



# Azure Synapse Analytics

## The unified, cloud-native platform for converged analytics

Orapin Anonthanasap  
Digital Data & AI Specialist

# Agenda

- Introduction to Azure Synapse Analytics
- Explore Azure Synapse Analytics
- Azure Synapse Analytics Resources
- Q&A



# Orapin Anonthanasap (Fon)

Digital Data & AI Specialist

Digital Sales Enterprise Azure Australia

- Former Machine Learning Engineer + Data Analyst
- Data Science/ Machine Learning / AI Lover ❤️

<https://www.linkedin.com/in/orapina/>  
orapina@microsoft.com



**Jirachai Chansivanon (Job)**

Digital Sale Enterprise

[linkedin.com/in/jirachai-c](https://linkedin.com/in/jirachai-c)  
jchansivanon@microsoft.com



**Jirachai Chansivanon (Job)**

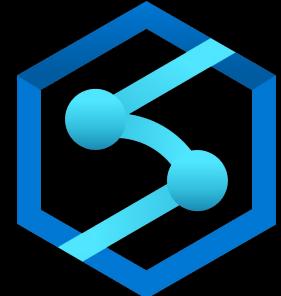
(Acting) Digital App & Infra Specialist



**Orapin Anonthanasap (Fon)**

Digital Data & AI Specialist

[linkedin.com/in/orapina/](https://linkedin.com/in/orapina/)  
orapina@microsoft.com



# Azure Synapse Analytics

The unified, cloud-native platform for converged analytics

# Common customer journey

Big Data and advanced analytics

## Modern data warehouse



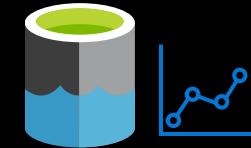
"Integrate all our data—including Big Data—with our data warehouse for analytics and reporting"

## Advanced analytics



"Predict next best offer and customer churn"

## Real-time analytics



"Derive insights from our devices and data streams in real-time"

# Common customer use cases

Modern data warehouse

Advanced analytics

Real-time analytics

## Sources (available to migrate to Azure)

- SQL, MySQL, PostgreSQL
- SAP on Azure
- Oracle to PostgreSQL
- File storage/Backup
- DB2, AS/400 Migration



## Data Analytics Use Cases

- Sales Forecast
- Customer Segmentation
- Customer Lifetime Value
- Churn Prediction
- Recommendation
- Promotion Effectiveness
- Cross-sell/Upsell
- Enterprise Search



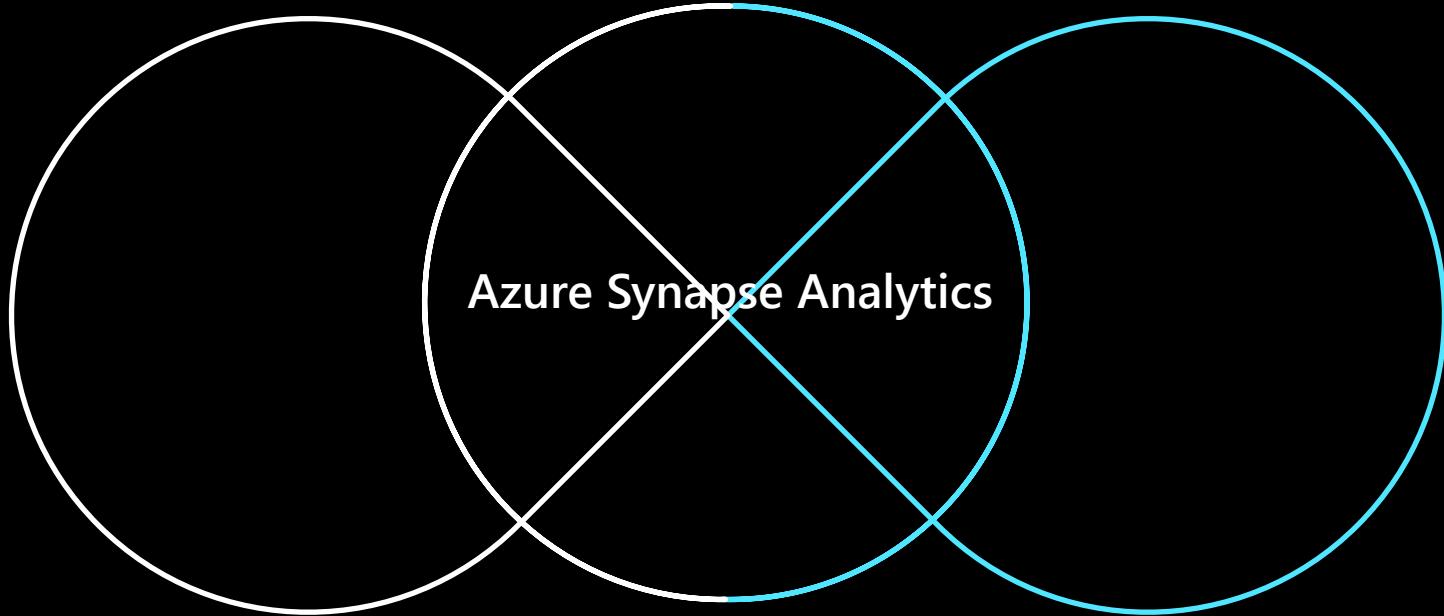
## Real-time Analytics Use Cases

- Dynamic Pricing
- Fraud Detection
- Predictive Maintenance
- Workplace Security (CCTV)
  - Face recognition
  - Security alert
- Digital Twin (simulation)

## Centralized Big Data

- Modern Data Warehouse
- Interactive Dashboard & Self-service BI

Azure brings these two worlds together, in a single service,  
to provide limitless analytics



**Welcome to limitless**  
Data warehousing & big data analytics—all in one service

This is a result of businesses being forced to maintain two critical, yet independent analytics systems

Big Data



Experimentation  
Fast exploration  
Semi-structured data

Relational Data



Proven security & privacy  
Dependable performance  
Operational data

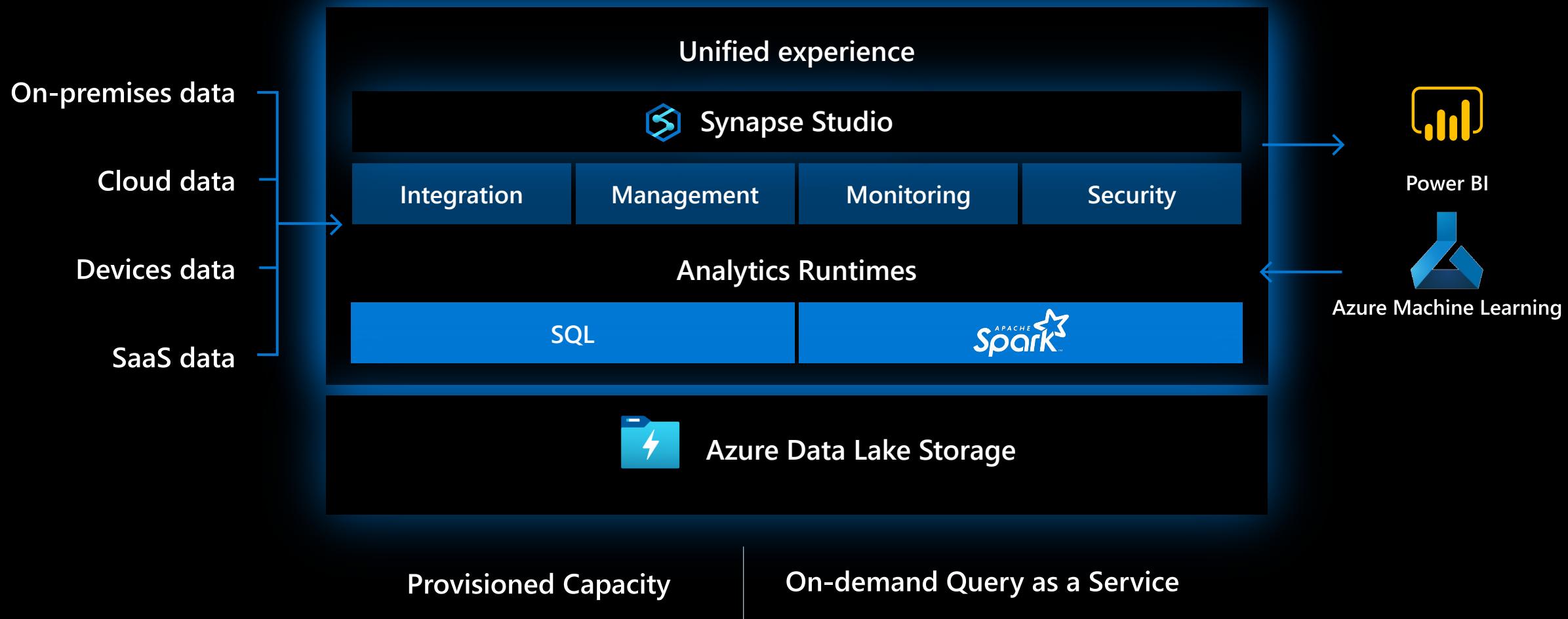
OR

Data Lake

Data Warehouse

# Azure Synapse Analytics

Limitless MPP data warehouse with unmatched time to insights



# Azure Synapse Analytics

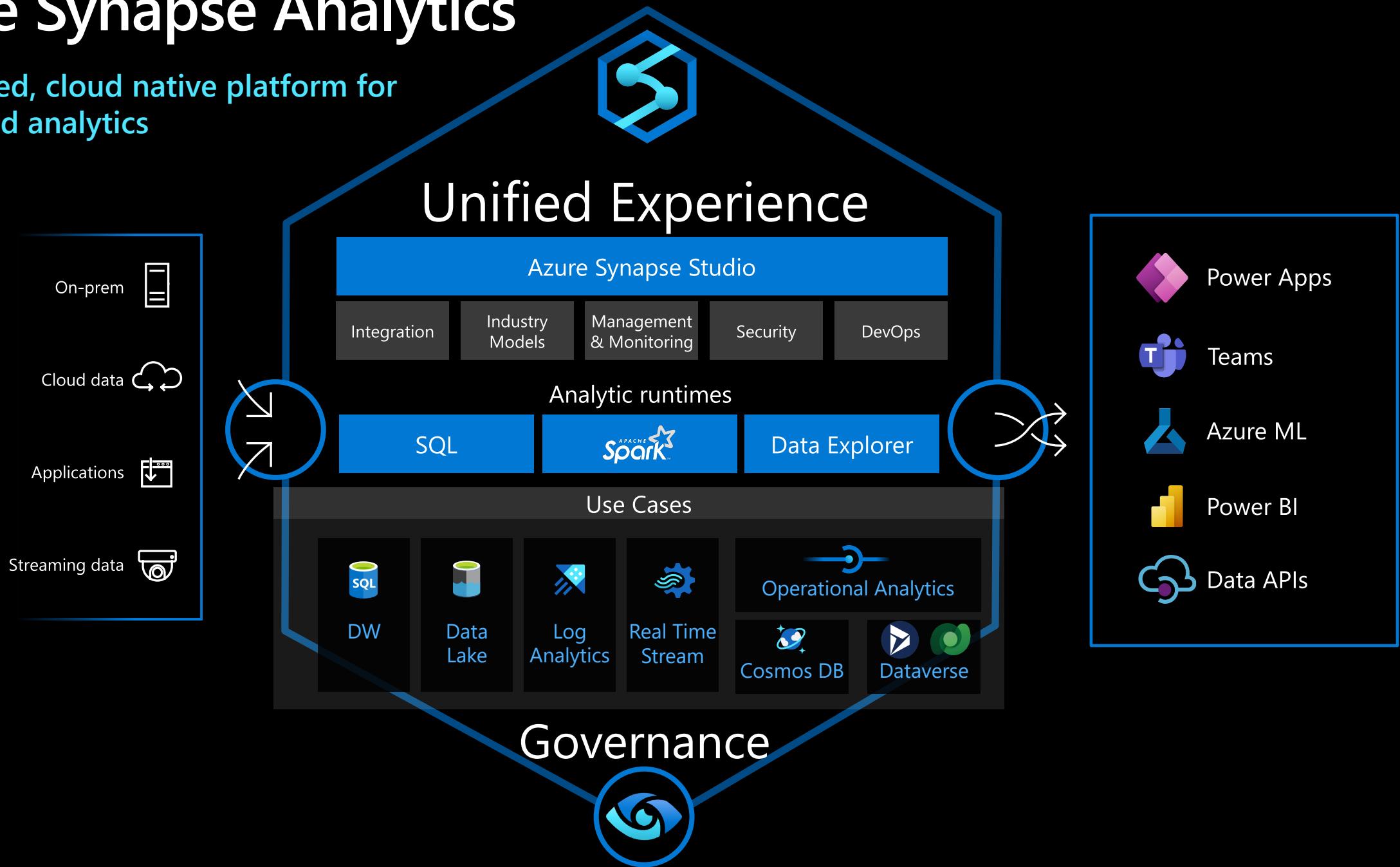
The first unified, cloud native platform for converged analytics



