



Optical Character Recognition Using Azure Services

Project Documentation

Created By: Ashwinkumar Naidu & Aryan jaiswar

ABSTRACT

The primary objective of this project is to develop a robust OCR solution that leverages Azure Virtual Machines (VMs), Azure Storage, and Virtual Network capabilities. The system is designed to process images, extract text content, and store the results in a secure and scalable manner.

Table of Contents

Chapter 1: Introduction	2
1.1. Project Overview	3
1.2. Objectives	3
1.3. Scope	3
Chapter 2: System requirements and specifications	4
2.1. Software Requirements	4
2.2. Hardware Requirements	4
Chapter 3: Azure Services Used	8
Chapter 4: Implementation	12
Chapter 5: Benefits	25
Chapter 6: Conclusion	26

Chapter 1

1.1 Project Overview

The project, titled "OCR using ASP.NET C# and Azure Services," aims to develop a robust Optical Character Recognition (OCR) system leveraging the capabilities of ASP.NET C# and various Azure services. This system is designed to efficiently extract text information from images, enabling businesses and users to digitize and process documents seamlessly.

Key Points:

- Introduction to the OCR project
- Brief description of the core functionalities
- Importance of implementing OCR in business processes

1.2 Objectives

The primary objectives of the OCR project are outlined to guide the development team and stakeholders toward achieving specific goals. These objectives are crafted to ensure the successful implementation of the OCR system with optimal performance and reliability.

Key Objectives:

- Develop a scalable and reliable OCR solution
- Integrate ASP.NET C# for efficient web-based application development
- Utilize Azure services for enhanced functionality and cloud-based capabilities
- Ensure accurate and fast text extraction from images
- Implement best practices for security, performance, and maintainability

1.3 Scope

The scope of the OCR project defines the boundaries and limitations within which the development team will operate. It outlines the features, functionalities, and components that will be included in the initial release of the OCR system.

Scope Highlights:

- Image-to-text conversion using OCR techniques
- Integration with Azure services like Availability Set, Virtual Network, Storage Account, Document Intelligence, and Recovery Services Vault
- Implementation of a user-friendly Desktop interface using ASP.NET C#
- Support for a predefined set of document types and formats
- Testing and validation of OCR accuracy under various scenarios

These sections provide an initial understanding of the project, its objectives, and the defined scope. Feel free to further customize and elaborate on each point based on the specific details and goals of your OCR project.

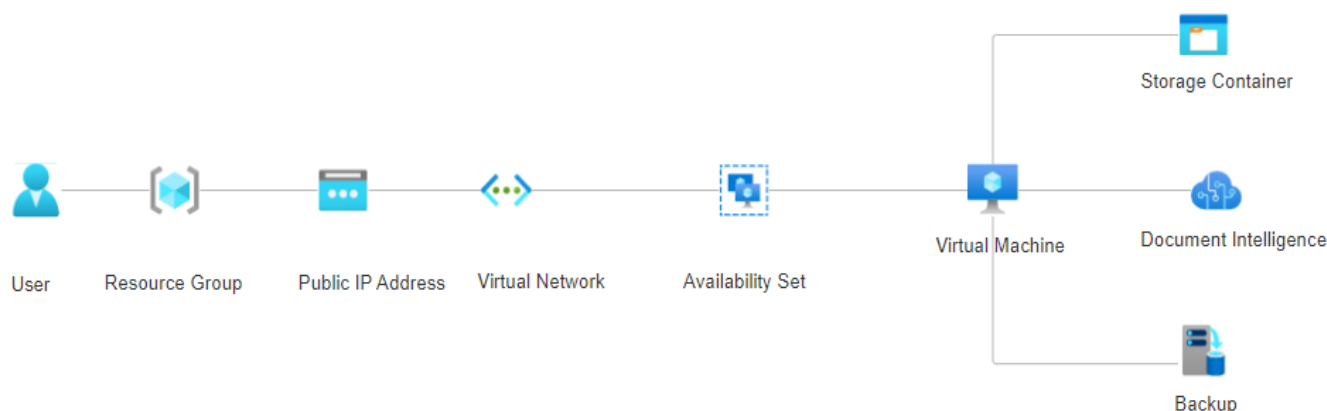
Chapter 2

2.1 Software Requirements:

- Operating system: Windows 10/11 or MAC OS.
- Platform: Microsoft Azure
- Microsoft azure subscription (Free Trial or Azure for student or Pay-as-you-go)
- Visual Studio 2022 (For coding in C# and designing GUI)
- Virtual machine OS: Windows 11 version 22H2

2.2 Hardware Requirements:

- Processor: Intel core i3 and above
- Hard disk: 256 GB or above
- RAM: 8GB or above
- Internet: 1 Mbps or above
- Virtual machine ram: 8GB
- Virtual machine storage: 30Gb



3.1 Availability Set: An Availability Set in Azure is a logical grouping of virtual machines (VMs) that helps ensure high availability. VMs within an Availability Set are placed in separate fault domains and update domains, minimizing the risk of simultaneous failures during planned or unplanned maintenance events. This architectural choice enhances the overall resilience of applications deployed on Azure, providing increased availability and reliability.

3.2 Virtual Network: Azure Virtual Network allows you to create private, isolated networks in the Azure cloud. It enables the segmentation of resources, facilitates secure communication between virtual machines, and connects on-premises networks to the cloud. Virtual Network provides the foundation for various Azure services, offering features like subnetting, network security groups, and virtual private network (VPN) connectivity for establishing secure connections between different environments.

3.3 Storage Account: Azure Storage Account is a scalable and secure cloud storage solution. It supports various types of storage services such as Blob, File, Queue, and Table storage. Blob storage, in particular, is commonly used for storing large amounts of unstructured data like images, videos, and documents. Azure Storage Account provides redundancy options, security features, and easy integration with other Azure services, making it a versatile choice for data storage needs.

3.4 Document Intelligence: Azure Document Intelligence is a set of AI services designed to analyze and extract valuable insights from unstructured documents. Using features like Optical Character Recognition (OCR), Document Intelligence can extract text, key-value pairs, and tables from images and scanned documents. This service is valuable for applications dealing with document processing, content extraction, and information retrieval, enhancing the efficiency of data extraction workflows.

3.5 Recovery Services Vault: Azure Recovery Services Vault is a cloud-based service that facilitates backup and disaster recovery for Azure resources. It allows you to protect your virtual machines, databases, and other workloads by creating backup policies and recovery plans. The Recovery Services Vault also supports features like site recovery, ensuring business continuity in the event of data loss or system failures. It provides a centralized management platform for

implementing robust backup and recovery strategies.

3.1 Architecture Overview

Azure Components:

- Availability Set: Enhances application resilience by grouping VMs to distribute outages and planned maintenance.
- Virtual Network: Isolates the application's network for security and control over connectivity.
- Storage Account: Stores documents for OCR processing and processed text for retrieval.
- Document Intelligence: Provides OCR capabilities for recognizing text in images and documents.
- Recovery Services Vault: Implements backup and disaster recovery for application components and data.

