



# Open Education Analytics

a modern education data estate

# Dispersed data prevents strategic use of data and AI



## For the Use Cases You Need



**Personalized Learning Pathways**



**Real Time Data Dashboards**



**Open Data for Research and Evaluation**



**Identifying and supporting at risk students**



**State, Province and National Longitudinal Reporting**

Data Engineers



Data Scientists



Contracted Data Scientists



Analysts



Learners, Educators



School Leaders, Researchers



## 1) Authentication/Authorization



Azure Active Directory

## 2) Data Processing Tools



Azure Synapse Analytics

## 3) Storage



Azure Data Lake Storage



On-premises data



Cloud data



SaaS data



Personalized Learning Pathways



Realtime Data Dashboards



Open Data for Research and Evaluation

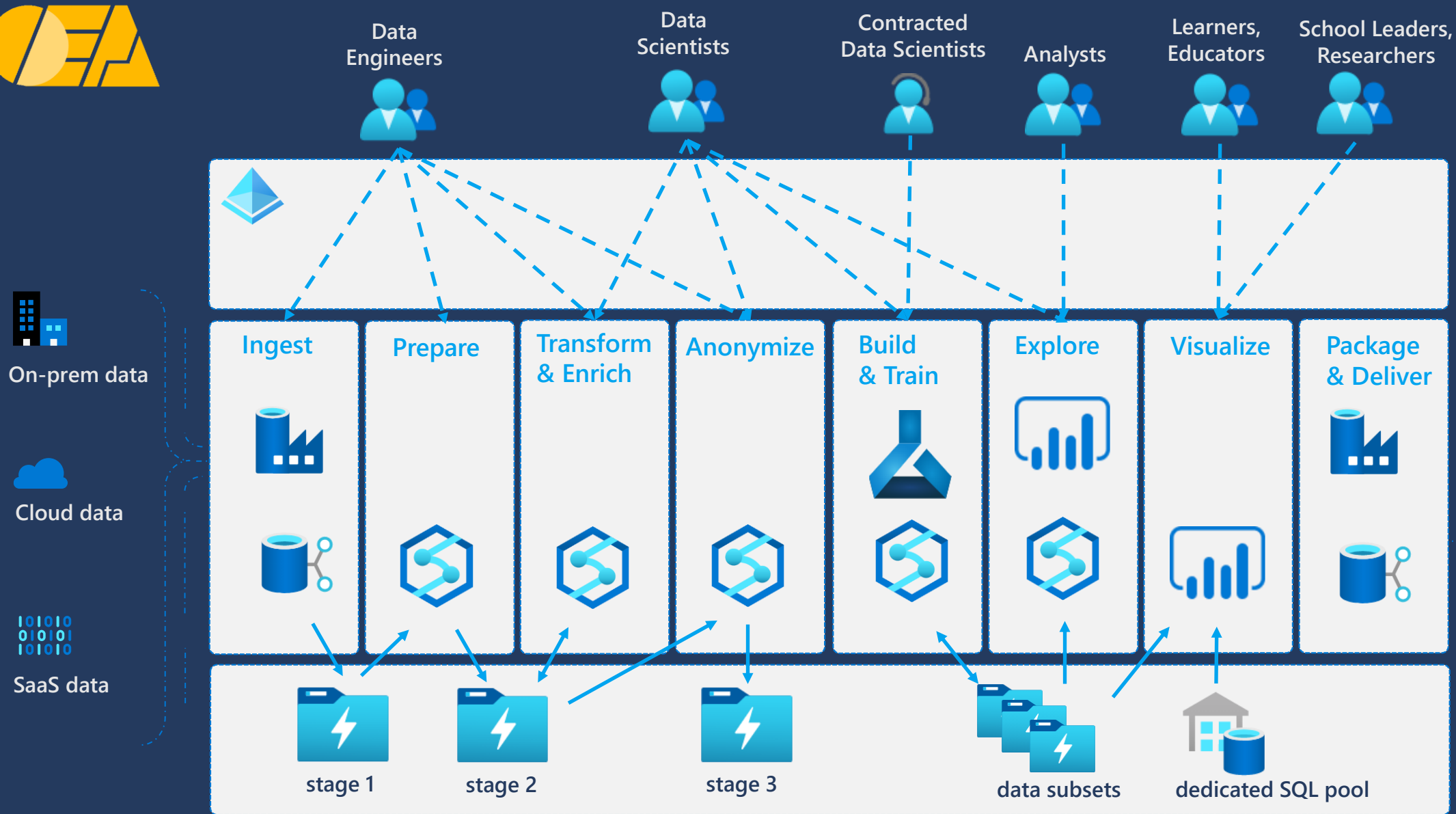


State, Province and National Reporting

1) Collect

2) Compute

3) Communicate



1) Collect

2) Compute

3) Communicate