Azure Synapse is the only unified platform for analytics, blending big data, data warehousing, and data integration into a single cloud native service for end-to-end analytics at cloud scale.

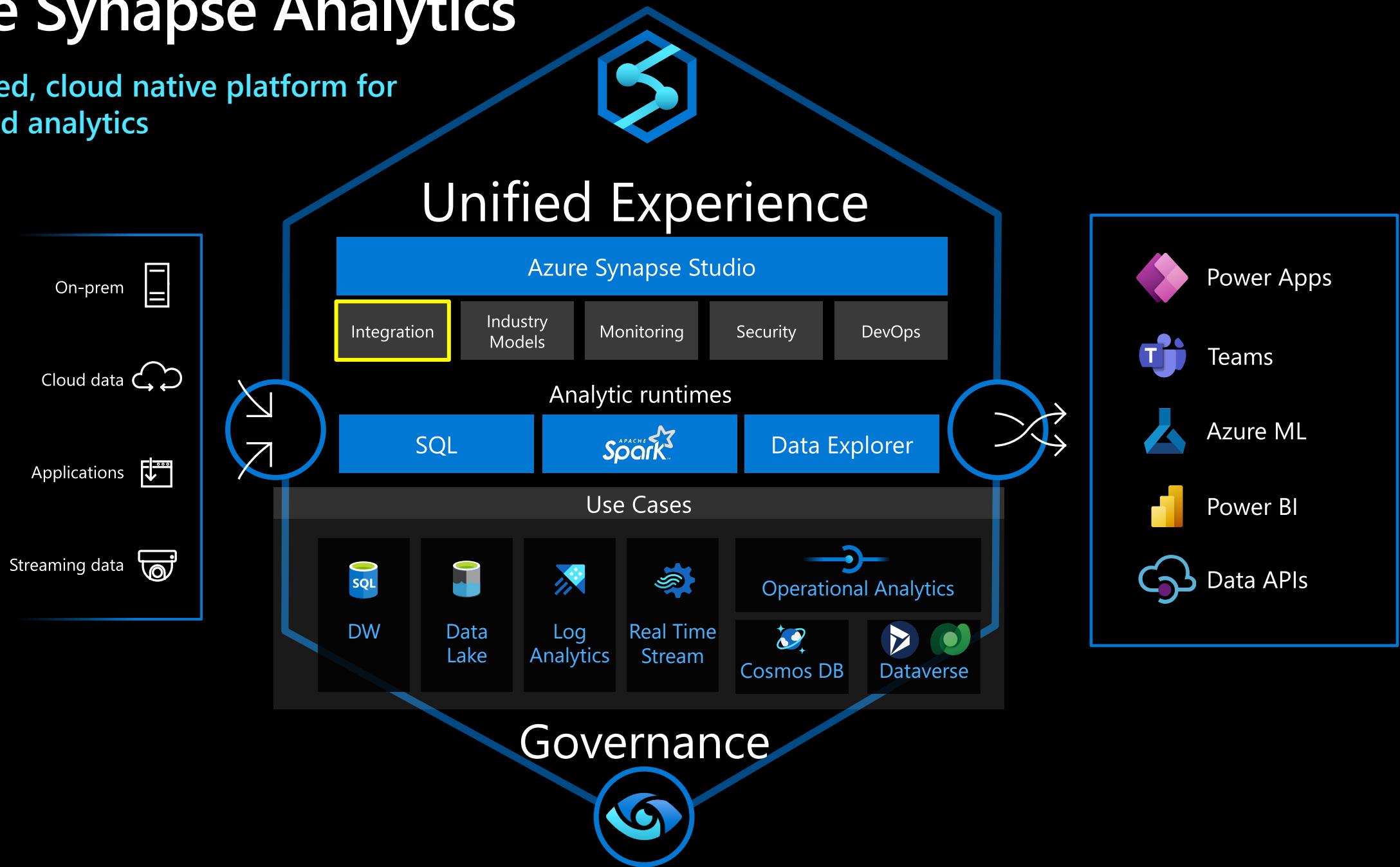
# Azure Synapse Analytics

The unified, cloud native platform for converged analytics



# Azure Synapse Analytics

The unified, cloud native platform for converged analytics



# Hybrid Data Integration

## Cloud native ETL/ELT

95+ connectors available

Secure connectivity to on-premise data sources, other clouds, and SaaS applications

Code-first and low/no code design interfaces

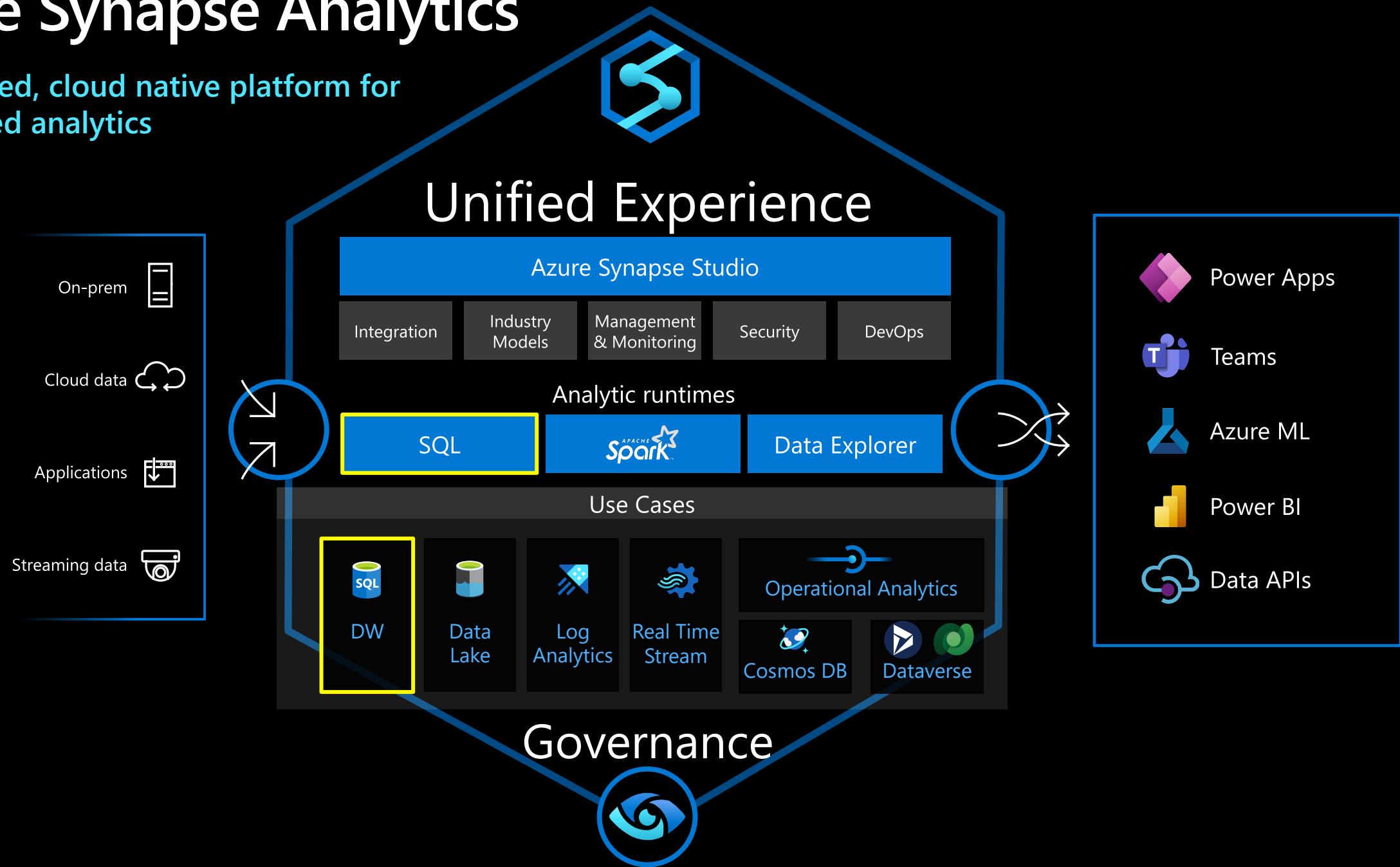
Schedule and Event based triggering



Code-free

# Azure Synapse Analytics

The unified, cloud native platform for converged analytics

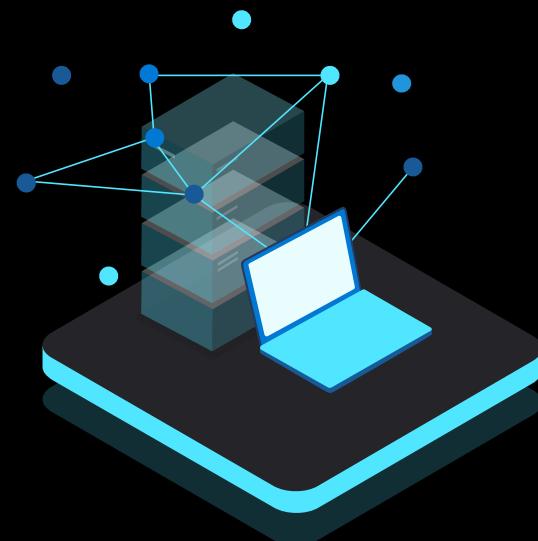


# Dedicated + serverless SQL

## Flexible consumption models

Serverless pay-per-query ideal for ad-hoc data lake exploration and transformation