Application Structure:

- ASP.NET C# Web Application: Frontend for user interaction and document uploads.
- Document Intelligence Processing Logic: C# code handles document submission, OCR processing, and text storage.

Communication:

- Azure SDKs and APIs: Facilitate communication between the application and Azure services.

3.2 Key Components

Azure Services:

- Document Intelligence: Core component for OCR functionality.
- Storage Account: Persistence for document storage and retrieval.

Application Components:

- Web UI: Enables user interaction, document uploads, and results display.
- OCR Processing Logic: Manages document submission, interacts with Document Intelligence, and stores processed text.

3.3 Code Structure

Desktop Application:

- Controllers: Handle user requests, initiate OCR processing, and manage document flow.
- Views: Present user interface for document uploads and results.

OCR Processing Logic:

- Classes/Methods: Encapsulate interaction with Azure services and OCR workflow.

4.1 Purpose

An Availability Set in Azure is a logical grouping of virtual machines (VMs) that helps ensure high availability and resilience to failure. The primary purpose of configuring an Availability Set is to distribute VM instances across different physical hardware in a data center. By doing so, you mitigate the risk of a single point of failure and enhance the overall reliability of your application.

Key purposes include:

- **Fault Tolerance:** Distributing VM instances across multiple fault domains and update domains helps protect your application from hardware failures and planned maintenance events.
- **High Availability:** In the event of a failure or maintenance operation, only one subset of your VM instances is affected, while the others continue to run, providing continuous service availability.

4.2 Configuration Steps

Configuring an Availability Set involves several steps to ensure proper distribution and fault tolerance. Below are the key configuration steps:

Step 1: Access the Azure Portal

Log in to the Azure Portal using your credentials.

Step 2: Navigate to Virtual Machines

- Select the target virtual machine for which you want to configure an Availability Set.
- Navigate to the "Settings" section of the VM.

Step 3: Configure Availability Set

- Choose "Availability" from the settings menu.
- Create a new Availability Set or select an existing one.
- Configure the number of fault domains and update domains based on your availability requirements.

Step 4: Save Configuration

Save the changes to apply the Availability Set configuration to the virtual machine.

4.3 Best Practices

Ensuring optimal performance and resilience requires adherence to best practices when configuring Availability Sets:

- **Distribute Across Fault and Update Domains:** Utilize the fault and update domains effectively to distribute VM instances across different physical servers and update sequences.
- **Use Standard Load Balancers:** Combine Availability Sets with Azure Standard Load Balancers for load distribution and additional fault tolerance.
- **Consider Application Affinity:** For applications with specific requirements, consider the affinity or anti-affinity rules when placing VM instances within the same or separate

Availability Sets.

- **Regularly Test Failover Scenarios:** Conduct regular tests to validate the failover and recovery capabilities provided by the configured Availability Set.
- **Leverage Proximity Placement Groups:** For workloads that require low-latency network connectivity, consider using Proximity Placement Groups in conjunction with Availability Sets.

5. Virtual Network Setup

Virtual networking is a crucial aspect of your OCR project, enabling secure communication between different components. This section covers the networking requirements, virtual network configuration, and security considerations.

5.1 Networking Requirements

In this subsection, outline the specific networking requirements for your OCR project. This may include:

- **Bandwidth Requirements:** 1MBPS
- **IP Addressing:** 1 Public IP Address
- **Subnetting:** Make a Subnet for the usage purpose, in this project we have used /24 (250 IP Address)

5.2 Virtual Network Configuration

This subsection should provide step-by-step guidance on setting up the virtual network for your OCR project. Include the following details:

- **Creating a Virtual Network:** The process of creating a virtual network in the Azure portal. Include screenshots or code snippets if applicable.
- **Subnet Configuration:** The creation and configuration of subnets within the virtual network. Specify the purpose of each subnet and any specific configuration settings.
- **Network Security Groups (NSG):** NSGs are utilized to control inbound and outbound traffic to network interfaces, VMs, and subnets. Include examples of NSG rules relevant to your project.
- **Azure VPN Gateway (if applicable):** To configure the Azure VPN Gateway.

5.3 Security Considerations

Security is paramount in any network setup. In this subsection, address security considerations related to the virtual network:

- **Firewall Configuration:** If applicable, the firewalls are configured within the virtual network. The rules governing traffic between different components.
- **Network Isolation:** The network isolation is achieved within the virtual network, emphasizing the principle of least privilege.
- **Encryption:** Outline any encryption mechanisms in place for data in transit and at rest within the virtual network.
- **Monitoring and Logging:** Network activities are monitored, and logs are maintained for auditing and security analysis.

6.1 Storage Account Introduction

In this section, provide a comprehensive introduction to the Azure Storage Account. Cover key concepts such as:

- **Storage Account Overview:** Azure Storage Account is, its core features, and the various types of data it can store.
- **Use Cases:** OCR project that necessitate the use of a Storage Account. For example, storing image files, text data, or any other relevant information.
- **Scalability and Performance:** The Storage Accounts provide scalability and performance benefits for your application.

6.2 Configuration Steps

This subsection should provide a step-by-step guide on how to set up a Storage Account for your OCR project. Include detailed instructions, screenshots, or code snippets where necessary. The steps might include:

- **Accessing Azure Portal:** Guide users on how to log in to the Azure Portal and navigate to the Storage Account creation section.
- **Creating a Storage Account:** Provide detailed steps on creating a new Storage Account. Include considerations such as choosing the storage account type (Standard vs. Premium), replication strategy, and selecting the appropriate performance tier.
- **Configuring Security:** Discuss any security considerations, such as setting up access control, managing keys, or integrating with Azure Active Directory for authentication.
- **Setting up Blob Storage:** If your OCR project involves storing images or binary data, guide users through the process of setting up Blob Storage within the Storage Account.

6.3 Blob Storage Usage

Delve into the specific usage of Blob Storage within your OCR project. Include information such as:

- **Blob Storage Overview:** The Blob Storage is and how it differs from other storage types within Azure Storage.
- **Uploading and Retrieving Data:** Provide code examples or detailed steps on how to upload files (e.g., images or documents) to Blob Storage and retrieve them when needed.
- **Accessing Blob Storage from ASP.NET C#:** If your ASP.NET C# application interacts with Blob Storage, include code snippets or guidelines on how to perform CRUD (Create, Read, Update, Delete) operations.
- **Best Practices:** Best practices or optimization tips for utilizing Blob Storage effectively within the context of your OCR application.

8. Recovery Services Vault

8.1 Purpose and Benefits

Purpose: A Recovery Services Vault in Azure is a management entity that stores recovery points created over time, providing data protection and recovery for various Azure services. Its primary purpose is to safeguard your data against loss or corruption, ensuring business continuity.

