**TASK REPORT**

1. What is Microsoft Azure?

**Microsoft Azure** is a cloud computing service from Microsoft. It offers SaaS, PaaS, IaaS and supports multiple programming languages and tools.With Azure , businesses can run software applications, services, and hardware on the internet. This helps businesses avoid large expenditures on on-site hardware installation, operation and maintenance.

**Advantage:**

* High availability
* Flexible expansion capabilities
* Encrypt data anytime, anywhere
* Cost optimization
* Provides enterprise-grade development tools

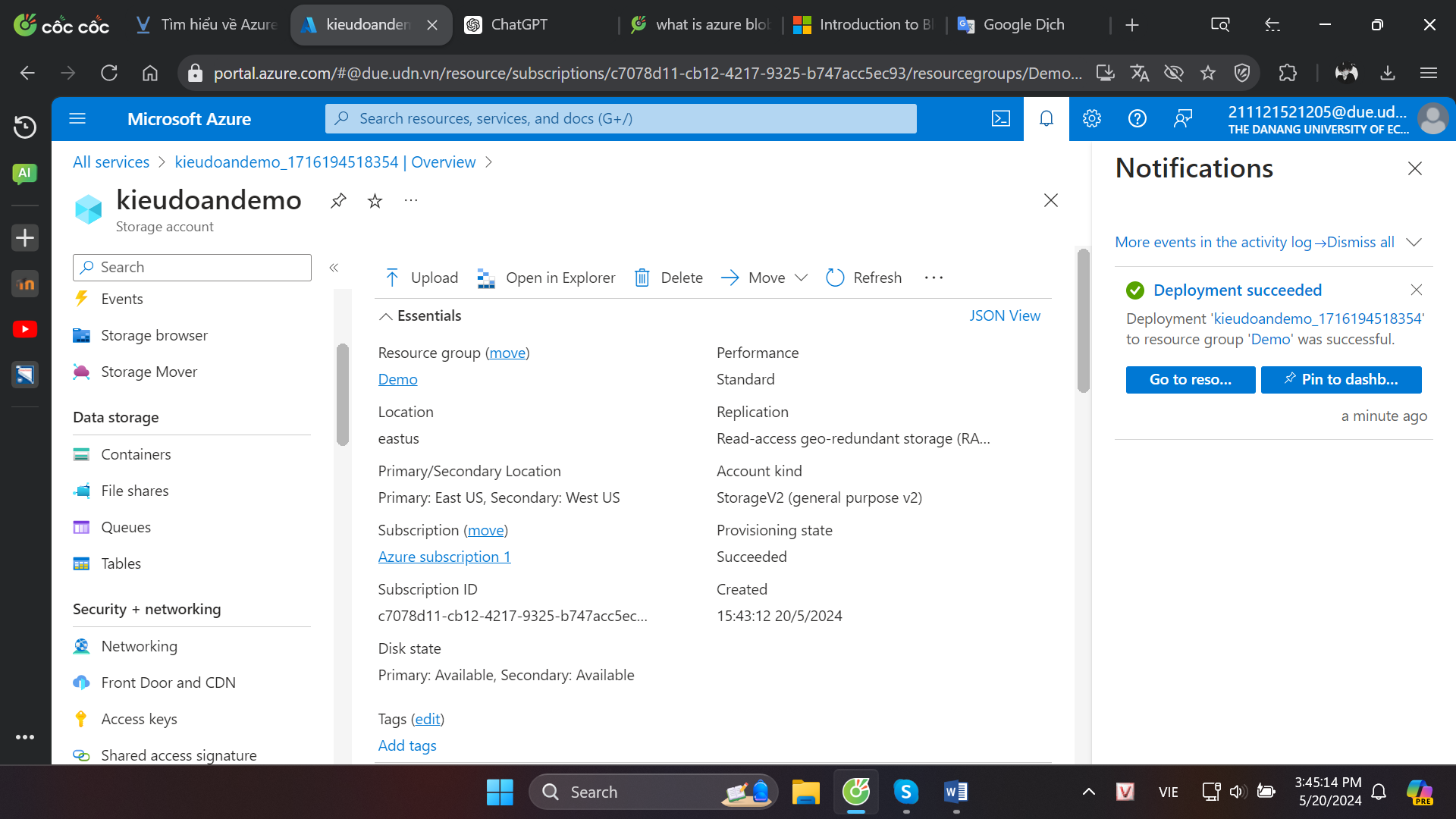
1. Azure blob storage

Azure Blob Storage is Microsoft's object storage solution for the cloud. Blob Storage is optimized for storing massive amounts of unstructured data.