Dedicated clusters optimized mission-critical data warehouse workloads



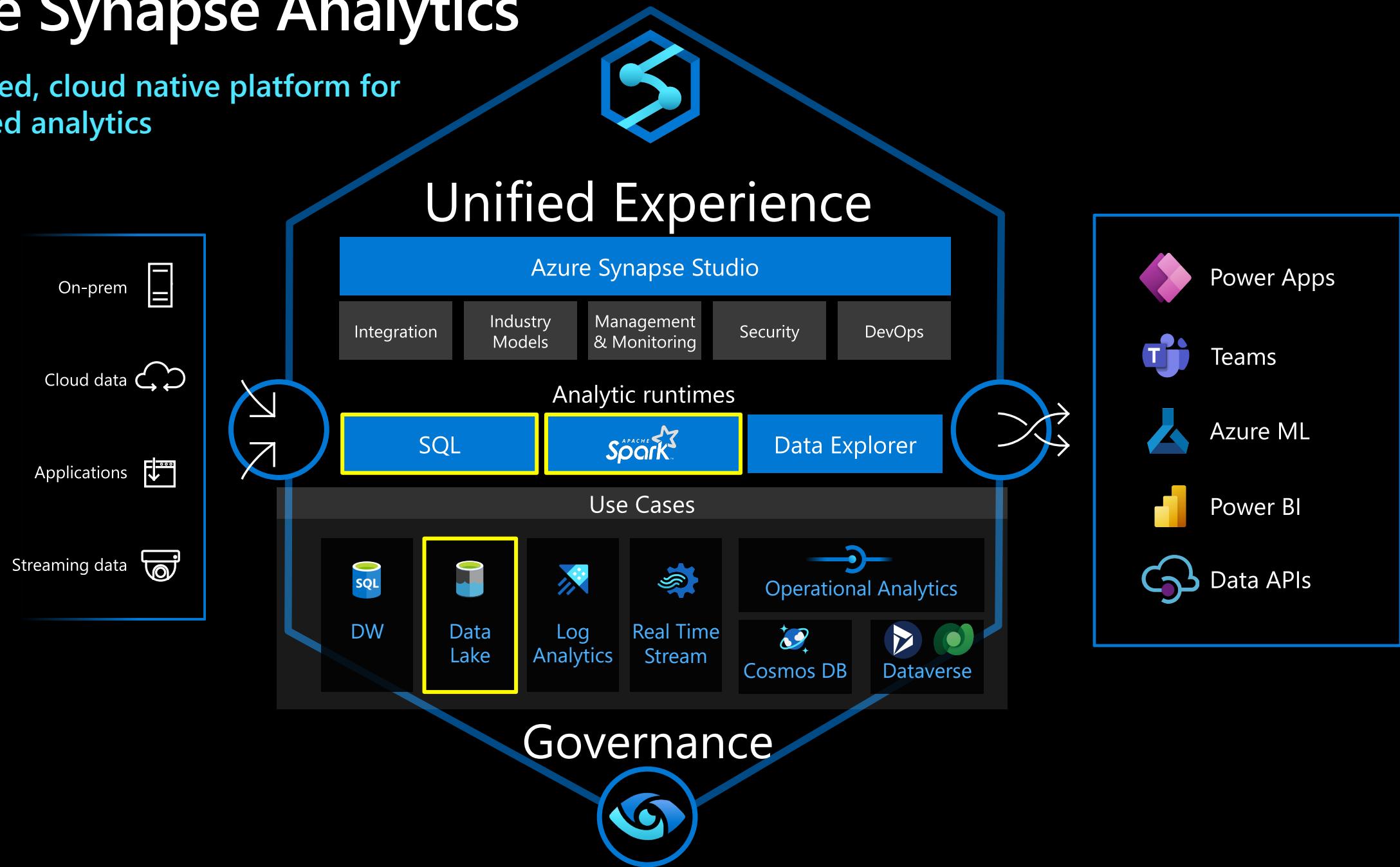
Serverless



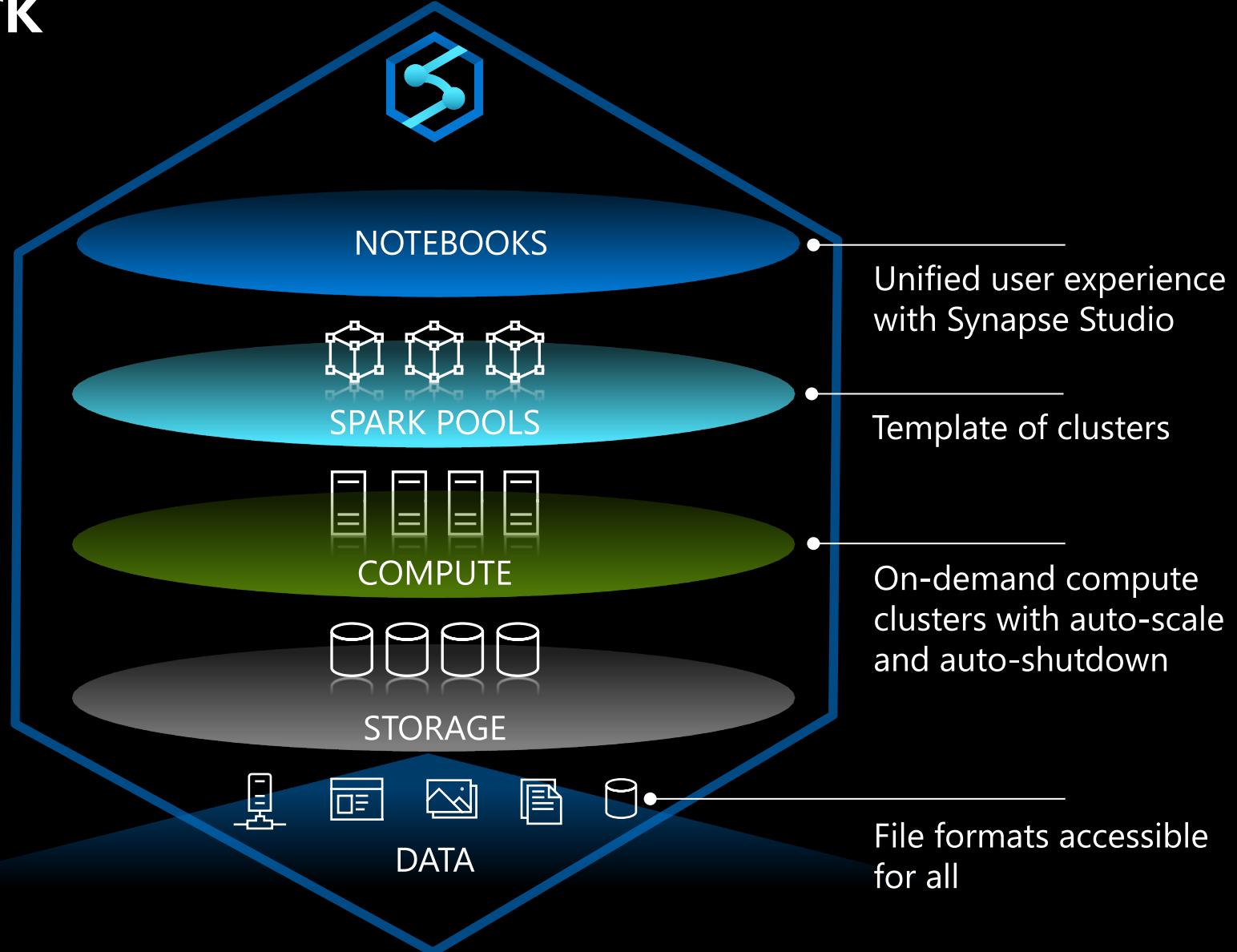
Dedicated

# Azure Synapse Analytics

The unified, cloud native platform for converged analytics

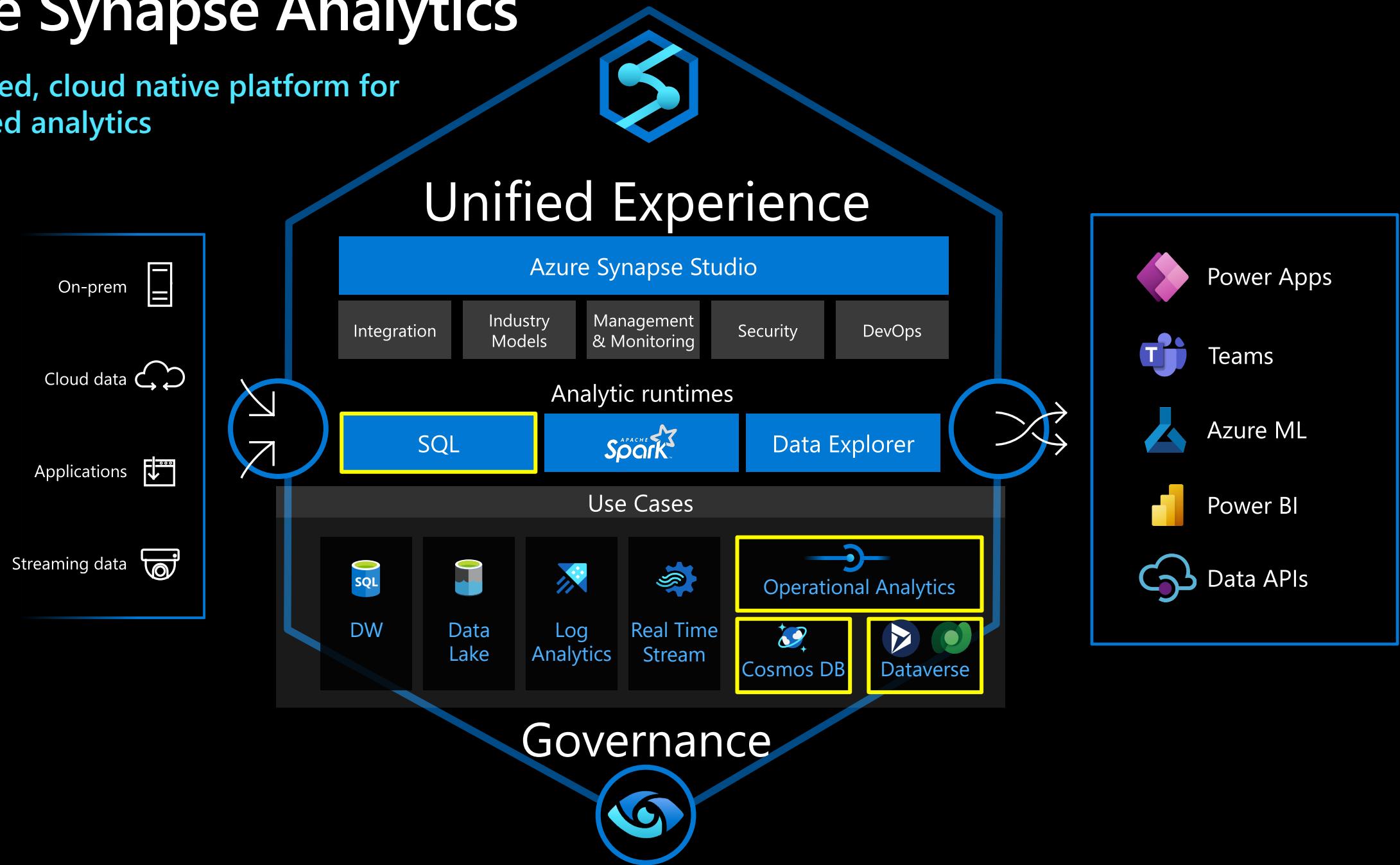


# Serverless Spark Job Service



# Azure Synapse Analytics

The unified, cloud native platform for converged analytics



# Integrating operational data with analytics systems



## Azure Synapse Link

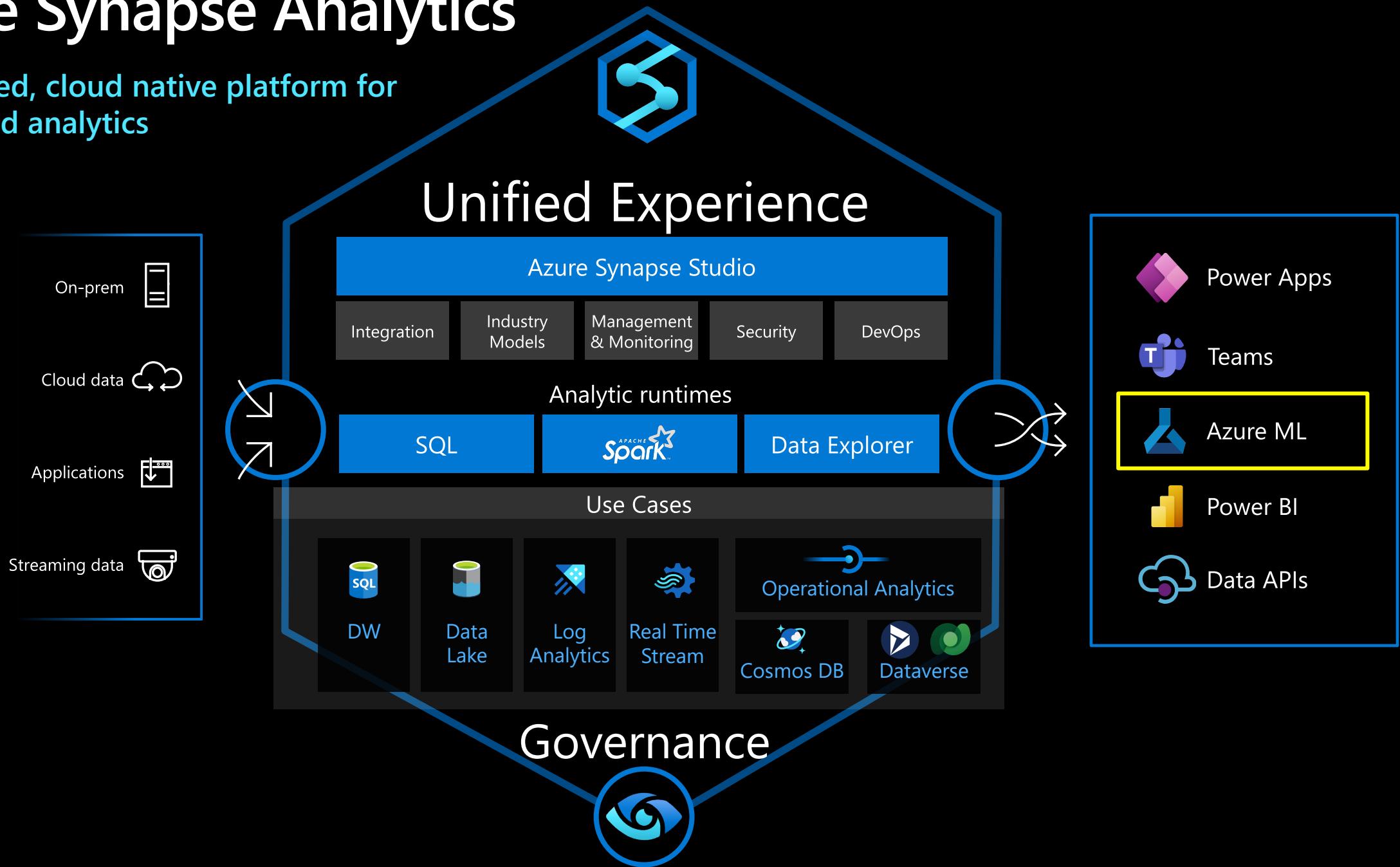
Real-time data analytics

No ETL required

No performance impact on transactions

# Azure Synapse Analytics

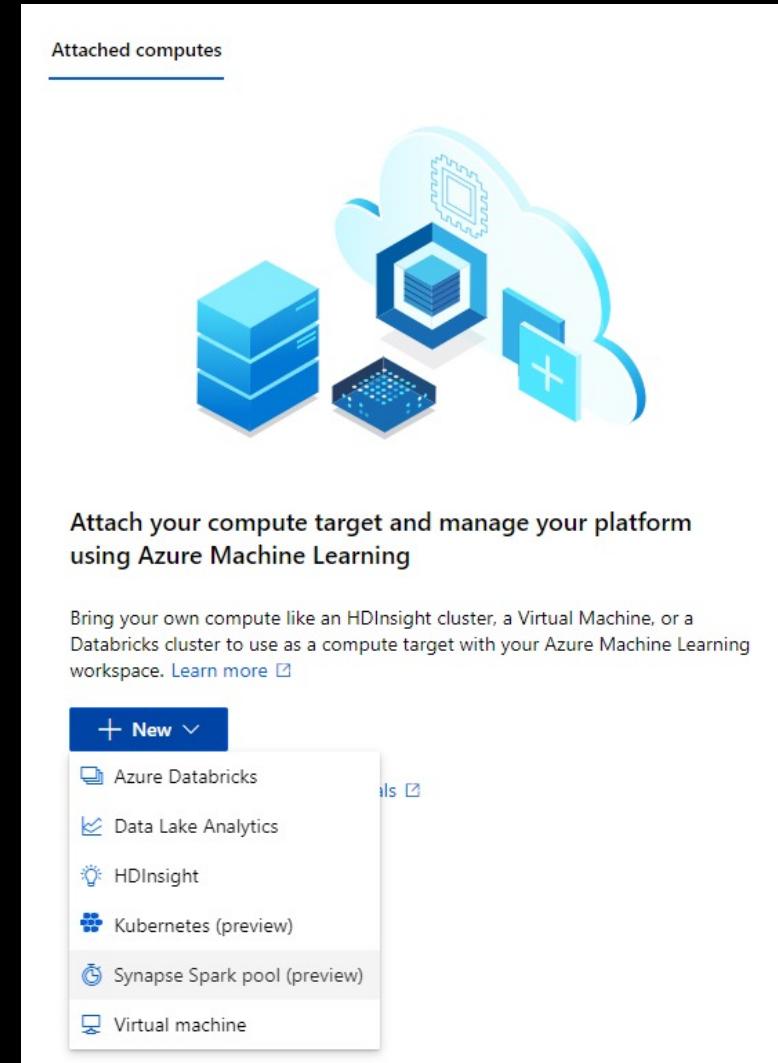
The unified, cloud native platform for converged analytics