Benefits:

- **Backup and Restore Capability:** Recovery Services Vault enables regular backups and streamlined restores, reducing downtime in case of data loss or system failures.
- **Cross-Region Replication:** The vault supports cross-region replication, allowing you to replicate your data to a secondary region for additional resilience.
- **Application Consistency:** It ensures application consistency during backups, maintaining the integrity of the data.
- **Long-Term Retention:** You can configure long-term retention policies for your backups, complying with regulatory and compliance requirements.
- **Integration with Azure Services:** It seamlessly integrates with various Azure services, making it a versatile solution for comprehensive data protection.

8.2 Configuration Steps

Step 1: Create a Recovery Services Vault: In the Azure Portal, navigate to the Recovery Services Vaults section and create a new vault. Specify the settings such as the subscription, resource group, and vault name.

Step 2: Configure Backup Policies: Define backup policies based on your project requirements. This includes specifying the backup frequency, retention range, and the types of items to be backed up.

Step 3: Register Resources: Register the resources you want to protect with the Recovery Services Vault. This may include virtual machines, storage accounts, or any other relevant Azure services used in your OCR project.

Step 4: Install Azure Backup Agent (if applicable): For on-premises or IaaS (Infrastructure as a Service) scenarios, install the Azure Backup Agent on the relevant machines to facilitate communication with the Recovery Services Vault.

Step 5: Trigger Initial Backup: Initiate the initial backup to create the first recovery point. This may involve a full initial backup followed by incremental backups according to the defined policies.

8.3 Backup and Restore Strategies

Backup Strategies:

- **Incremental Backups:** Take advantage of incremental backups to reduce the amount of data transferred and storage requirements for each backup operation.
- **Regular Backup Schedule:** Establish a consistent and regular backup schedule to ensure that critical data is captured frequently.
- **Retention Policies:** Configure retention policies based on your project's needs, considering factors such as compliance requirements and data value.

Restore Strategies:

- **Point-in-Time Recovery:** Leverage point-in-time recovery options to restore data to a specific state, allowing for granular recovery.
- **Test Restores:** Periodically perform test restores to validate the backup process and ensure the recoverability of data.
- **Cross-Region Restore:** Understand and plan for cross-region restore capabilities for additional resilience and redundancy.

Chapter 4: Implementation

First Create a Resource Group.

Resource groups - Microsoft Azure

portal.azure.com/#view/HubsExtension/BrowseResourceGroups

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

Home > Resource groups

Default Directory (ashwinjanmagicbusoutlook.onmicrosoft.com)

+ Create Manage view Refresh Export to CSV Open query Assign tags

Filter for any field... Subscription equals all Location equals all Add filter

Showing 1 to 2 of 2 records. No grouping List view

Name	Subscription	Location
NetworkWatcherRG	Free Trial	Central India
OCR	Free Trial	Central India

< Previous Page 1 of 1 Next >

Give feedback

Created one Subnet.

new_NTconfig - Microsoft Azure

portal.azure.com/#@ashwinjanmagicbusoutlook.onmicrosoft.com/resource/subscriptions/3ca8fc99-4b10-4d5c-9fad-489553f3420c/resourceGroups/OCR/providers/Microsoft.Network/virtualNetworks/...

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

Home > Virtual networks > new_NTconfig

Virtual networks

Default Directory (ashwinjanmagicbusoutlook.onmicrosoft.com)

+ Create Manage view

Filter for any field... Name

new_NTconfig

new_NTconfig | Subnets

Virtual network

+ Subnet + Gateway subnet Refresh Manage users Delete

Search subnets

Name	IPv4	IPv6	Available IPs	Delegated to	Security group	Route table
default	10.63.240.0/24	-	250	-	-	-

< Previous Page 1 of 1 Next >

Give feedback

Next Created 1 VM size D2as_v4

Virtual machines - Microsoft Azure

portal.azure.com/#view/HubsExtension/BrowseResource/resourceType/Microsoft.Compute%2FVirtualMachines

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

ashwinjanmagicbus@ou... DEFAULT DIRECTORY (ASHWINJA...)

Home

Virtual machines

Default Directory (ashwinjanmagicbusoutlook.onmicrosoft.com)

+ Create Switch to classic Reservations Manage view Refresh Export to CSV Open query Assign tags Start Restart Stop Delete Services Maintenance

Filter for any field... Subscription equals all Type equals all Resource group equals all Location equals all Add filter

Showing 1 to 1 of 1 records. No grouping List view

Name	Type	Subscription	Resource group	Location	Status	Operating system	Size	Public IP address	Disks
OCR-DOC	Virtual machine	Free Trial	OCR	Central India	Running	Windows	Standard_D2as_v4	4.188.231.123	1

< Previous Page 1 of 1 Next >

Give feedback

Windows File Explorer Edge Word PowerPoint OneDrive

ENG US 2:19 PM 1/9/2024

Created 1 Availability set for fault Tolerance

Availability sets - Microsoft Azure

portal.azure.com/#view/HubsExtension/BrowseResource/resourceType/Microsoft.Compute%2FavailabilitySets

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

ashwinjanmagicbus@ou... DEFAULT DIRECTORY (ASHWINJA...)

Home

Availability sets

Default Directory (ashwinjanmagicbusoutlook.onmicrosoft.com)

+ Create Manage view Refresh Export to CSV Open query Assign tags

Filter for any field... Subscription equals all Resource group equals all Location equals all Add filter

Showing 1 to 1 of 1 records. No grouping List view

Name	Resource group	Location	Subscription
new_av_set	OCR	Central India	Free Trial

< Previous Page 1 of 1 Next >

Give feedback

Windows File Explorer Edge Word PowerPoint OneDrive

ENG US 2:21 PM 1/9/2024

Created 1 Storage Account

Storage accounts - Microsoft A

portal.azure.com/#view/HubsExtension/BrowseResource/resourceType/Microsoft.Storage%2FStorageAccounts

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

ashwinjanmagicbus@ou...
DEFAULT DIRECTORY (ASHWINJA...)

Home >

Storage accounts

Default Directory (ashwinjanmagicbusoutlook.onmicrosoft.com)

Create Restore Manage view Refresh Export to CSV Open query Assign tags Delete

Filter for any field... Subscription equals all Resource group equals all Location equals all Add filter

Showing 1 to 1 of 1 records.

No grouping List view

Name	Type	Kind	Resource group	Location	Subscription
newstorageazure	Storage account	StorageV2	OCR	Central India	Free Trial

< Previous Page 1 of 1 Next >

Give feedback

Windows Taskbar

ENG US 2:21 PM 1/9/2024

Created 1 File Share Account for Mounting inside VM

newstorageazure - Microsoft A

portal.azure.com/#@ashwinjanmagicbusoutlook.onmicrosoft.com/resource/subscriptions/3ca8fc99-4b10-4d5c-9fad-489553f3420c/resourceGroups/OCR/providers/Microsoft.Storage/storageAccounts/...

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

ashwinjanmagicbus@ou...
DEFAULT DIRECTORY (ASHWINJA...)

Home > Storage accounts > newstorageazure

Storage accounts

Default Directory (ashwinjanmagicbusoutlook.on...

Create Restore

Filter for any field...