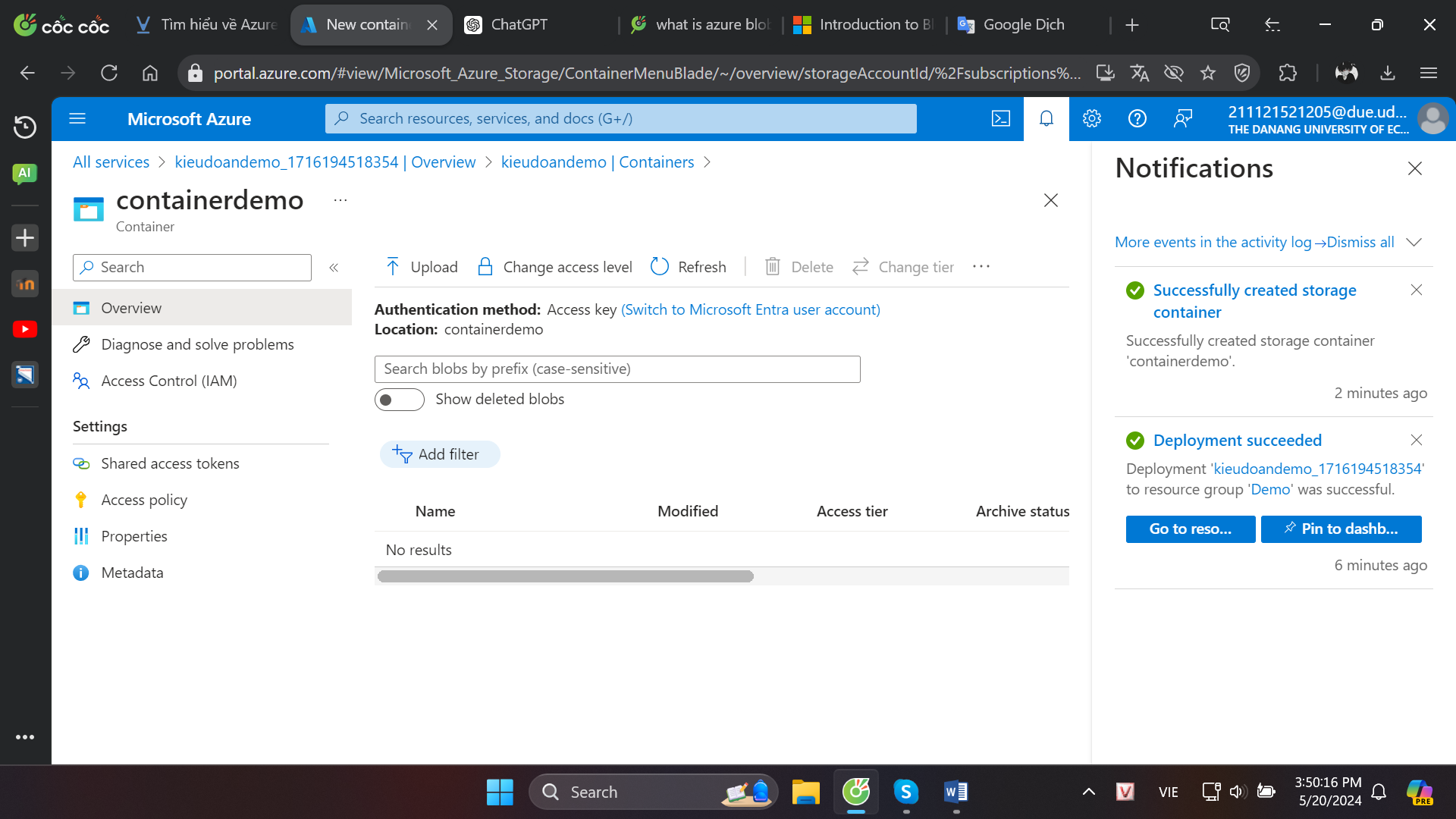
Azure storage supports 3 types of blob:

* Block blobs: stores text and binary data.
* Append blobs: ideal for writing data from virtual machines
* Page blobs: stores random access files up to 8 TB in size