# Spark-enabled Azure Machine Learning

## Spark Compute

The Azure Synapse Analytics integration with Azure Machine Learning (preview) allows you to attach an Apache Spark pool backed by Azure Synapse for interactive data exploration and preparation. With this integration, you can have a dedicated compute for data wrangling at scale.



# Machine Learning

## Democratize data science to all

Synapse makes predictive analytics accessible to all

Notebooks provides a code authoring experience for complex predictive models

Automatic ML graphical interface provides a no-code experience for creating ML models

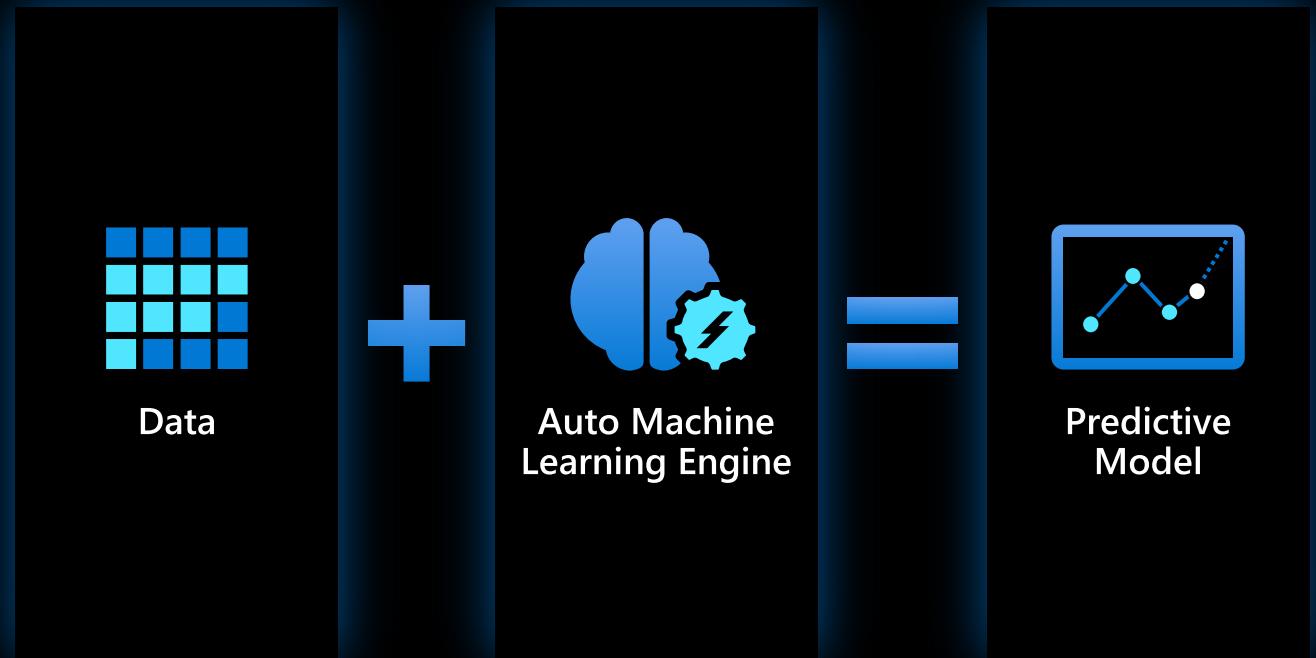
Native integration with Azure Cognitive Search provides access to pre-built models



