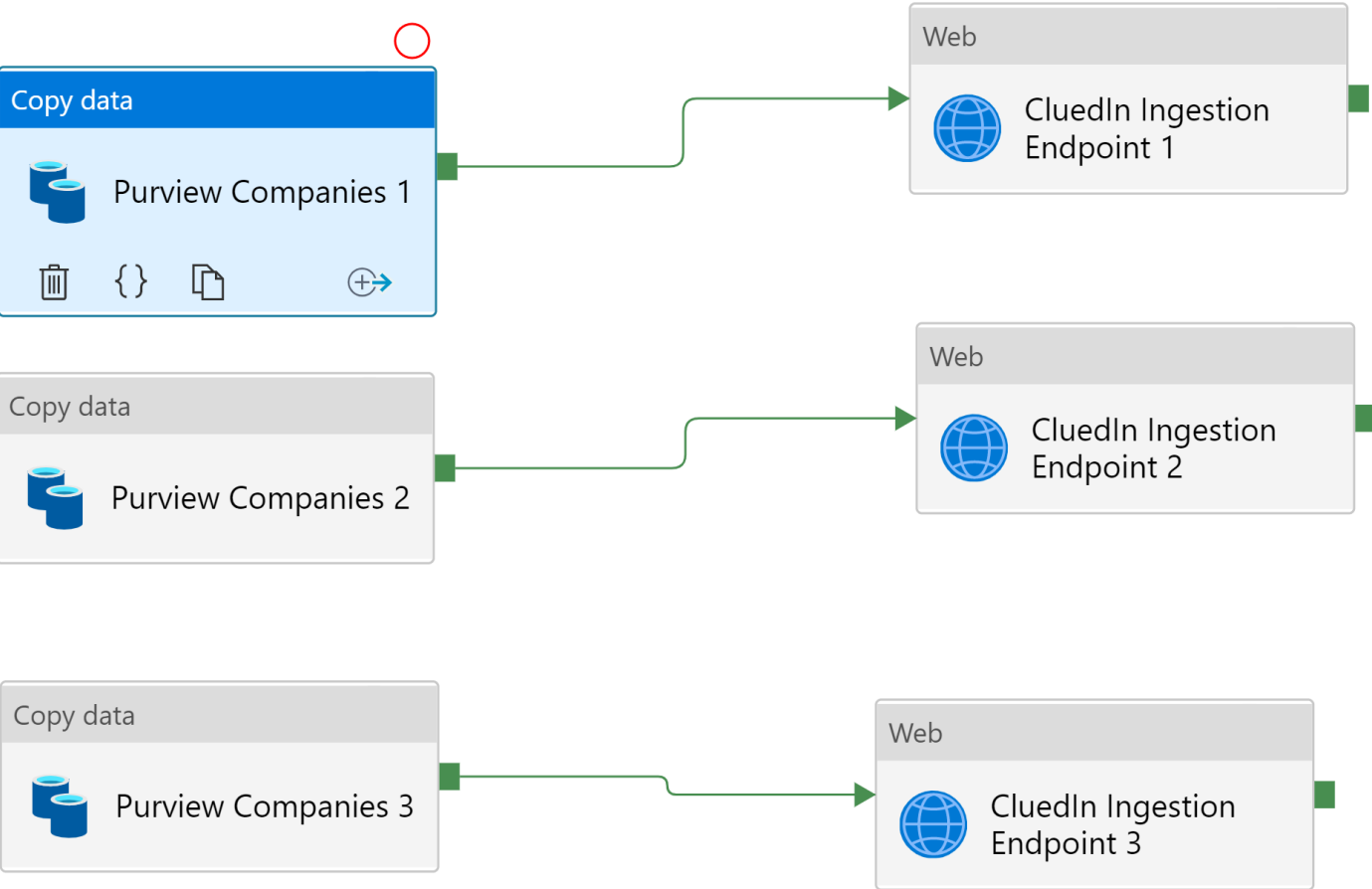
Purview - Scan Base URL * ⓘ

Purview - Client Id * ⓘ

Purview - Tenant Id * ⓘ

Purview - Client Secret * ⓘ

Purview - Confirm Client Secret * ⓘ



CluedIn

Integrations

Governance

Preparation

Engine Room

Management

Entity Types

Data Catalog

Duplicate

Rule Builder

Annotation

Glossary

Consume

Administration

All

Search

All

Clipboard

Refresh

Notifications

Help

User

< Back

Management / ... / Organization / Address City

Address City

Vocabulary key organization.address.city

Edit

Configuration

Usage

All Values

Lineage

Pending Changes

Audit Log

10 k +

Entities

2

Streams

View stream usage

1

Glossary Terms

View glossary usage

2

Rules

View rule usage

© CluedIn 2022

v 3.3.0-alpha.733

 Address CityVocabulary key organization.address.city  Edit

Configuration

Usage

All Values

Lineage

Pending Changes

Audit Log

**INTEGRATIONS** 2**File Data Source**

Some text about File Data Sources

**HubSpot**

HubSpot is all-in-one software for inbound marketing, CRM and Sales.

SOURCE VOCABULARY KEYS 3

gleif.organization.legalAddress+city



hubspot.company.City



mdm.facility.city

CORE VOCABULARY KEYS

organization.address.city

STREAMS 2

Fac



Fac

Create Vocabulary

Common

Reset 1 filter

All status

All Entity types

Name	Entity type	Key Prefix	Source
> Common Data Model Account	Account	commonDataModel.account	No source
> Common Data Model Address	Address	commonDataModel.address	No source