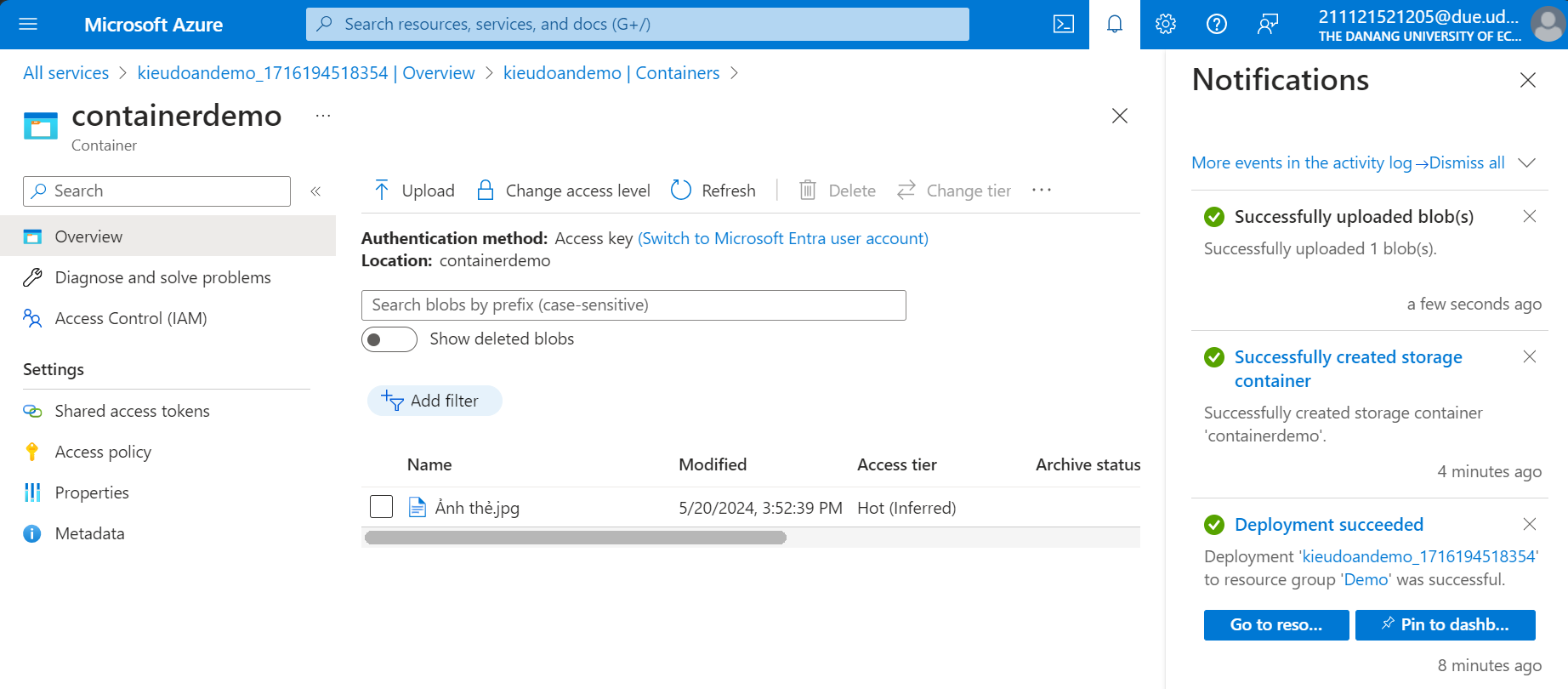
1. Create account



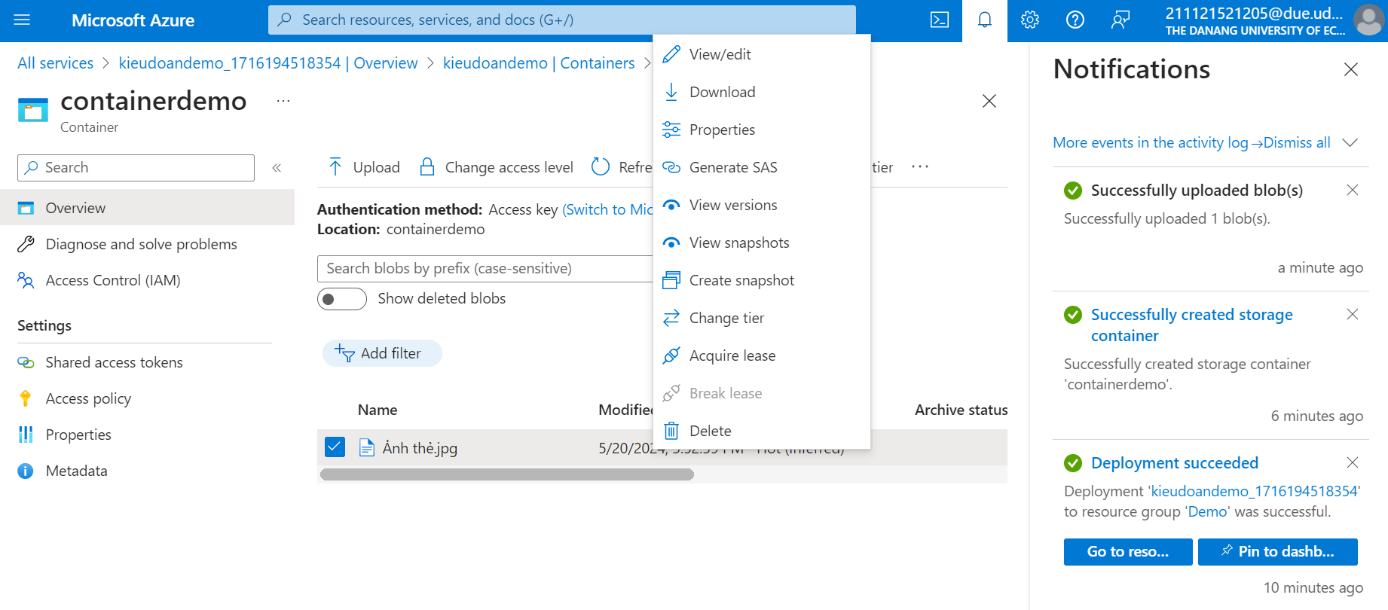
1. Create container



1. Upload file



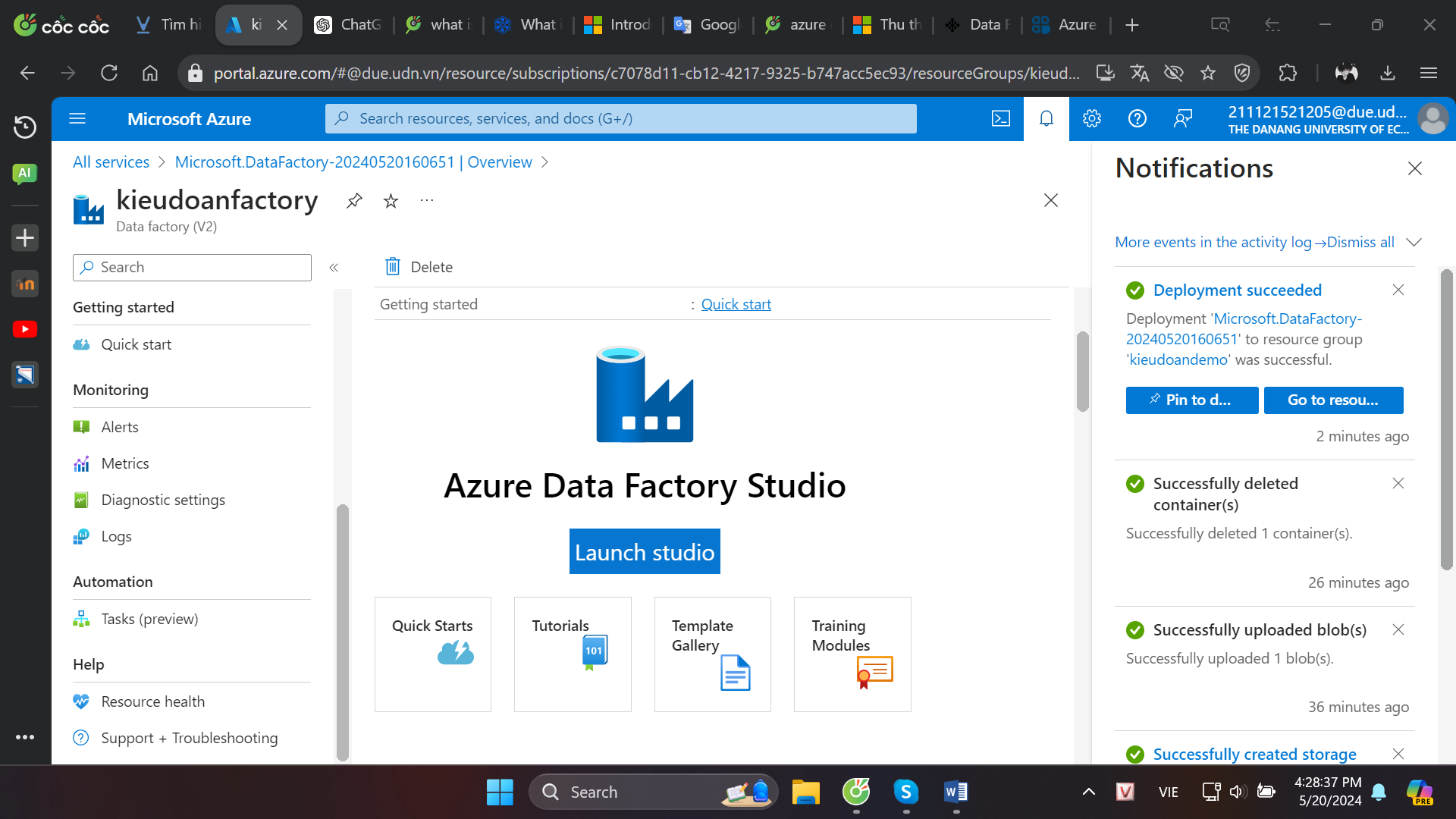
1. Operations with uploaded data



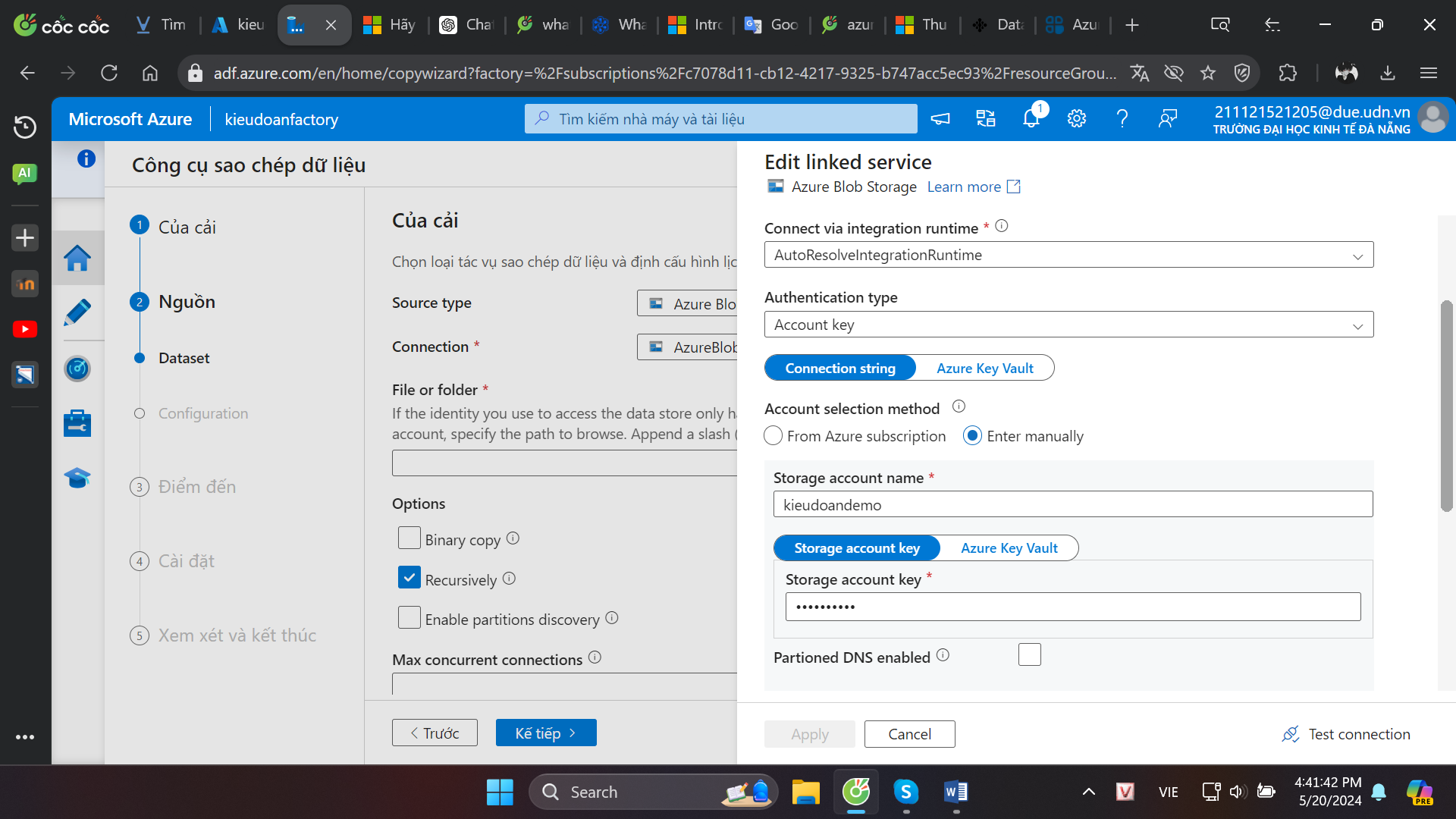
1. Azure datafactory

* Azure Data Factory is a cloud-based data integration service that allows you to create data-driven workflows in the cloud for orchestrating and automating data movement and data transformation.
* The Data Factory service allows you to create data pipelines that move and transform data and then run the pipelines on a specified schedule (hourly, daily, weekly, etc.)
* ADF can be used for:
* Supporting data migrations
* Getting data from a client’s server or online data to an Azure Data Lake
* Carrying out various data integration processes
* Integrating data from different ERP systems and loading it into Azure Synapse for reporting