Name

newstorageazure

newstorageazure | File shares

Storage account

Search File share Refresh

File share settings

Identity-based access: Not configured Default share-level permissions: Disabled Soft delete: 14 days Maximum capacity: 5 TiB

Security: Maximum compatibility

Search file shares by prefix (case-sensitive) Show deleted shares

Name	Modified	Tier	Quota
newstorageaccount	1/5/2024, 12:09:12 AM	Cool	5 TiB

< Page 1 of 1 >

Windows Taskbar

ENG US 2:22 PM 1/9/2024

Created 1 Document Intelligence .

Microsoft Azure

Upgrade

Search resources, services, and docs (G+)

ashwinjanmagicbus@ou...
DEFAULT DIRECTORY (ASHWINJA...)

Home > Azure AI services

Azure AI services | Document intelligence

Azure AI services

Search

Create

Manage commitment plans

Manage deleted resources

Manage view

Refresh

Export to CSV

Open query

Assign tags

Delete

Overview

All Azure AI services

Azure AI services

Azure AI services

Azure OpenAI

AI Search

Computer vision

Face API

Custom vision

Speech service

Language service

Translator

Document intelligence

Bot services

Anomaly detector

Content moderator

Health Insights

Personalizer

Filter for any field...

Subscription equals all

Type equals all

Resource group equals all

Location equals all

Add filter

Showing 1 to 1 of 1 records.

No grouping

List view

Name	Kind	Location	Custom Domain	Pricing tier	Status	Created date
OCRMAGICBUS	FormRecognizer	Central India	ocrmagicbus	F0	Succeeded	2024-01-01T18:50:46.4...

< Previous

Page 1 of 1

Next >

Give feedback

Copied Both Key1 & Endpoints .

Microsoft Azure

Upgrade

Search resources, services, and docs (G+)

ashwinjanmagicbus@ou...
DEFAULT DIRECTORY (ASHWINJA...)

Home > Azure AI services | Document intelligence > OCRMAGICBUS

OCRMAGICBUS | Keys and Endpoint

Document intelligence

Search

Regenerate Key1

Regenerate Key2

Overview

Activity log

Access control (IAM)

Tags

Diagnose and solve problems

Resource Management

Keys and Endpoint

Encryption

Pricing tier

Networking

Identity

Cost analysis

Properties

Locks

Monitoring

Alerts

Metrics

Diagnostic settings

These keys are used to access your Azure AI service API. Do not share your keys. Store them securely-- for example, using Azure Key Vault. We also recommend regenerating these keys regularly. Only one key is necessary to make an API call. When regenerating the first key, you can use the second key for continued access to the service.

Hide Keys

KEY 1

b848d168793b413d8333eea723983499

KEY 2

d67e62468eea497bab4ed0824e4dd5a2

Location/Region

centralindia

Endpoint

https://ocrmagicbus.cognitiveservices.azure.com/

Created 1 Recovery Service Vault.

Recovery Services vaults - Microsoft Azure

portal.azure.com/#view/HubsExtension/BrowseResource/resourceType/Microsoft.RecoveryServices%2FVaults

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

Home > Recovery Services vaults

Default Directory (ashwinjanmagicbusoutlook.onmicrosoft.com)

+ Create Manage view Refresh Export to CSV Open query Assign tags

Filter for any field... Subscription equals all Resource group equals all Location equals all Immutability equals all Add filter

Showing 1 to 1 of 1 records.

Name	Type	Immutability	Resource group	Location	Subscription
vault-lqzjs4v1	Recovery Services vault	Disabled	OCR	Central India	Free Trial

< Previous Page 1 of 1 Next >

Give feedback

Now Taking VM access via Public IP address

OCR-DOC - Microsoft Azure

portal.azure.com/#@ashwinjanmagicbusoutlook.onmicrosoft.com/resource/subscriptions/3ca8fc99-4b10-4d5c-9fad-489553f3420c/resourceGroups/OCR/providers/Microsoft.Compute/virtualMachines/OCR-DOC

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

Home > Virtual machines > OCR-DOC

Virtual machine

Filter for any field... Name

Overview Activity log Access control (IAM) Tags Diagnose and solve problems

Connect Connect Bastion

Networking Network settings Load balancing Application security groups Network manager

Settings Disks Extensions + applications Configuration Advisor recommendations

Enter your credentials

These credentials will be used to connect to 4.188.231.123.

azureadmin

Remember me

More choices

OK Cancel

Subscription ID 3ca8fc99-4b10-4d5c-9fad-489553f3420c

Tags (edit) Add tags

Properties Monitoring Capabilities (7) Recommendations Tutorials

Virtual machine

Computer name	OCR-DOC
Operating system	Windows (Windows 11 Pro)
Image publisher	microsoftwindowsdesktop
Image offer	windows-11
Image plan	win11-22h2-pro

Networking

Public IP address	4.188.231.123 (Network interface ocr-doc144)
Public IP address (IPv6)	-
Private IP address	10.63.240.4
Private IP address (IPv6)	-

Operating system Windows (Windows 11 Pro)