# Studio

A single place for Data Engineers, Data Scientists, and IT Pros to collaborate on enterprise analytics

Microsoft Azure | Synapse Analytics | internalsandboxwe5

prlangad@microsoft.com

Home

Data

Develop

Orchestrate

Monitor

Manage

Synapse workspace  
internalsandboxwe5

New

**Ingest**  
Use the copy data tool to import data once or on a schedule.

**Explore**  
Learn how to navigate and interact with your data.

**Analyze**  
Learn how to use SQL or Spark to get insights from your data.

**Visualize**  
Build interactive reports with integrated Power BI capabilities.

**Resources**

Recent Pinned

NAME	LAST OPENED BY YOU
CopyPipeline_0313	13 minutes ago

**Useful links**

[Synapse Analytics overview](#)  
Discover the capabilities offered by Synapse and learn how to make the most of them.

[Pricing](#)  
Learn about pricing details for Synapse capabilities.

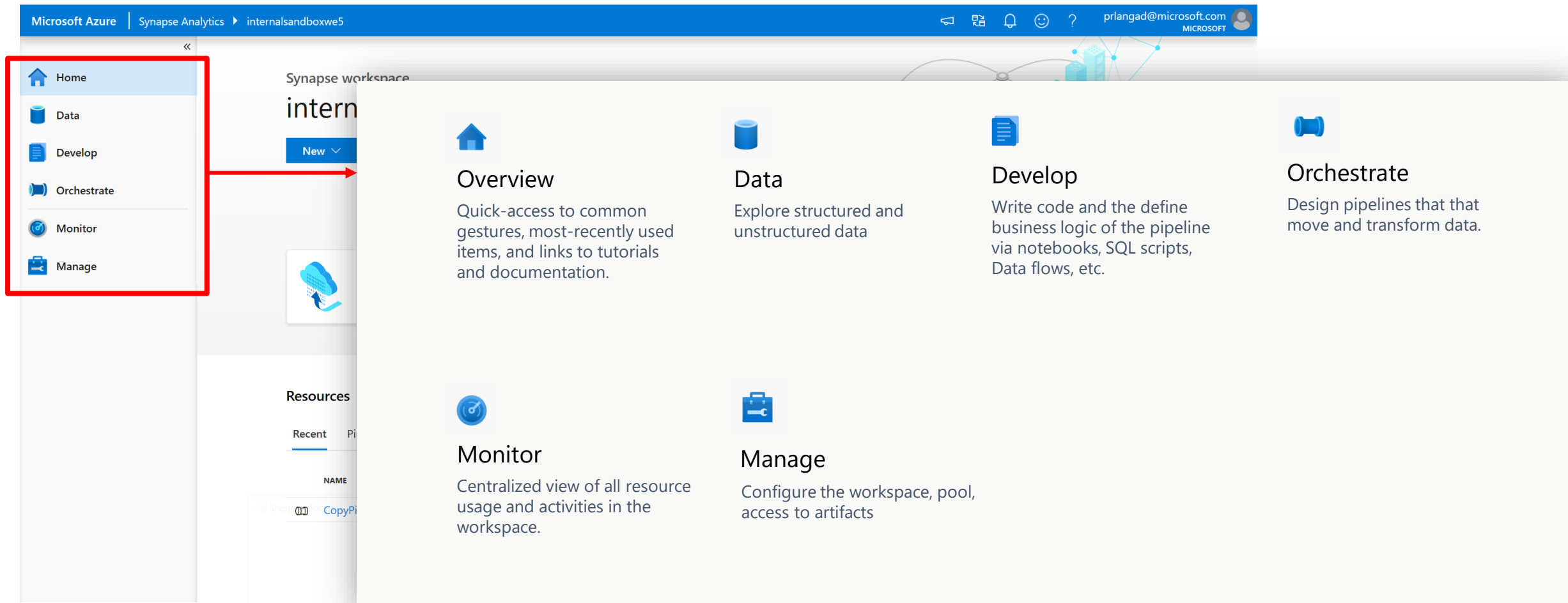
[Documentation](#)  
Visit the documentation center for quickstarts, how-to guides, and references for PowerShell, APIs, etc.