# Automatic Machine Learning

All you need is data

Fully automated feature exploration



# Automatic Machine Learning

## Code-free in Synapse Studio

No-code creation on Machine Learning models

Democratize ML to everyone since no data science domain knowledge required

Support for ensemble models

Supports classification, regression, and time-series forecasting

The screenshot shows the Microsoft Azure Synapse Analytics Studio interface. On the left, there's a sidebar with icons for Data, Workspace, and Linked resources. Under 'Data', it lists Databases (newpoll, NYCTaxi\_Pool, Predict\_Pool, Streaming\_Pool, WWI\_Pool, NYT2020, SQLServerlessDB) and Tables (retaildata, surfacesalesdb). A specific table named 'retailsales' is highlighted. On the right, the main area shows a code editor with Python code for an experiment. The code includes imports for pandas, azureml, matplotlib, and automl, followed by a function definition for 'align\_'. The interface also features sections for 'Enrich with new model' (with a 'retailsales' dataset selected), 'Choose a model type' (with options for Classification, Regression, and Time series forecasting), and buttons for 'Continue', 'Back', and 'Cancel'.

```
from pandas import *
from azureml import *
from matplotlib import *
from automl import *

def align_():
    if (horizon == '1 hour'):
        df = df.set_index('Time')
        df['Time'] = df['Time'].apply(lambda x: x.replace(minute=0))
        df['Time'] = df['Time'].apply(lambda x: x.replace(second=0))

    else:
        df = df.set_index('Time')
        df['Time'] = df['Time'].apply(lambda x: x.replace(minute=0))
        df['Time'] = df['Time'].apply(lambda x: x.replace(second=0))

    # y and df_fcst
    # align X_test
    X_test = df[['Time', 'retailsales']]
    X_test = pd.get_dummies(X_test)
    X_test = X_test.drop(['Time'], axis=1)

    # X_te
    df_fcst = X_test[X_test['Time'].isna()]
    X_test = X_test[X_test['Time'].notna()]
    X_test = X_test.groupby('Time').mean()
    X_test = X_test.reset_index()

    # drop clean :
    df_all = df[['Time', 'retailsales']]
    df_all = df_all.drop(['Time'], axis=1)
    df_all = df_all.groupby('Time').mean()
    df_all = df_all.reset_index()

    return df_all, df_fcst, X_test, X_te

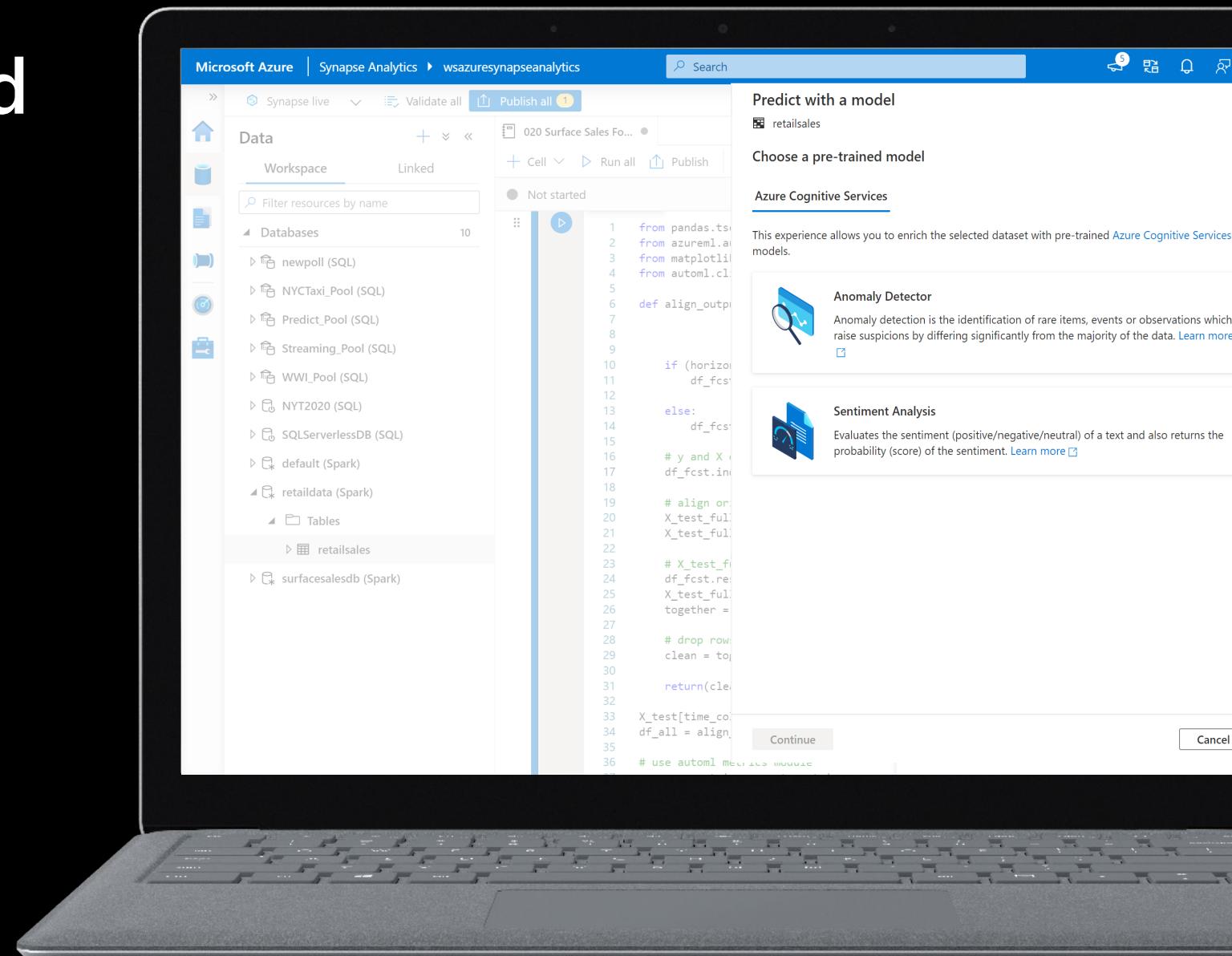
X_test[time] = df_all[time]
df_all = df_all.drop(['Time'], axis=1)
```

# Predict with Pre-trained Model

## Code-free in Synapse Studio

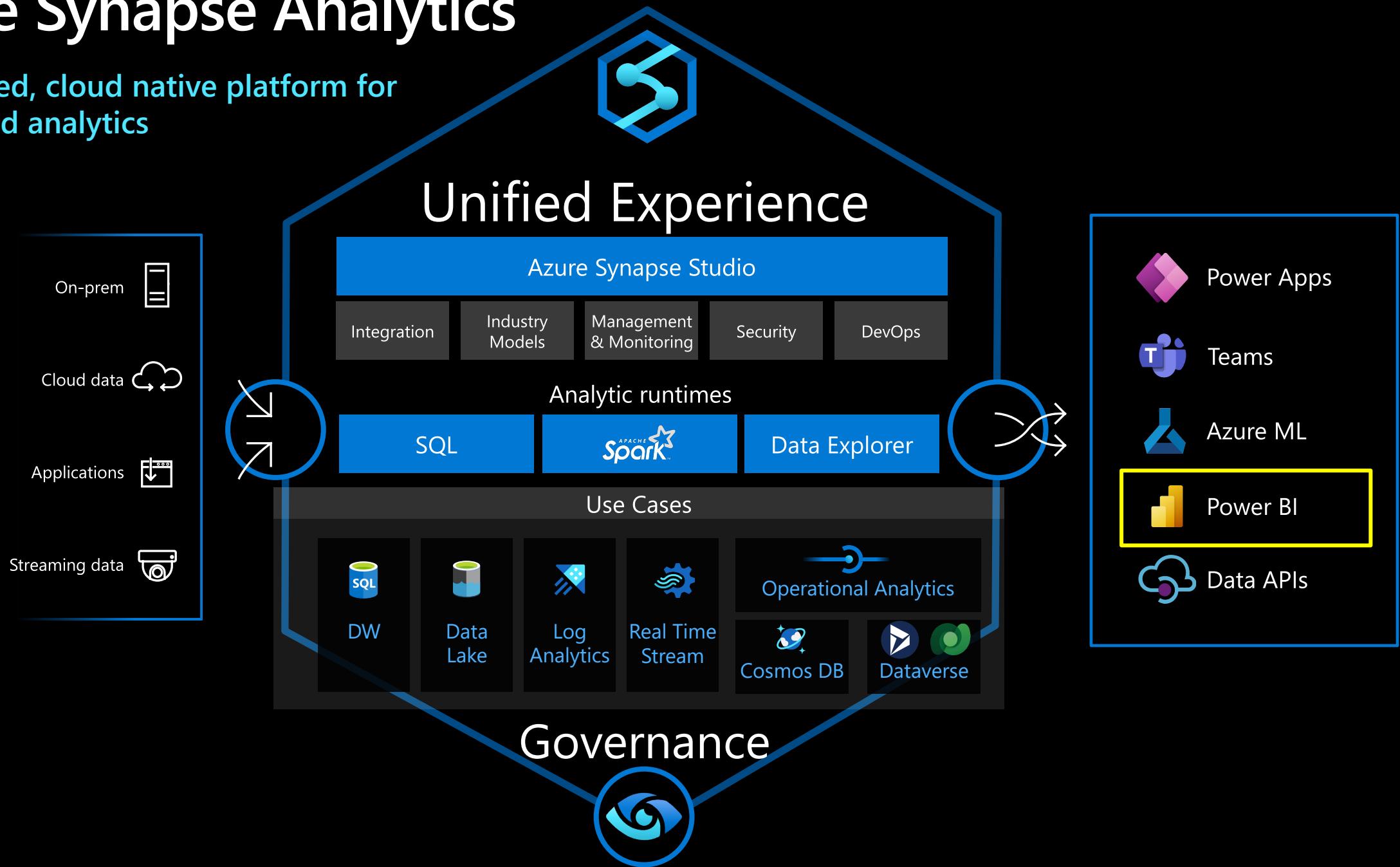
Native integration with Azure Cognitive Search provides access to pre-built models