> Common Data Model CarePlan	CarePlan	commonDataModel.careplan	No source
> Common Data Model CodeableConcept	CodeableConcept	commonDataModel.codeableconc...	No source
> Common Data Model Contact	Contact	commonDataModel.contact	No source
> Common Data Model Device	Device	commonDataModel.device	No source
> Common Data Model EmrAppointment	EmrAppointment	commonDataModel.emrappointm...	No source
> Common Data Model Encounter	Encounter	commonDataModel.encounter	No source
> Common Data Model EpisodeOfCare	EpisodeOfCare	commonDataModel.episodeofcare	No source
> Common Data Model Location	Location	commonDataModel.location	No source

CluedIn

Integrations

Governance

Preparation

Engine Room

Management

Consume

GraphQL

Streams

Export Targets

Administration

© CluedIn 2022

v 3.3.0-alpha.733

All

Search

All

Consume / ... / All Streams / Fac

Fac

Created by

A

 admin@foobar.com

Synchronized

Active

Cancel

Save

Configuration

Export Target Configuration

Preview Condition

Data

Monitoring

Pending Changes

Audit Log

Name	organization.address.city
BALANCE RESIDENTIAL CARE CORP II	BAKERSFIELD
BASHAM FUNERAL CARE	BAKERSFIELD
ARCADIA FAMILY CARE 1V	BAKERSFIELD
CLINICA SIERRA VISTA	BAKERSFIELD
BAHAMAS SURGERY CENTER LLC	BAKERSFIELD
ARCADIA GARDENS #4	BAKERSFIELD
BAKERSFIELD SENIOR CENTER	BAKERSFIELD
BAKERSFIELD SENIOR CENTER	BAKERSFIELD
ARCADIA GARDENS #4	BAKERSFIELD
ARCADIA FAMILY CARE 1V	BAKERSFIELD