[Give feedback](#)  
Share your comments or suggestions with us to improve Synapse.

# Synapse Studio

Synapse Studio divided into **Activity hubs**.

These organize the tasks needed for building analytics solution.





On-prem data



Cloud data



SaaS data

Ingest



Azure Data  
Factory



Azure Data  
Share

Microsoft Azure | Synapse Analytics | syn-oea-cisd3demo4

Search

Synapse live | Validate all | Publishing 7

Home | Data | Develop | Integrate | Monitor | Manage

**Integrate**

Filter resources by name

Pipelines 7

- M365
  - Get\_logins
  - Get\_members
  - Get\_schools\_classes\_users
  - Get\_users\_from\_AAD
- On prem sources

**Activities**

Search activities

- Synapse
- Move & transform
- Azure Data Explorer
- Azure Function
- Batch Service
- Databricks
- Data Lake Analytics
- General
- HDInsight
- Iteration & conditionals
- Machine Learning

Get\_schools\_classes...

Validate | Debug | Add trigger

Copy data

- Retrieve class
- Retrieve\_schools
- Retrieve\_users

Notebook

- AAD\_classes\_users

1) Collect



Microsoft Azure | Synapse Analytics | syn-oea-cisd3demo4

Synapse live | Validate all | Publish all

Home | Data | Develop | Integrate | Monitor | Manage

Data

Workspace | Linked

Filter resources by name

- Azure Data Lake Storage Gen2 1
  - public
  - stage1
  - stage2
  - stage3
  - synapse (Primary)
  - test-env
- Integration datasets 8

Get\_schools\_classes... | stage1

New SQL script | New notebook | New data flow | New integration dataset | More

stage1 > AAD > classes

Name	Last Modified	Content Type	Size
data_24d8bb08-003e-4989-975e-996746266ef9_37c0be75-10b2-443b-8b69-e725b52e9cf0....	6/11/2021, 4:53:30 PM		159.9 KB