Supports Anomaly Detector and Sentiment Analysis



# Azure Synapse Analytics

The unified, cloud native platform for converged analytics

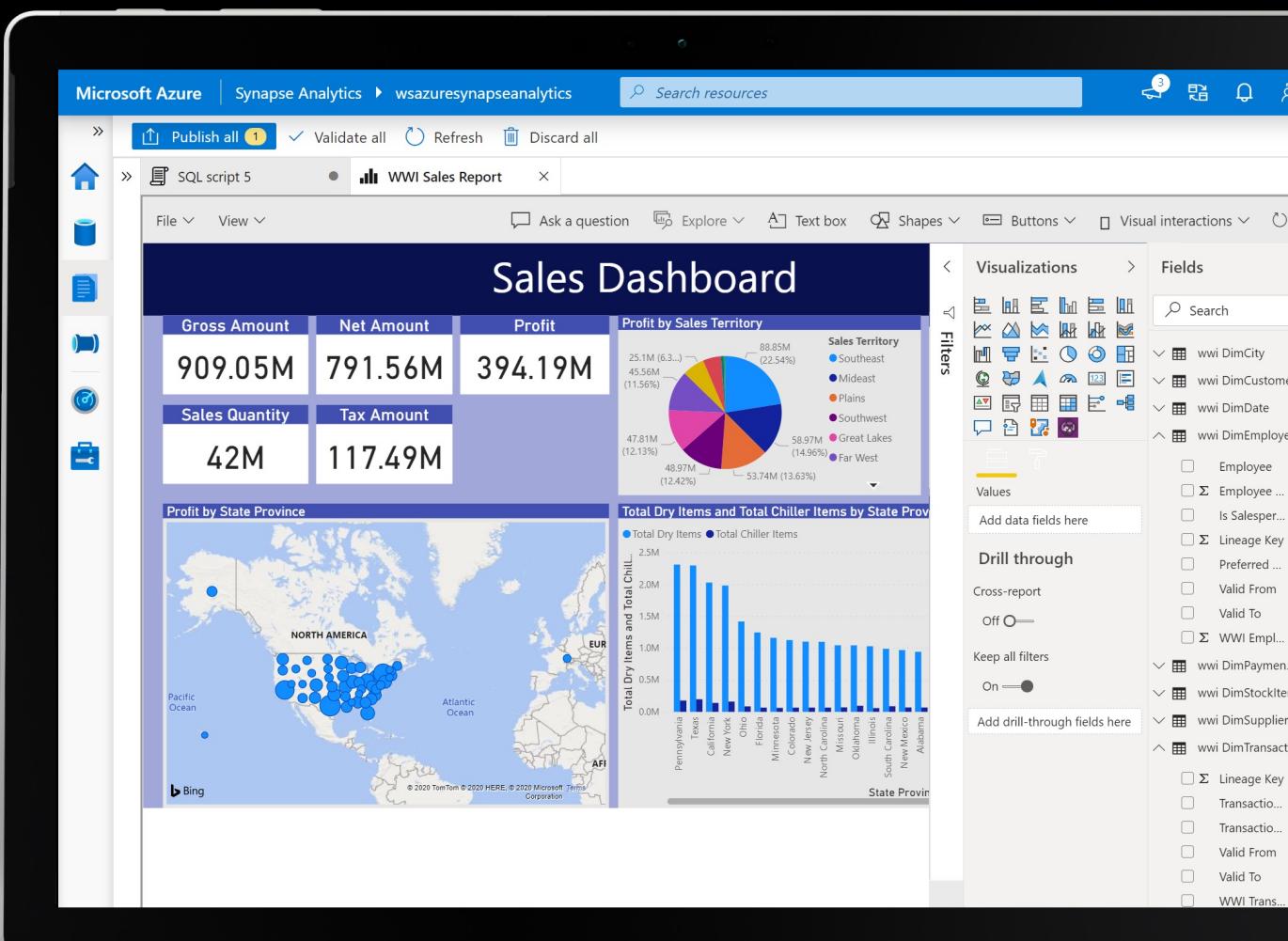


# Power BI integration

## Build dashboard in Synapse Studio

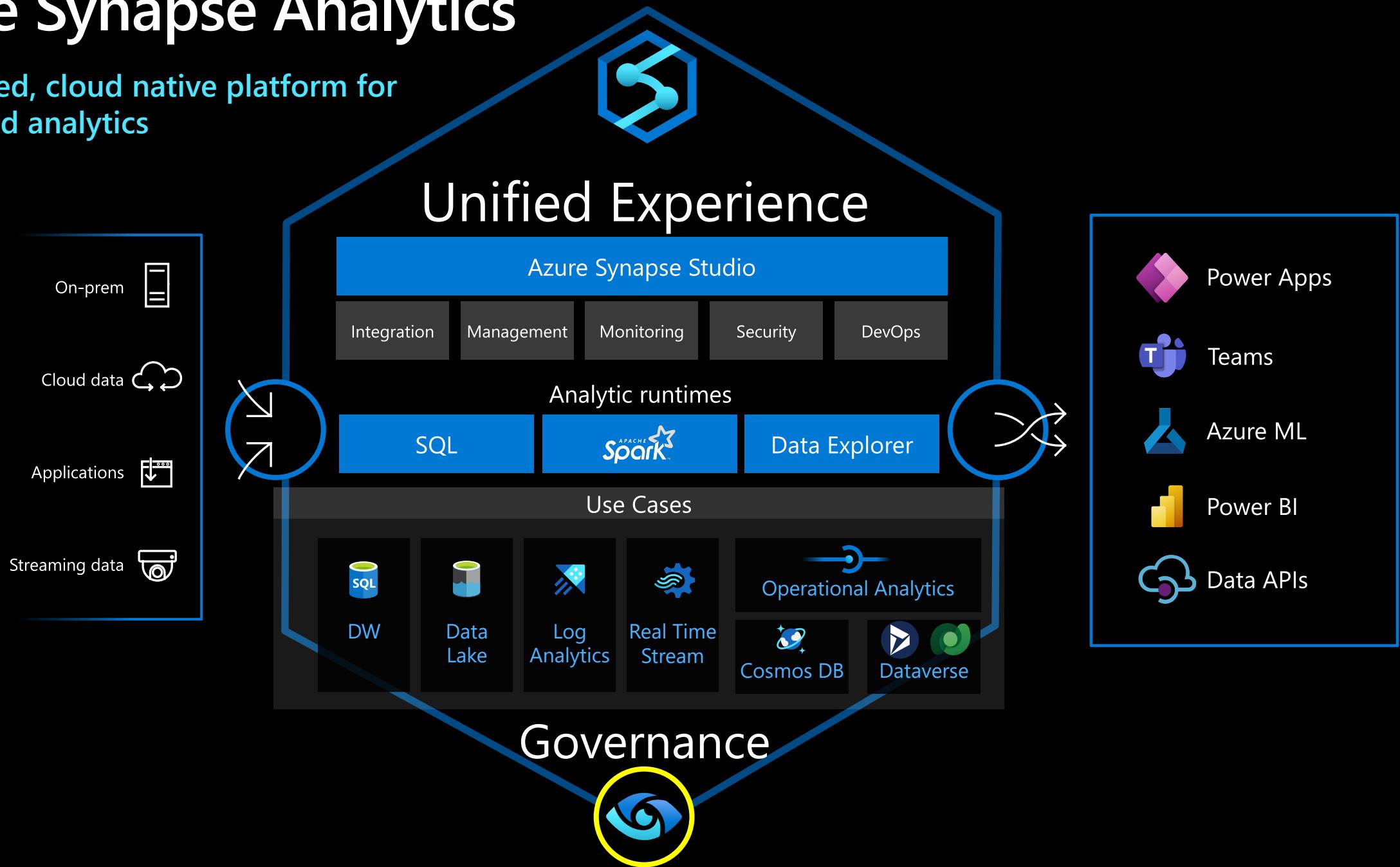
Code-free experience for development rich visualizations