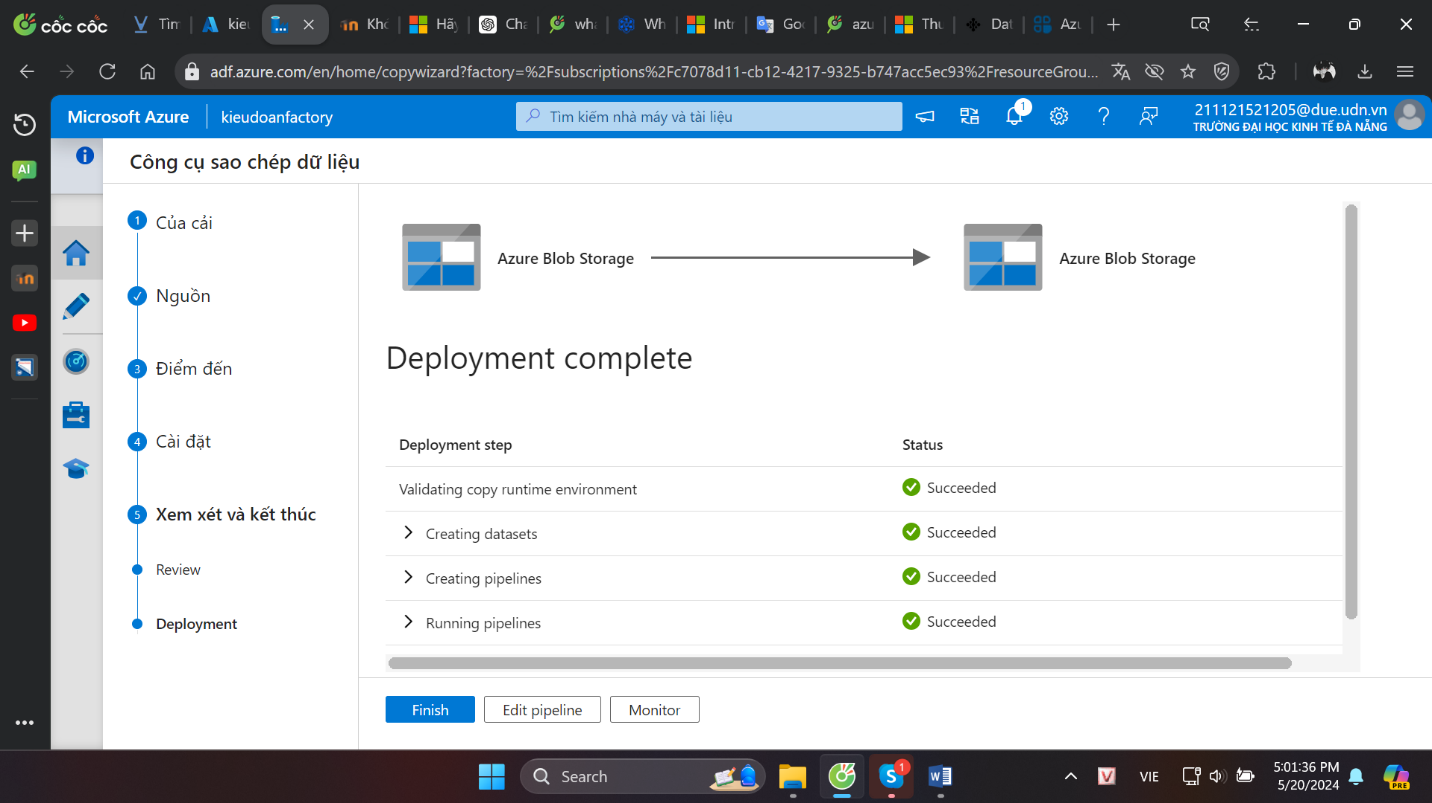
1. Create Data Factory

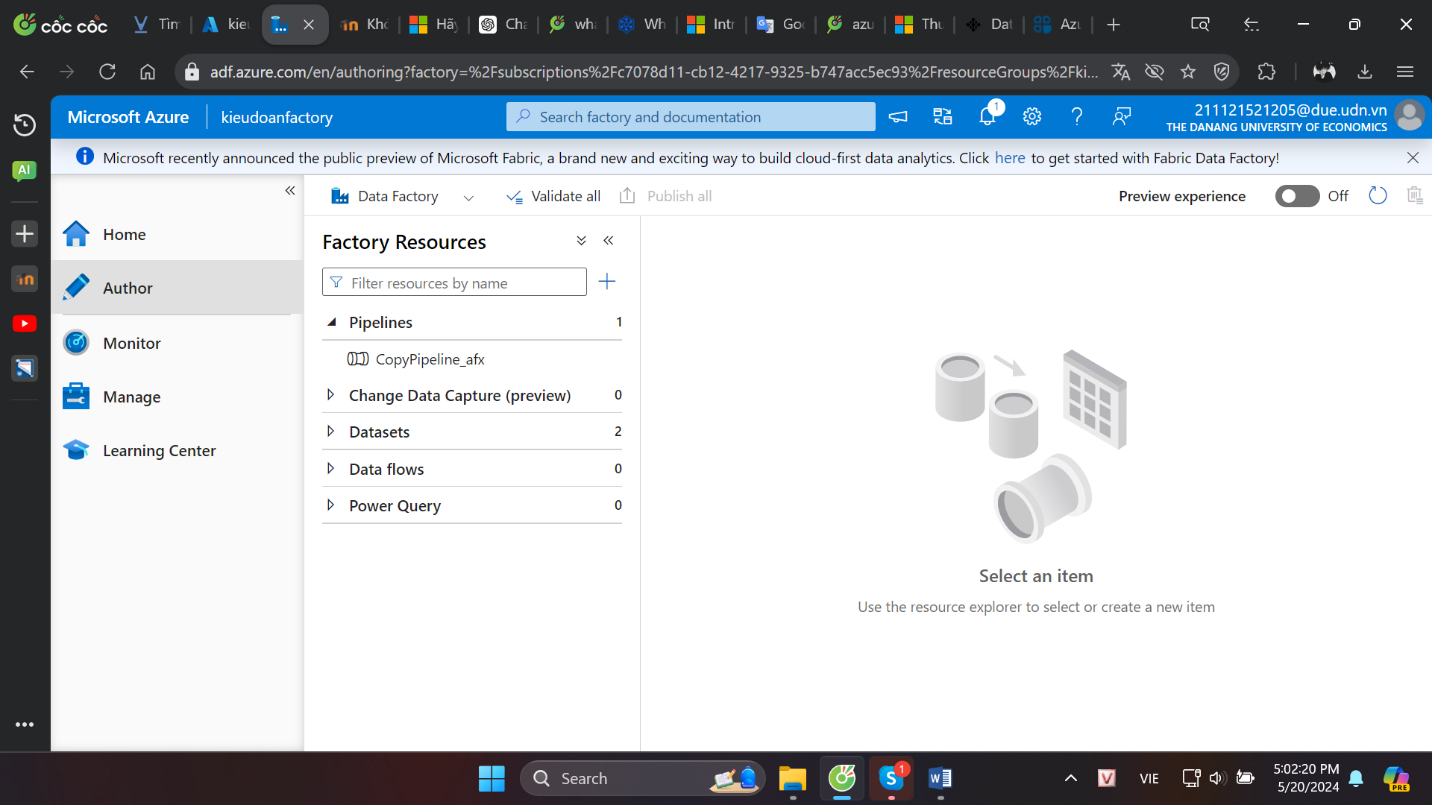


1. Create Linked Services

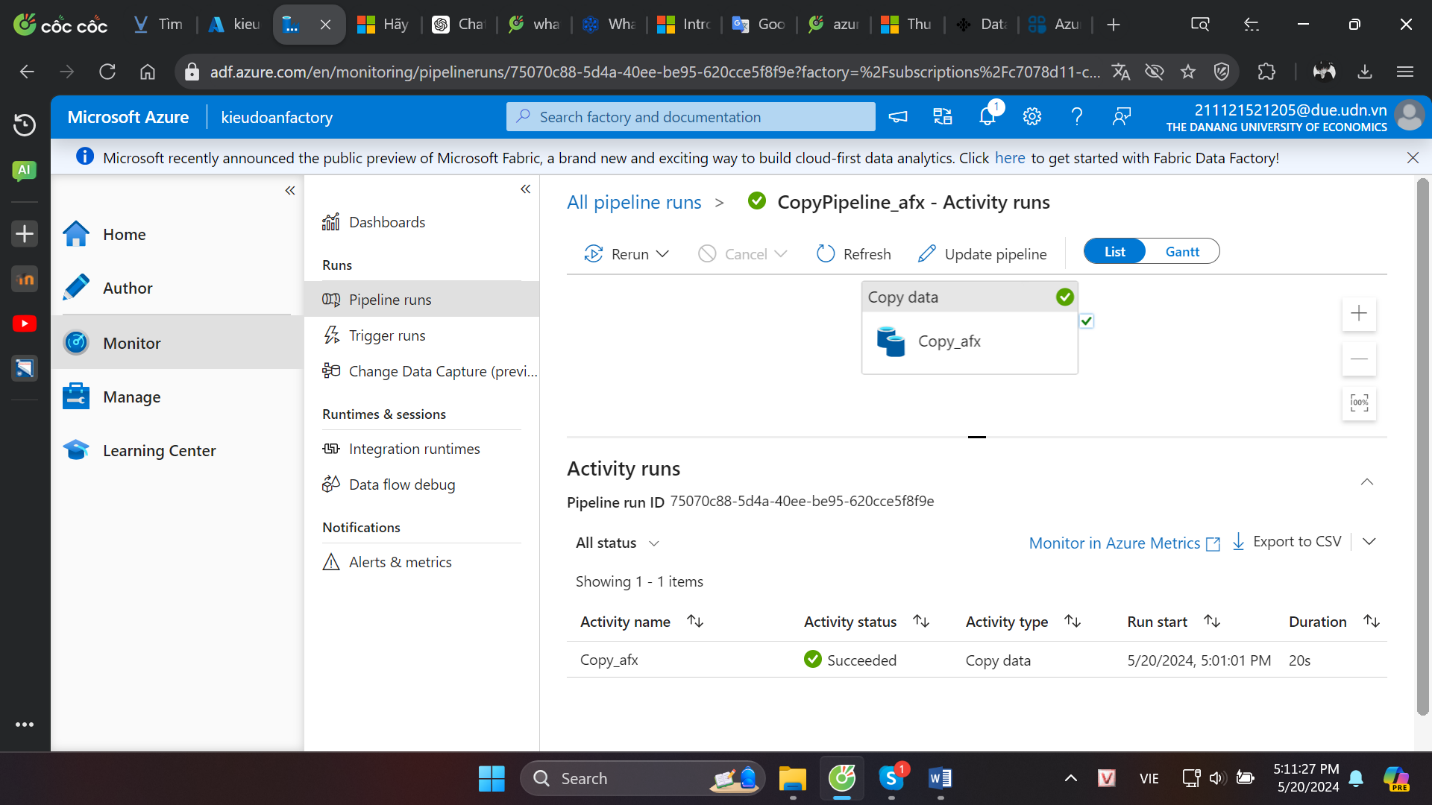


1. Create Pipelines



* 

1. Monitor the pipelines



1. Azure databricks

* Azure Databricks is a unified, open analytics platform for building, deploying, sharing, and maintaining enterprise-grade data, analytics, and AI solutions at scale. The Databricks Data Intelligence Platform integrates with cloud storage and security in your cloud account, and manages and deploys cloud infrastructure on your behalf.
* Azure Databricks cung cấp các công cụ giúp bạn kết nối nguồn dữ liệu của mình với một nền tảng để xử lý, lưu trữ, chia sẻ, phân tích, lập mô hình và kiếm tiền từ các tập dữ liệu bằng các giải pháp từ BI đến AI tổng hợp.
* The Azure Databricks workspace provides a unified interface and tools for most data tasks, including:
* Data processing scheduling and management, in particular ETL
* Generating dashboards and visualizations
* Managing security, governance, high availability, and disaster recovery
* Data discovery, annotation, and exploration
* Machine learning (ML) modeling, tracking, and model serving
* Generative AI solutions
* The Azure Databricks platform architecture comprises two primary parts:
* The infrastructure used by Azure Databricks to deploy, configure, and manage the platform and services.
* The customer-owned infrastructure managed in collaboration by Azure Databricks and your company.