On-prem data



Cloud data



SaaS data

Ingest

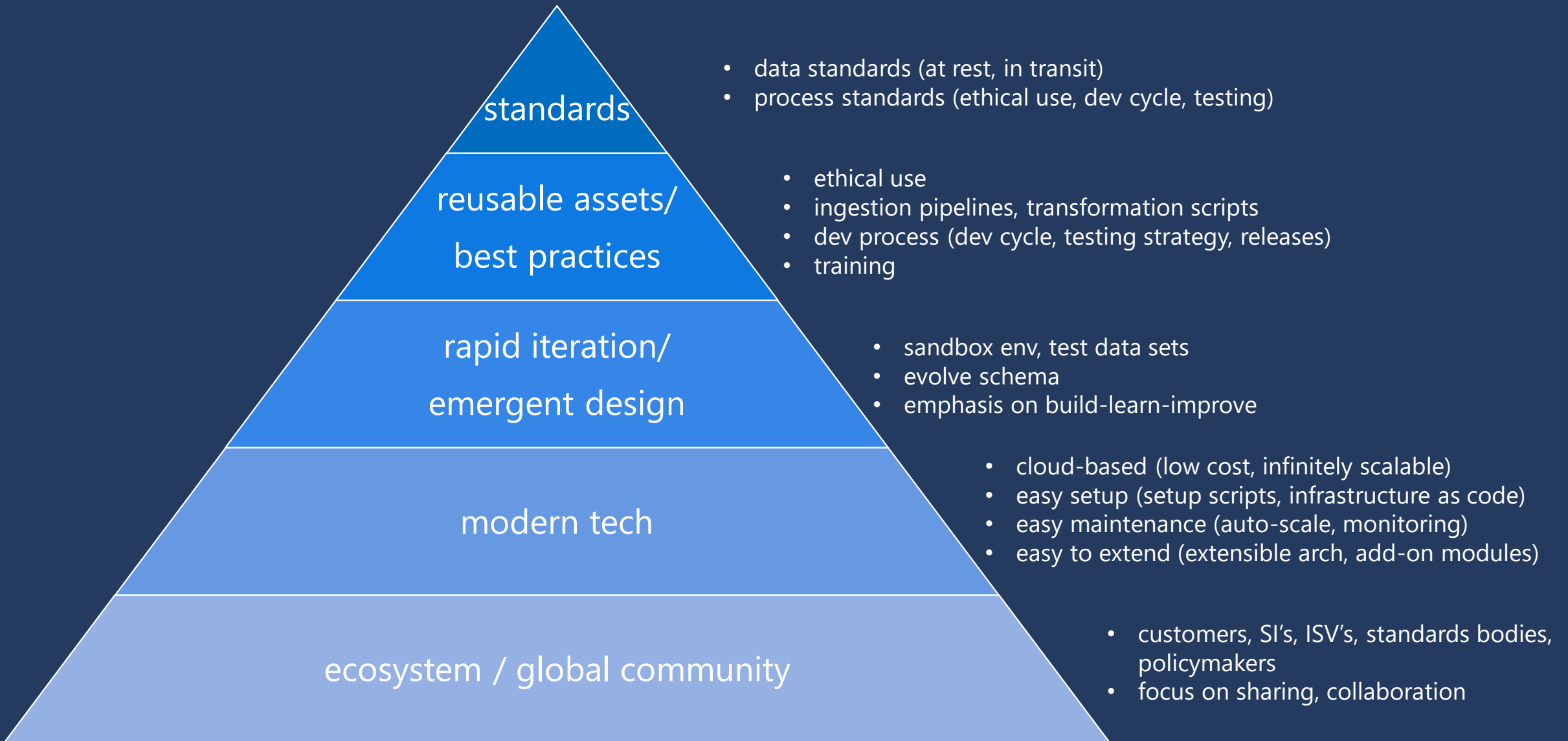


stage 1

1) Collect



## OEA Focus



Search or jump to... Pull requests Issues Marketplace Explore

microsoft / OpenEduAnalytics

<> Code ! Issues 1 Pull requests Discussions Actions Projects Wiki Security Insights Settings

main 2 branches 2 tags

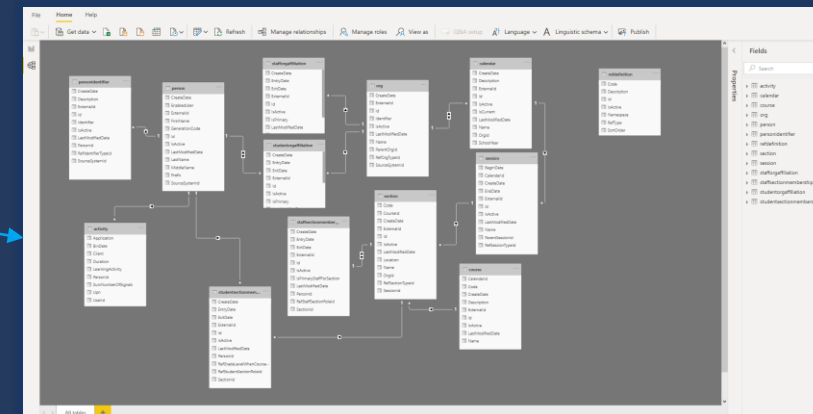
genegc Merge branch 'main' of https://github.com/microsoft/OpenEduAnalytics b009e63 5 days ago 115 commits