One-click publishing to for secure consumption across the enterprise



# Azure Synapse Analytics

The unified, cloud native platform for converged analytics



# Azure Purview

## UNIFIED DATA GOVERNANCE

### Data Map

- Automate and manage metadata at scale

### Data Catalog

- Enable effortless discovery for data consumers

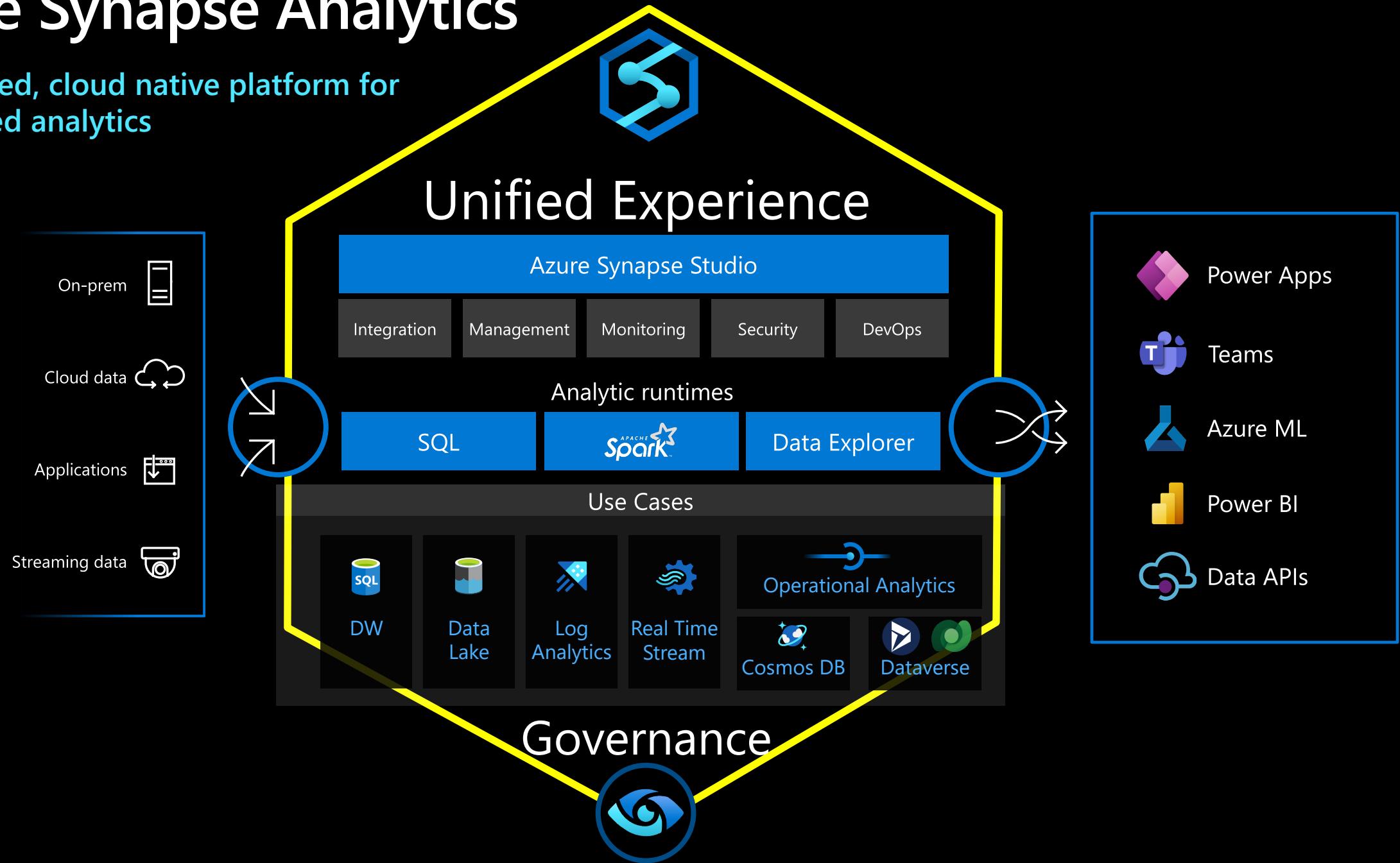
### Data Insights

- Assess data usage across your organization

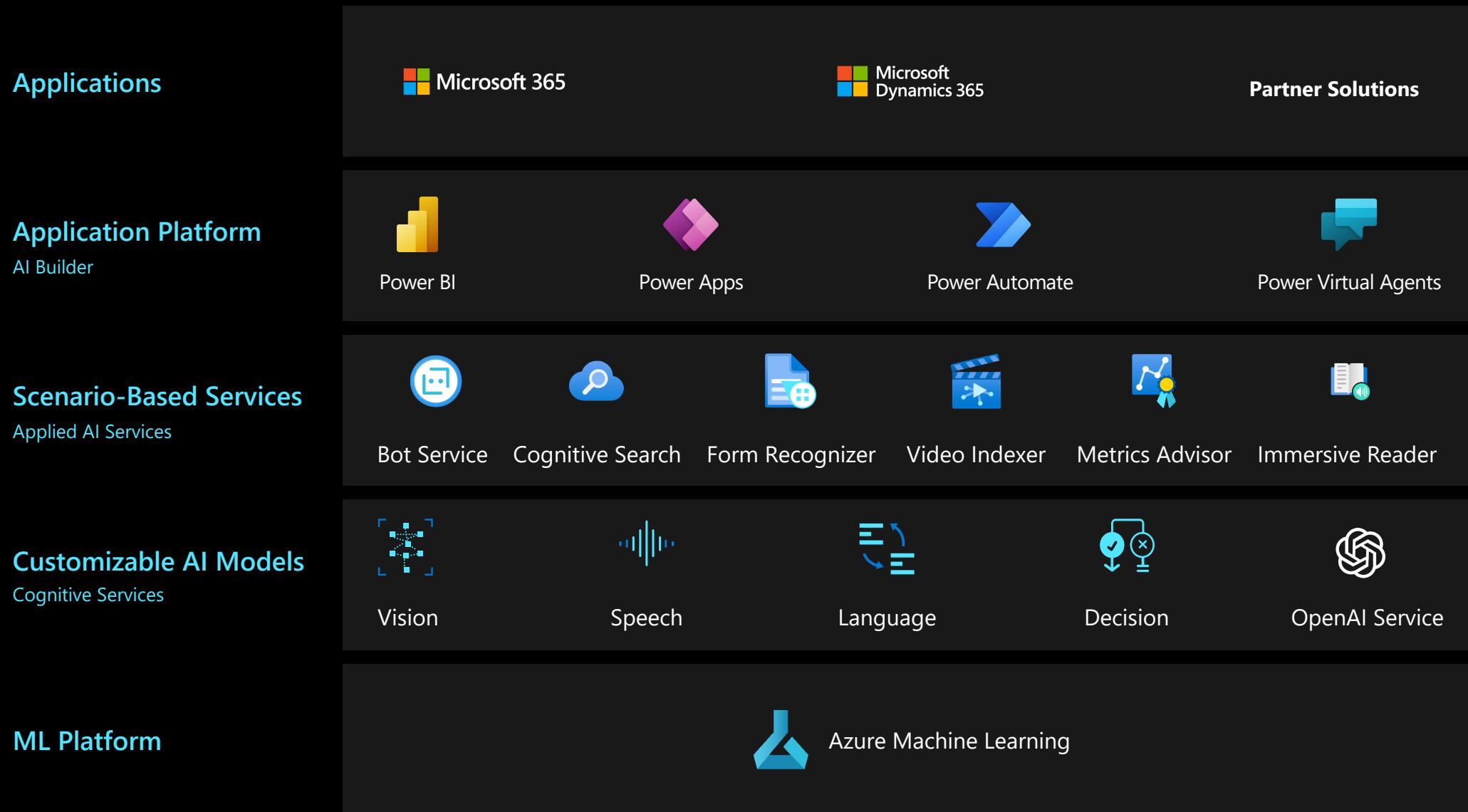


# Azure Synapse Analytics

The unified, cloud native platform for converged analytics



# Azure AI



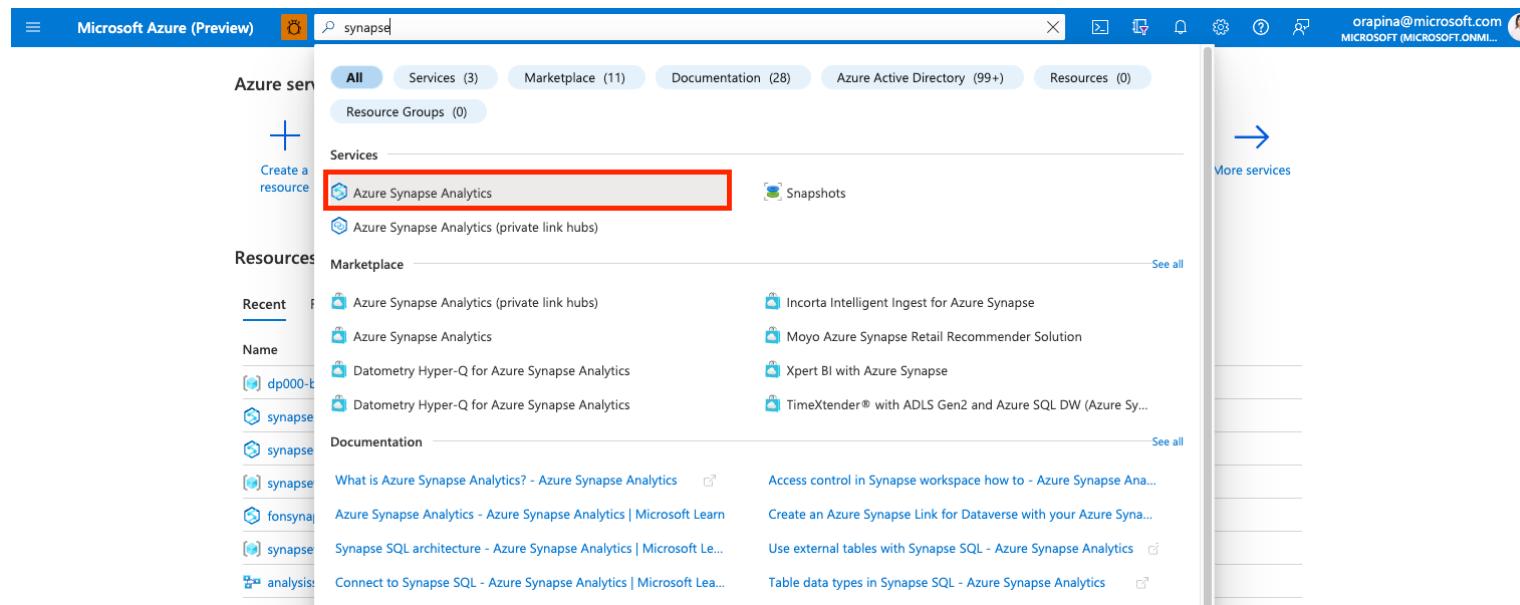
# Demo

# STEP 1 - Create and setup a Synapse workspace

## Create a Synapse workspace in the Azure portal

### Start the process

1. Open the Azure portal, in the search bar enter Synapse without hitting enter.
2. In the search results, under Services, select Azure Synapse Analytics.



# STEP 1 - Create and setup a Synapse workspace

## Create a Synapse workspace in the Azure portal

### Start the process

3. Select **Create** to create a workspace.

The screenshot shows the Microsoft Azure (Preview) portal with the URL [https://portal.azure.com/#blade/HubsBlade](#). The user is signed in as [orapina@microsoft.com](#). The main navigation bar includes 'Home', 'Search resources, services, and docs (G+)', and various icons for notifications and settings. Below the navigation bar, the title 'Azure Synapse Analytics' is displayed, along with the Microsoft logo and the URL [microsoft.onmicrosoft.com](#). The top navigation bar for the service includes 'Create', 'Manage view', 'Refresh', 'Export to CSV', 'Open query', and 'Assign tags'. There are also filter options for 'Subscription', 'Resource group', 'Location', and 'Add filter', as well as grouping and list view controls. The main content area displays a large hexagonal icon with a stylized 'S' and the text 'No Azure Synapse Analytics to display'. A descriptive paragraph explains that Synapse Analytics is a fully-managed service for building modern data warehouses. At the bottom of the page, there is a blue 'Create Synapse workspace' button, which is highlighted with a red rectangular border.

# STEP 1 - Create and setup a Synapse workspace

## Basics tab > Project Details

Fill in the following fields:

- 1. Subscription** - Pick the subscription that would like to use in this tutorial
- 2. Resource group** - Use any resource group.

**rg-synapse-xxxx**

*xx will refer to participant name*

- 3. Managed Resource group** - Leave this blank.

The screenshot shows the Microsoft Azure (Preview) interface for creating a Synapse workspace. At the top, there's a navigation bar with 'Microsoft Azure (Preview)', a search bar, and a user icon. Below the navigation bar, the URL 'Home > Azure Synapse Analytics >' is visible. The main title is 'Create Synapse workspace'. A descriptive text says: 'Select the subscription to manage deployed resources and costs. Use resource groups like folders to organize and manage all of your resources.' There are three configuration sections: 'Subscription \*' with a dropdown menu showing 'Orapin internal subscription', 'Resource group \*' with a dropdown menu showing 'rg\_synapse\_fon' and a 'Create new' link, and 'Managed resource group' with an empty input field labeled 'Enter managed resource group name'.

# STEP 1 - Create and setup a Synapse workspace

## Basics tab > Workspace details

1. **Workspace name** - Pick any globally unique name.

In this tutorial, we'll use **workspace-synapse-xx**

*xx will refer to participant name*

2. **Region** - East Asia

The screenshot shows the 'Create Synapse workspace' page in the Microsoft Azure (Preview) portal. The 'Workspace name' field contains 'workspace-synapse-fon'. The 'Region' dropdown is set to 'East Asia'. Other fields include 'Use Spark on Cosmos' (unchecked) and 'Data Lake Analytics account' (set to 'None').

### Note

A workspace that is not co-located with the client applications or storage can be the root cause of many performance issues. If your data or the clients are placed in multiple regions, you can create separate workspaces in different regions co-located with your data and clients.

# STEP 1 - Create and setup a Synapse workspace

## Select Data Lake Storage Gen 2

1. By **Account name**, select **Create New** and name the new storage account **xxxdatalake** or similar as the name **must be unique**.

2. By **File system name**, select **Create New** and name it **data**.

*This will create a storage container called **data**. The workspace will use this storage account as the "primary" storage account to Spark tables and Spark application logs.*

3. Check the "Assign myself the Storage Blob Data Contributor role on the Data Lake Storage Gen2 account" box.

Microsoft Azure (Preview) Search resources, services, and docs (G+)

Home > Azure Synapse Analytics > Create Synapse workspace

Name your workspace, select a location, and choose a primary Data Lake Storage Gen2 file system to serve as the default location for logs and job output.

Workspace name \*

workspace-synapse-fon

Use Spark on Cosmos ⓘ

Data Lake Analytics account

Region \*

None

East Asia

Select Data Lake Storage Gen2 \* ⓘ

From subscription  Manually via URL

Account name \* ⓘ

fondatalake

Create new

File system name \*

(New) data

Create new

# STEP 1 - Create and setup a Synapse workspace

## Security

≡ Microsoft Azure (Preview) Report a bug

Home > Azure Synapse Analytics >

### Create Synapse workspace

\* Basics \* **Security** Networking Tags Review + create

Configure security options for your workspace.

**Authentication**

Choose the authentication method for access to workspace resources such as SQL pools. The authentication method can be changed later on. [Learn more](#)

Authentication method  ⓘ

Use both local and Azure Active Directory (Azure AD) authentication  
 Use only Azure Active Directory (Azure AD) authentication

SQL Server admin login  ⓘ

SQL Password  ⓘ   ✓

Confirm password  ✓

**System assigned managed identity permission**

Select to grant the workspace network access to the Data Lake Storage Gen2 account using the workspace system identity. [Learn more](#)

Allow network access to Data Lake Storage Gen2 account.  ⓘ

ⓘ  The selected Data Lake Storage Gen2 account does not restrict network access using any network access rules, or you selected a storage account manually via URL under Basics tab. [Learn more](#)

# STEP 1 - Create and setup a Synapse workspace

The screenshot shows the Microsoft Azure (Preview) portal with the title 'Create Synapse workspace'. At the top, there's a green banner indicating 'Validation succeeded'. Below it, the 'Review + create' tab is selected. The 'Product Details' section shows an Azure Synapse Analytics workspace by Microsoft, with a Serverless SQL est. cost/TB of 5.00 USD. The 'Terms' section contains a legal agreement text. The 'Basics' section lists resource details: Subscription (Orapin internal subscription), Resource group (rg\_synapse\_fon), Region (East Asia), Workspace name ((new) workspace-synapse-fon), Data Lake Storage Gen2 account (https://fondatalake.dfs.core.windows.net), and Data Lake Storage Gen2 file system ((new) data). At the bottom, there are buttons for 'Create', '< Previous', 'Next >', and 'Download a template for automation'.