Size Standard

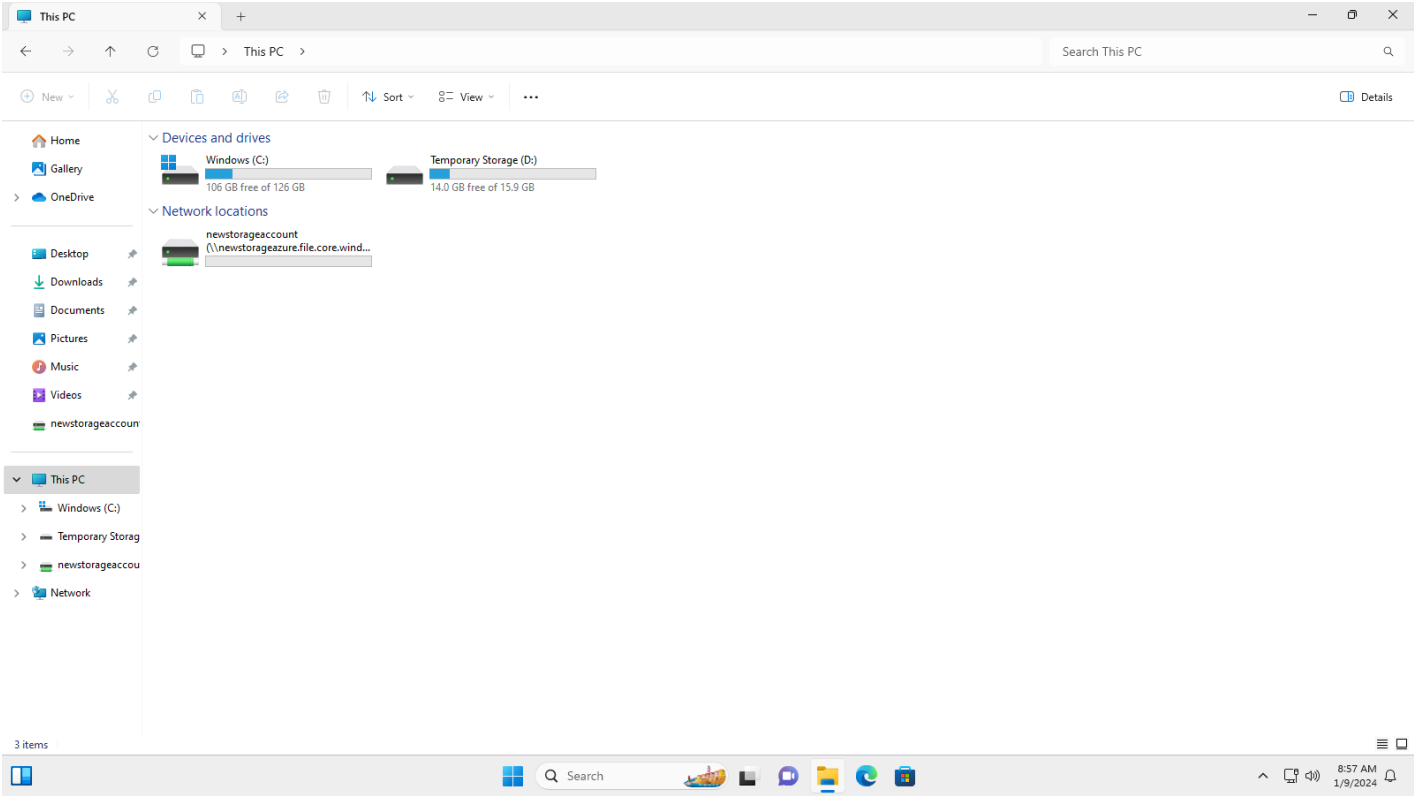
Public IP address 4.188.231.123

Virtual network/subnet new_NTconfig/default

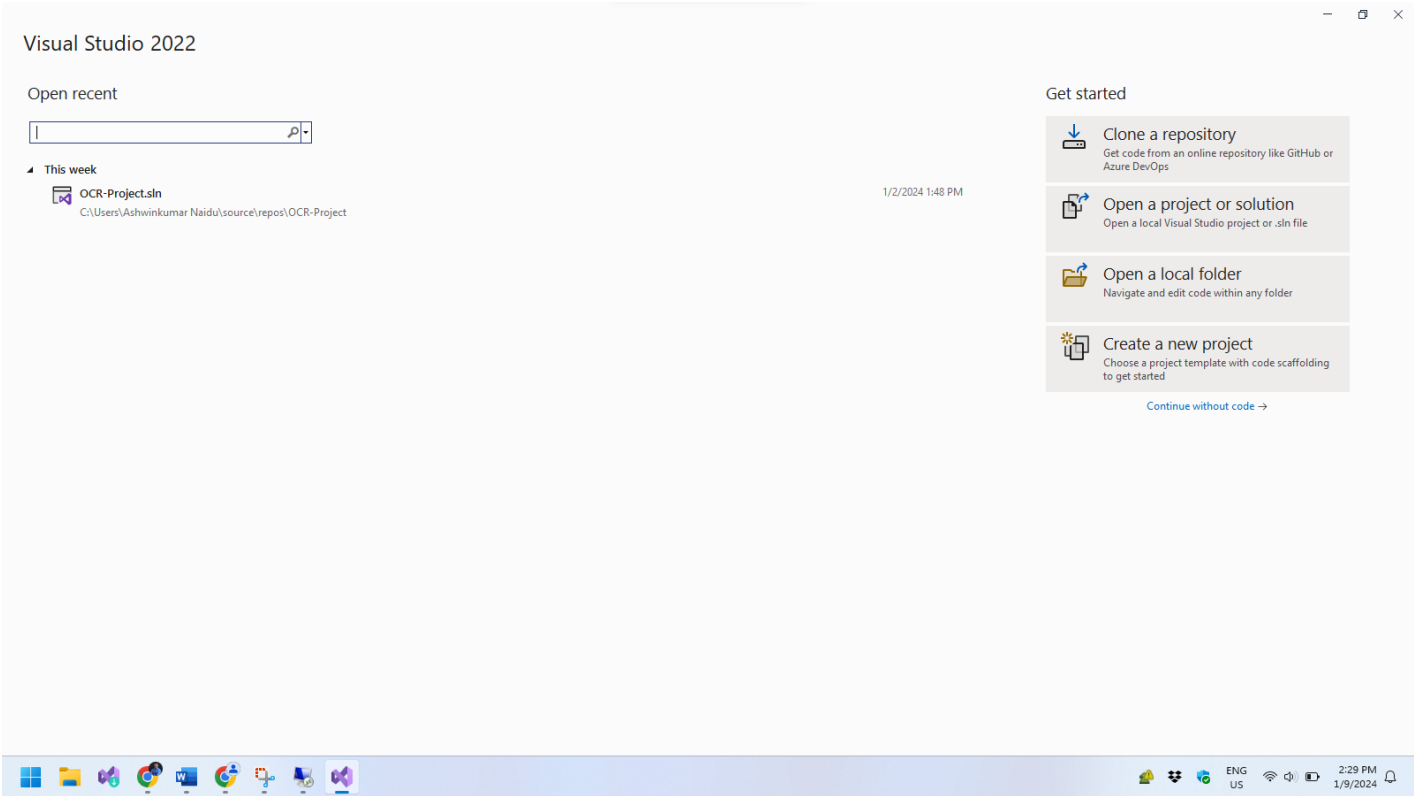
DNS name Not configured