Created by admin@foobar.com

Synchronized

Active ☒

Cancel

Save

Configuration

Export Target Configuration

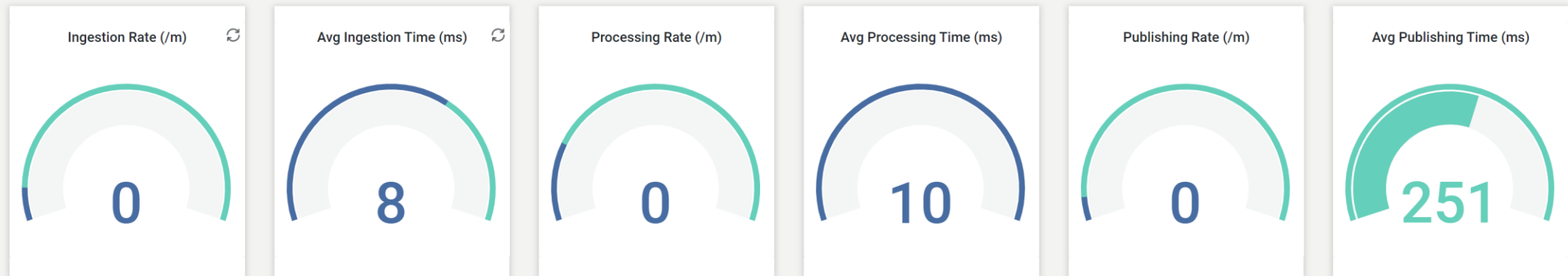
Preview Condition

Data

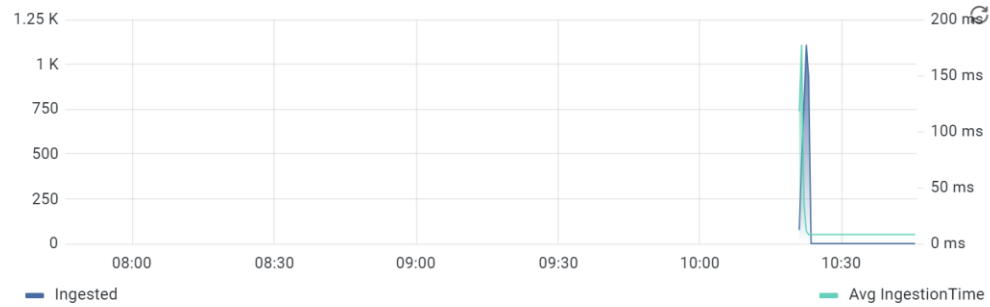
Monitoring

Pending Changes

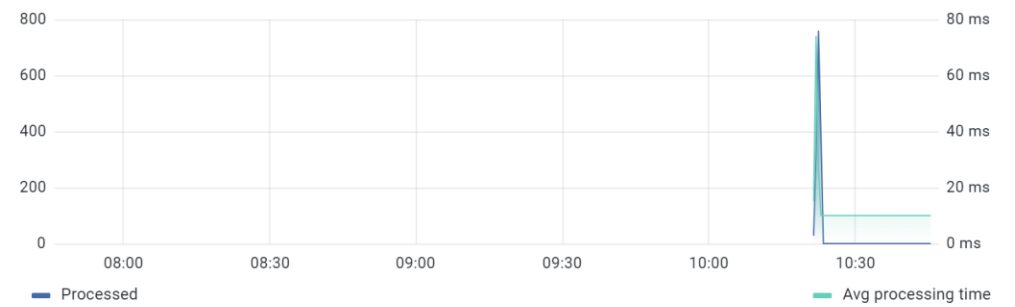
Audit Log



1. Ingestion vs Ingestion Time ?



2. Processed vs Processing Time ?



3. Published vs Publishing Time ?



4. Exceptions ?

Thank you for
your time

cluedin.com





Time for questions! 😊

Additional Resources (Purview)

- Whitepaper: [CDO Seat at the Cloud Table](#)
- Get started with Azure Purview: aka.ms/azurepurview/getstarted
- Microsoft Learn: Introduction to Azure Purview: aka.ms/intro-to-azure-purview
- Azure Purview Documentation: aka.ms/purviewdocs
- Customer Stories: [Microsoft Customer Stories Search](#)
- Cloud Adoption Framework: [Azure Purview for Enterprise Data Governance](#)
- Azure Purview Blog: [Azure Purview Blog - Microsoft Tech Community](#)

Additional resources (CluedIn):

- Company website: [CluedIn - Modern Master Data Management](#)
- CluedIn documentation: [What is CluedIn | CluedIn Documentation](#)
- Combine [CluedIn – University](#) with Azure Data Fundamentals <https://aka.ms/AAd5u07>
- Azure Architecture Contributions:
 - [Migrate master data services to Azure with CluedIn and Azure Purview - Azure Architectures | Microsoft Docs](#)
 - [Master data management with Azure and CluedIn - Azure Architectures | Microsoft Docs](#)

- Extended Free Offers (limited time):
 - Free scanning of on-premises SQL Servers
 - Free scanning of Power BI tenants
- Free Previews:
 - Catalog Insights
 - Data Access Control Policies
 - Self-Service Data Access Request
 - Business Workflows



Thanks for joining! 😊 Feel free to reach out