.vs	add to readme	4 months ago
docs	Add files via upload	last month
modules	Update readme.md	10 days ago
packages	updated setup scripts.	last month
CODE_OF_CONDUCT.md	Initial CODE_OF_CONDUCT.md commit	7 months ago
CONTRIBUTING.md	Create CONTRIBUTING.md	last month
LICENSE	Initial LICENSE commit	7 months ago

```

1 # Process Roster data from stage 1 to stage 2
2 #
3 # Sets up the edu_dl (stage 2 data lake) with whatever data is found in the DIP inbound folder.
4 # This includes:
5 # - adding column names
6 # - casting values into a schema
7
8 # Calendar
9 df = spark.read.csv(stage1_path + '/Roster/Calendar.csv', header='false')
10 if (df.count() > 0):
11     sqlContext.registerDataFrameAsTable(df, 'Calendar')
12 df = spark.sql("select _c0 Id, _c1 Name, _c2 Description, cast(_c3 as int) SchoolYear, cast(_c4 as int) as B")
13 df.write.format("parquet").mode("overwrite").save(f'{stage2_path}/Calendar')
14
15 # Course
16 df = spark.read.csv(stage1_path + '/Roster/Course.csv', header='false')
17 if (df.count() > 0):
18     sqlContext.registerDataFrameAsTable(df, 'Course')
19 df = spark.sql("select _c0 Id, _c1 Name, _c2 Code, _c3 Description, _c4 ExternalId, to_timestamp(_c5

```



Microsoft

# Open Edu Analytics Implementation Guide

Published October, 2020

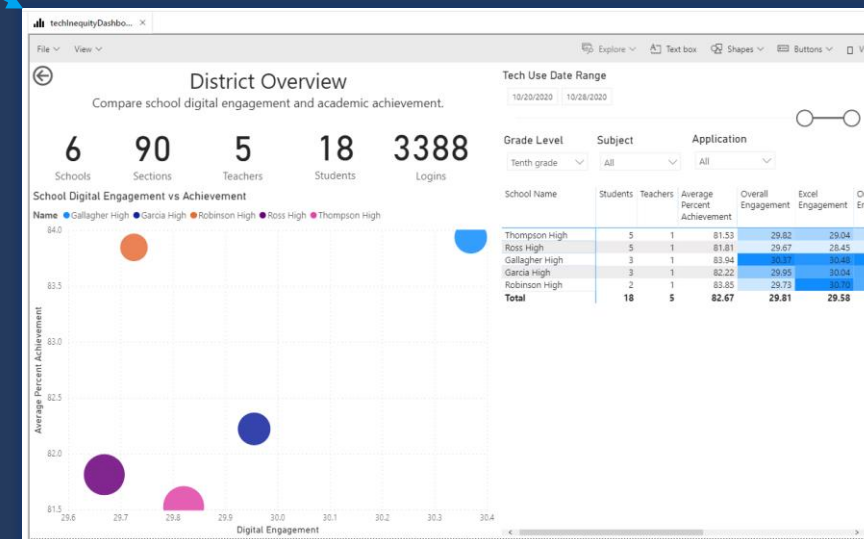
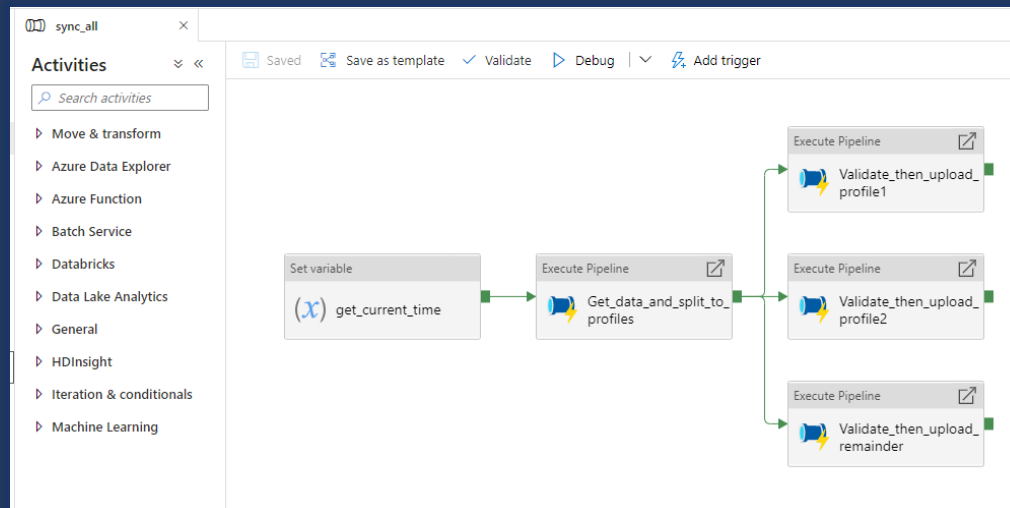
Introduction.....1