Health state -

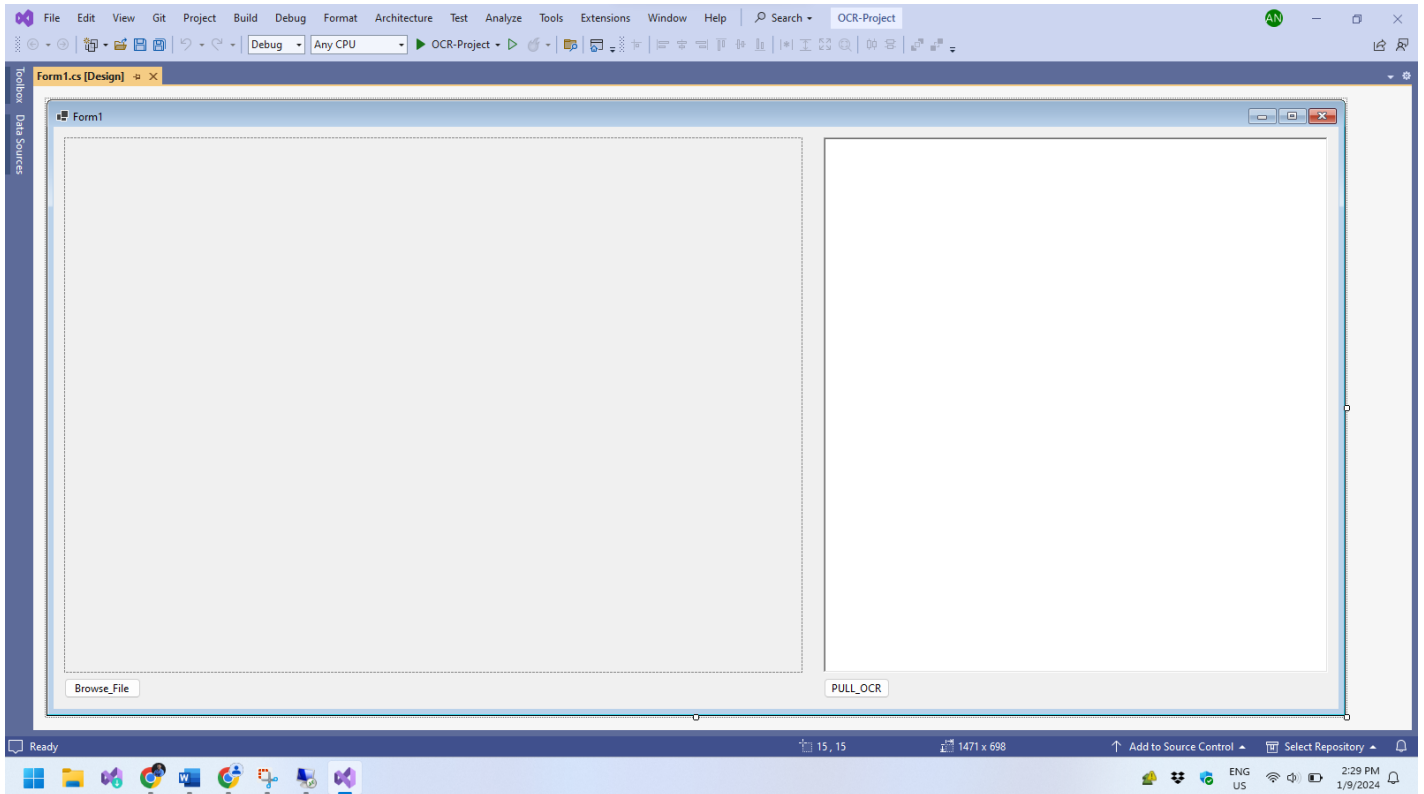
Successfully Mounted Storage Account



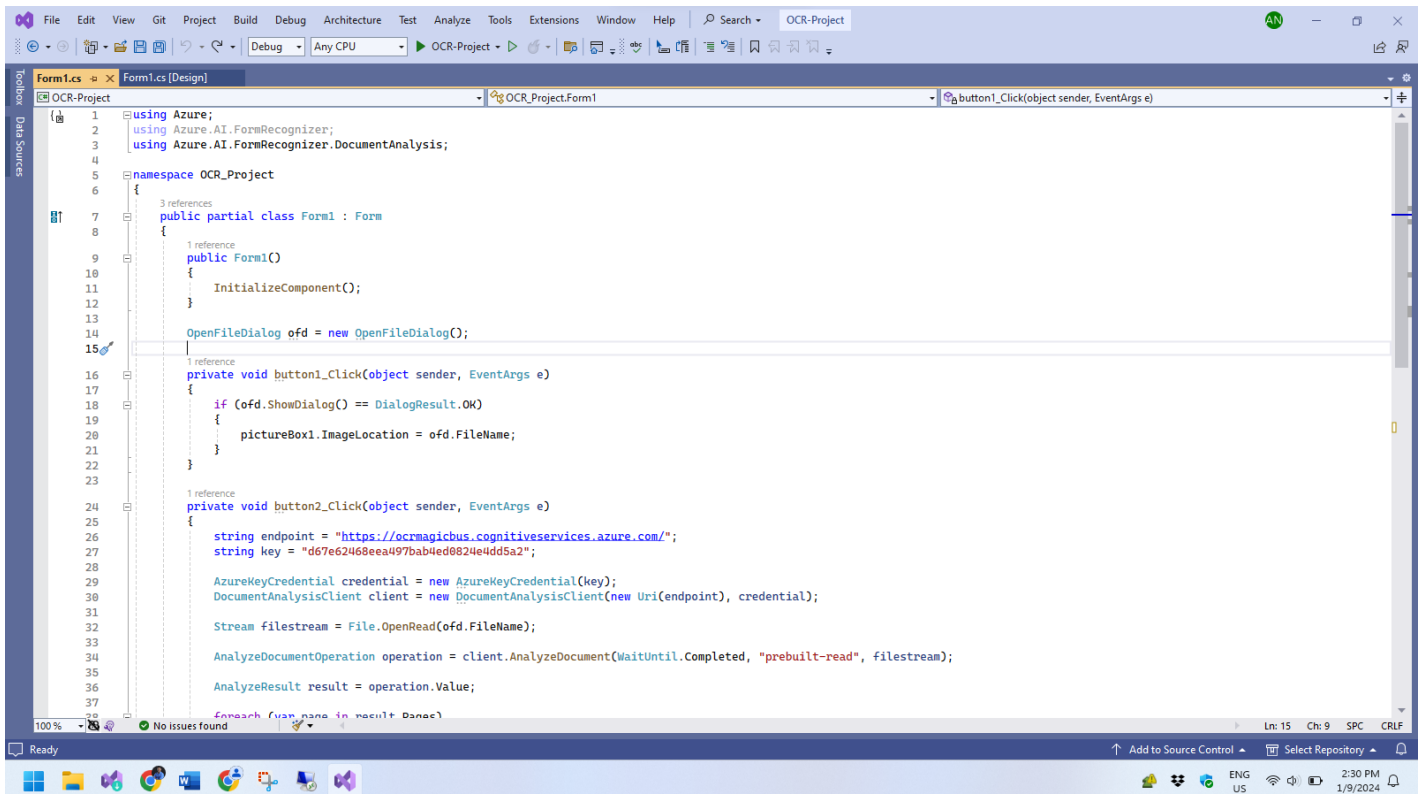
Open Visual Studio 2022 for creating OCR Project