Validation succeeded

\* Basics \* Security Networking Tags Review + create

**Product Details**

Azure Synapse Analytics workspace by Microsoft Serverless SQL est. cost/TB ①  
5.00 USD

[Terms of use](#) | [Privacy policy](#)

**Terms**

By clicking Create, I (a) agree to the legal terms and privacy statement(s) associated with the Marketplace offering(s) listed above; (b) authorize Microsoft to bill my current payment method for the fees associated with the offering(s), with the same billing frequency as my Azure subscription; and (c) agree that Microsoft may share my contact, usage and transactional information with the provider(s) of the offering(s) for support, billing and other transactional activities. Microsoft does not provide rights for third-party offerings. For additional details see [Azure Marketplace Terms](#).

**Basics**

Subscription	Orapin internal subscription
Resource group	rg_synapse_fon
Region	East Asia
Workspace name	(new) workspace-synapse-fon
Data Lake Storage Gen2 account	https://fondatalake.dfs.core.windows.net
Data Lake Storage Gen2 file system	(new) data

Create < Previous Next > Download a template for automation

# STEP 1 - Create and setup a Synapse workspace

The screenshot shows the Microsoft Azure (Preview) portal with the title "Microsoft.Azure.SynapseAnalytics-20221106234417 | Overview". The left sidebar includes "Home", "Overview" (which is selected), "Inputs", "Outputs", and "Template". The main content area displays a green checkmark icon followed by the text "Your deployment is complete". Below this, it shows deployment details: Deployment name: Microsoft.Azure.SynapseAnalytics-20221106234..., Start time: 11/6/2022, 11:48:27 PM, Subscription: Orapin internal subscription, Resource group: rg-gbdi-synapse101-orapin, and Correlation ID: ba503dcb-a787-436e-abca-90e6a8fd8b8d. There are also sections for "Deployment details" and "Next steps" with a "Go to resource group" button.

Select **Review + create > Create**. Your workspace is ready in a few minutes.

# STEP 1 - Open Synapse Studio

After your Azure Synapse workspace is created, you have two ways to open Synapse Studio:

1. Open your Synapse workspace in the [Azure portal](#), in the **Overview** section of the Synapse workspace, select **Open** in the Open Synapse Studio box.
2. Go to the <https://web.azuresynapse.net> and sign in to your workspace.

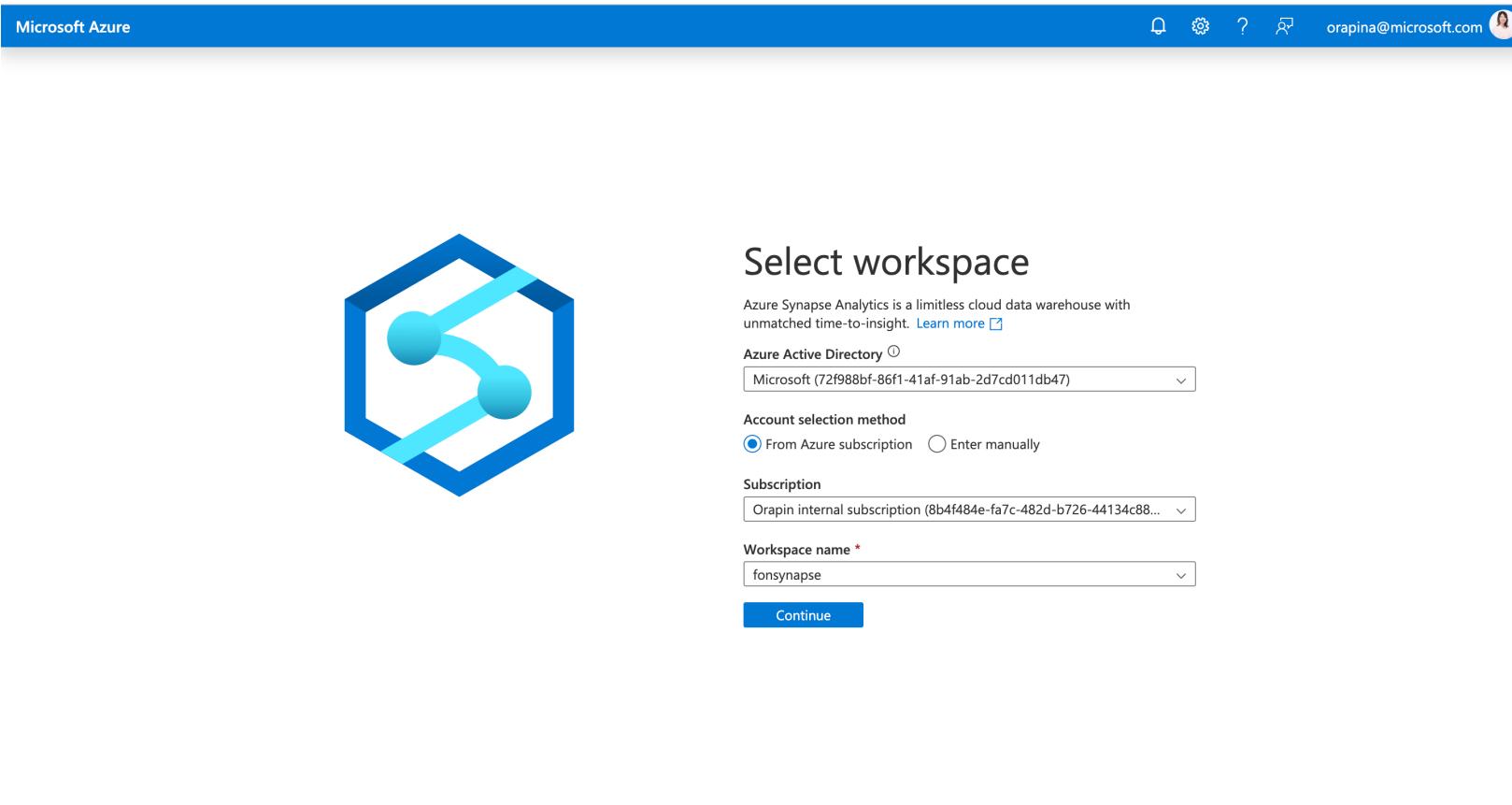
The screenshot shows the Azure Synapse workspace Overview page. The left sidebar includes sections for Overview, Activity log, Access control (IAM), Tags, Diagnose and solve problems, Settings (with sub-options for Azure Active Directory, Properties, and Locks), Analytics pools (with sub-options for SQL pools, Apache Spark pools, and Data Explorer pools (preview)), and Security (with sub-options for Encryption, Networking, Identity, and Private endpoint connections). The main content area features a search bar at the top, followed by a 'Getting started' section with 'Open Synapse Studio' and 'Read documentation' buttons, and an 'Analytics pools' section listing various pool types and their details. The 'Analytics pools' table has columns for Name, Type, and Size.

Name	Type	Size
Built-in	Serverless	Auto
sql6fj8srq	Dedicated	DW100c
spark6fj8srq	Apache Spark pool	Small
adx6fj8srq	Data Explorer pool	Medium

# STEP 1 - Open Synapse Studio

After your Azure Synapse workspace is created, you have two ways to open Synapse Studio:

1. Open your Synapse workspace in the [Azure portal](#), in the **Overview** section of the Synapse workspace, select **Open** in the Open Synapse Studio box.
2. Go to the <https://web.azuresynapse.net> and sign in to your workspace.



The screenshot shows the 'Select workspace' page of the Microsoft Azure Synapse Studio. At the top, there's a blue header bar with the Microsoft Azure logo, a search bar, and user information (orapina@microsoft.com). Below the header, the page title 'Select workspace' is displayed. To the left is the Azure Synapse Analytics logo, which consists of three blue spheres connected by lines within a hexagonal frame. The main form area contains the following fields:

- Azure Active Directory**: A dropdown menu showing 'Microsoft (72f988bf-86f1-41af-91ab-2d7cd011db47)'.
  - Account selection method**: Radio buttons for 'From Azure subscription' (selected) and 'Enter manually'.
  - Subscription**: A dropdown menu showing 'Orapin internal subscription (8b4f484e-fa7c-482d-b726-44134c88...)'.
    - Workspace name \***: An input field containing 'fonsynapse'.
- Continue**: A blue button at the bottom of the form.

# Explore Synapse Studio

Microsoft Azure | synapse6fj8srq

Synapse Analytics workspace  
synapse6fj8srq

New ▾

Ingest      Explore and analyze      Visualize

Discover more

Knowledge center      Browse partners

Recent resources

Name	Last opened by you
Aggregate product sales	4 days ago



The screenshot shows the Microsoft Azure Synapse Studio interface. At the top, there's a blue header bar with the Microsoft Azure logo, the workspace name 'synapse6fj8srq', and various navigation icons. Below the header is a sidebar with icons for Home, Ingest, Explore and analyze, Visualize, and other workspace management options. The main content area features a large, visually appealing chart that combines 3D bars with a network graph, symbolizing complex data analysis. Below the chart, there are three main buttons: 'Ingest' (cloud icon), 'Explore and analyze' (bar chart icon), and 'Visualize' (bar chart icon). Further down, there are links to 'Discover more' (Knowledge center and Browse partners) and a 'Recent resources' section showing a single item: 'Aggregate product sales' last opened 4 days ago.

# To Follow the exercise

<https://microsoftlearning.github.io/dp-203-azure-data-engineer/Instructions/Labs/01-Explore-Azure-Synapse.html#before-you-start>

Data:

<https://raw.githubusercontent.com/MicrosoftLearning/dp-203-azure-data-engineer/master/Allfiles/labs/01/adventureworks/products.csv>

# Delete Azure resources

Now that you've finished exploring Azure Synapse Analytics, you should delete the resources you've created to avoid unnecessary Azure costs.

- 1.Close the Synapse Studio browser tab and return to the Azure portal.
- 2.On the Azure portal, on the **Home** page, select **Resource groups**.
- 3.Select the **rg-synapse101-xxxx** resource group for your Synapse Analytics workspace (not the managed resource group), and verify that it contains the Synapse workspace, storage account, SQL pool, Data Explorer pool, and Spark pool for your workspace.
- 4.At the top of the **Overview** page for your resource group, select **Delete resource group**.
- 5.Enter the **rg-synapse101-xxxx** resource group name to confirm you want to delete it, and select **Delete**.
- 6.After a few minutes, your Azure Synapse workspace resource group and the managed workspace resource group associated with it will be deleted.

# Azure Synapse Analytics

## Resources

- [Watch Azure Synapse Analytics demo videos](#)
- [Get Started](#)
- [Realize Integrated Analytical Solutions with Azure Synapse Analytics \(Learning\)](#)
- [Integrate Azure Synapse Analytics with Azure Data and AI services \(Learning\)](#)
- [Documentation](#)
- [Knowledge Center](#)