1) Setup Data Share in School Data Sync.....1

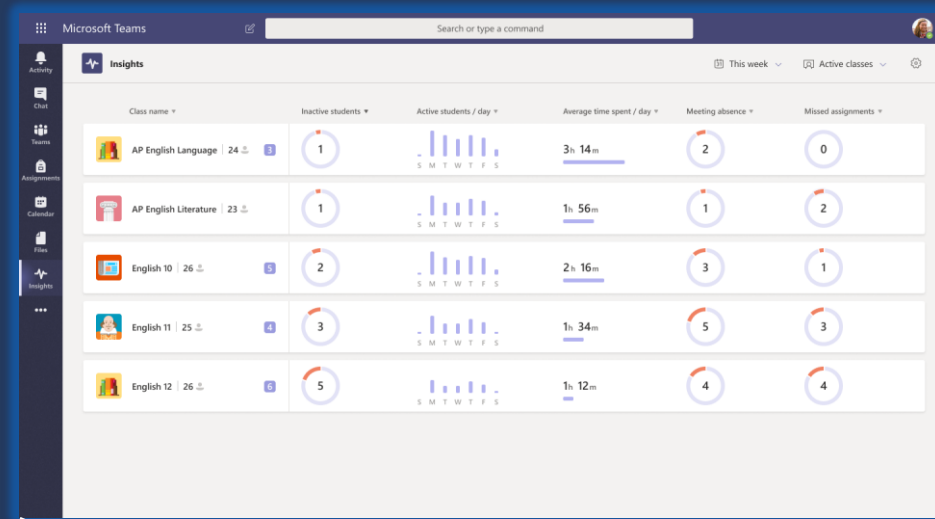
2) Setup Resources in Azure Portal.....2