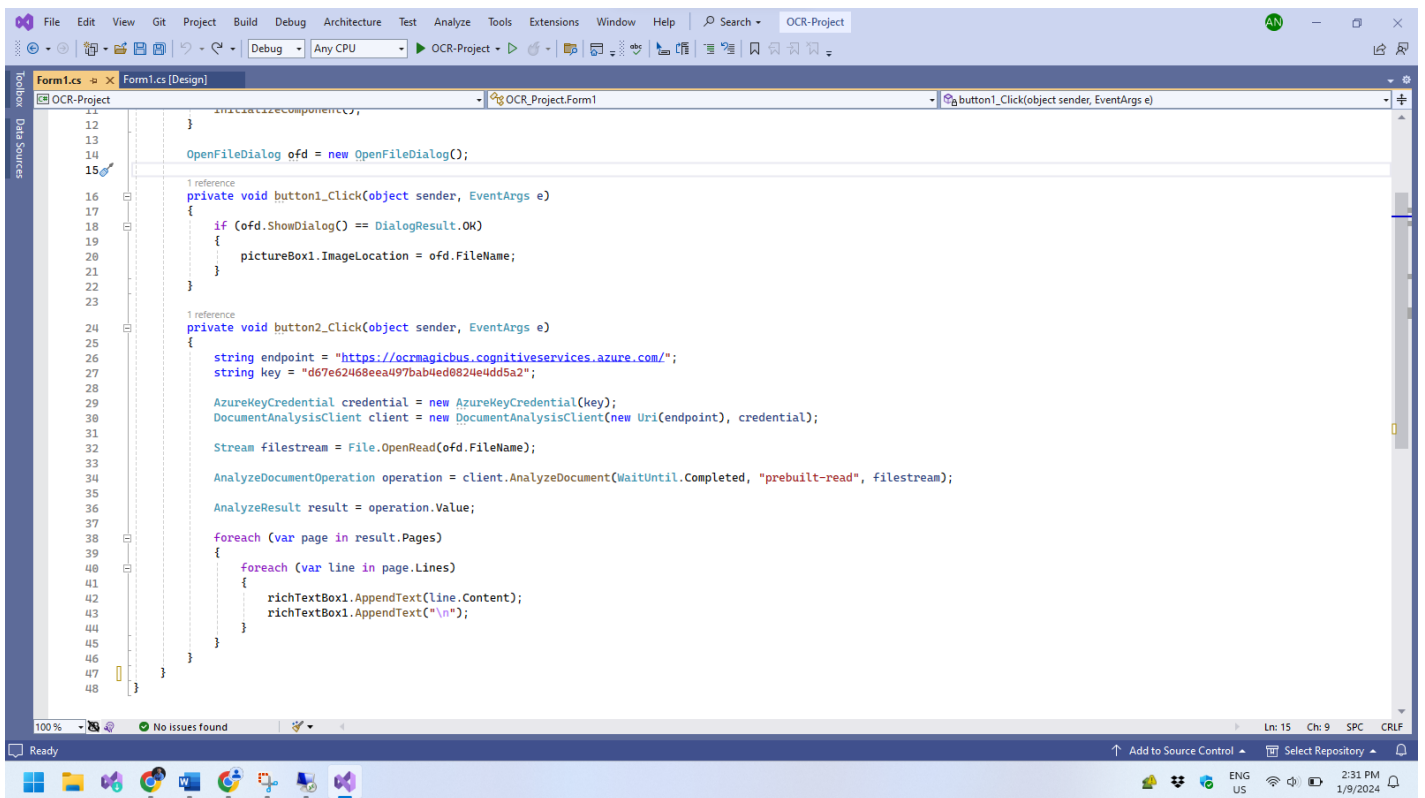


Now Create an GUI Design for User Interface

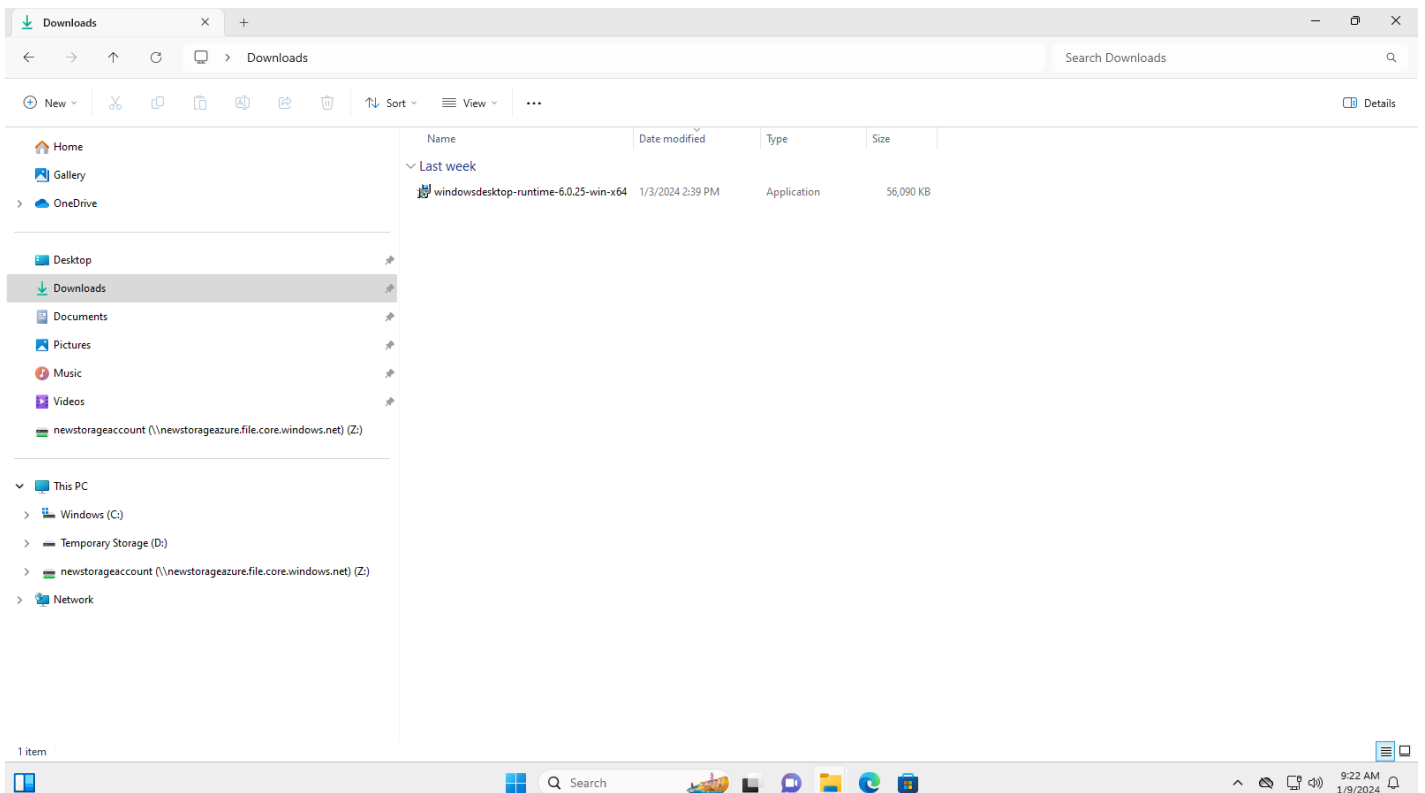


Write the Coding's on C# and add Endpoint & Key Value