3) Setup Synapse.....5

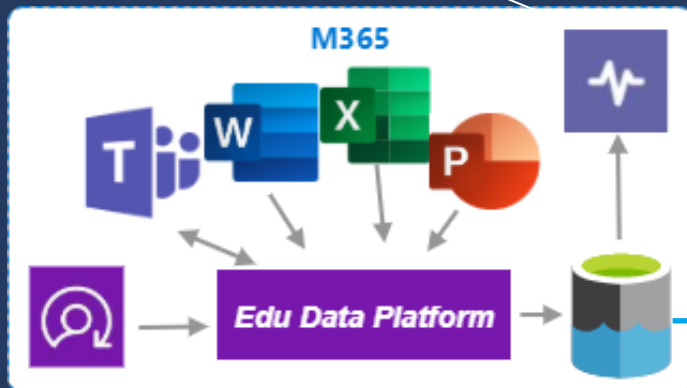
4) Setup Power BI.....7



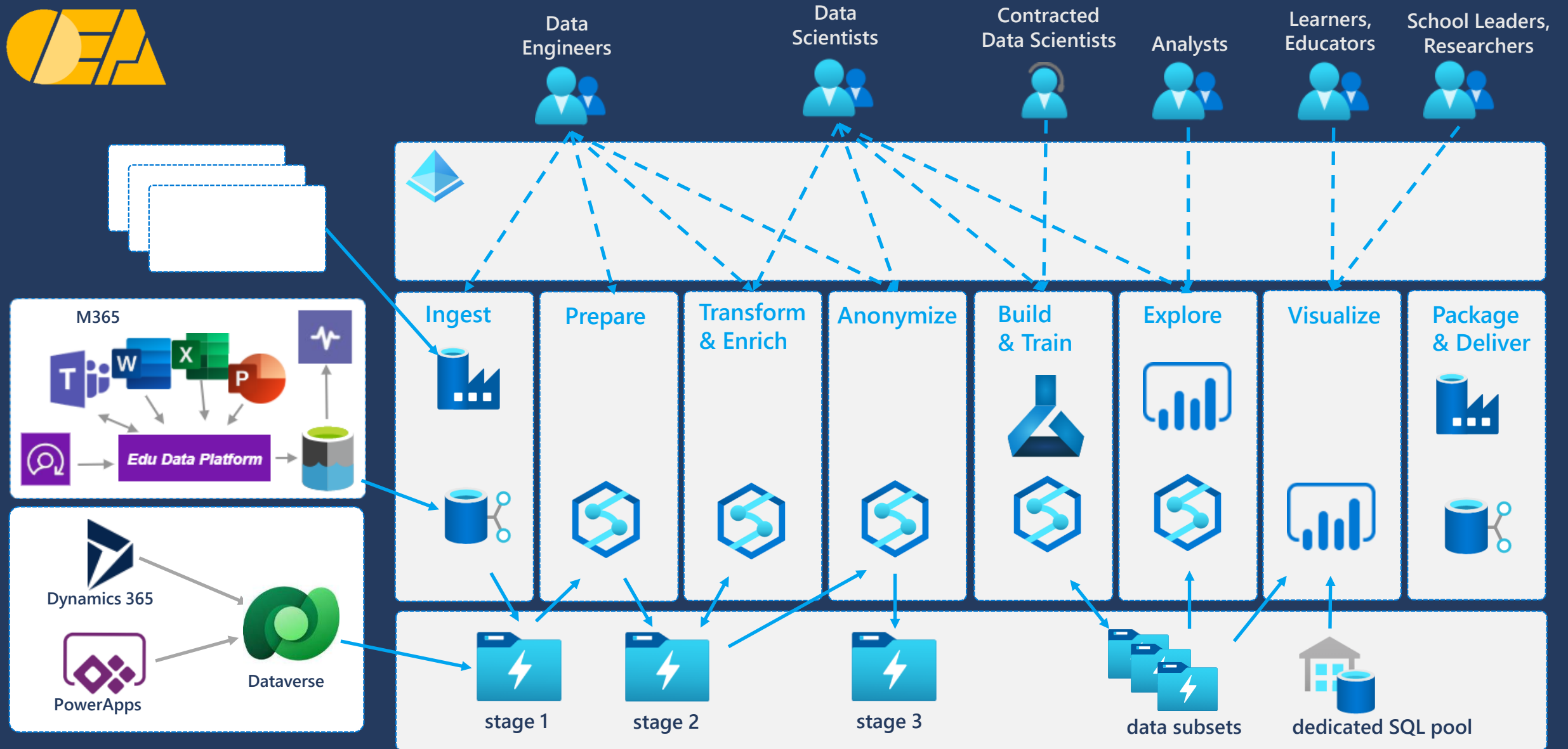
# Education Analytics and Insights



SIS



Customer's Azure Subscription



1) Collect

2) Compute

3) Communicate

# Multi-tenant architecture