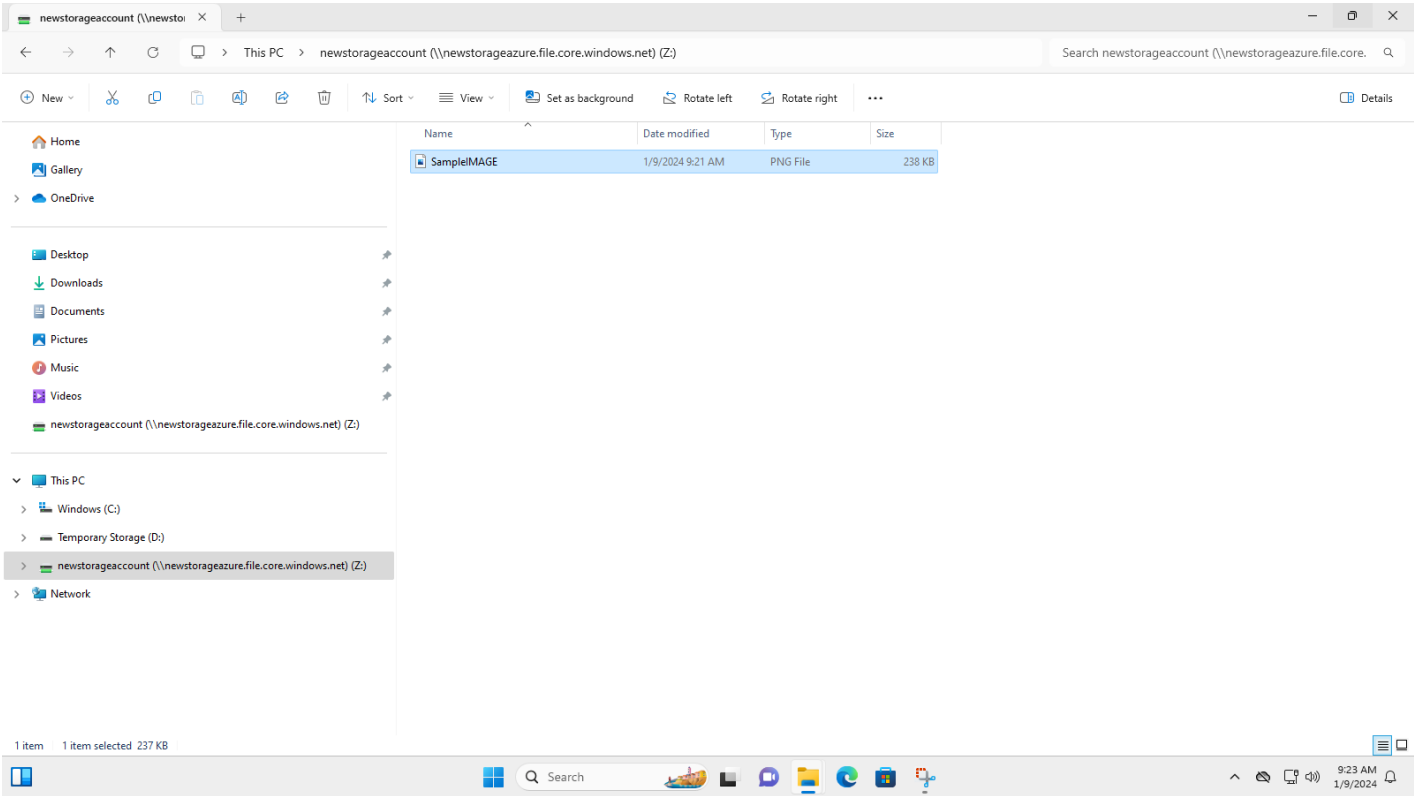




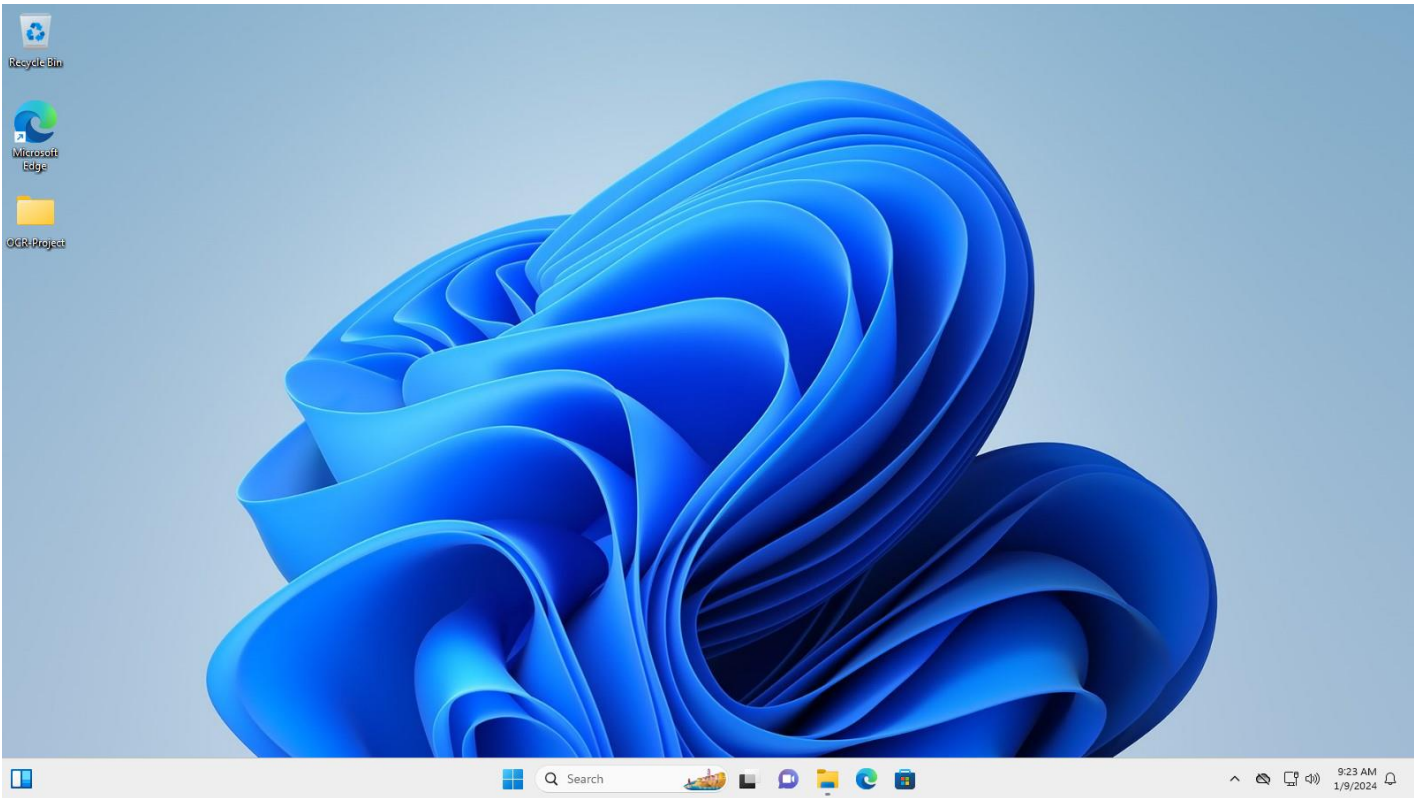
Install .Net Runtime to run the application on VM



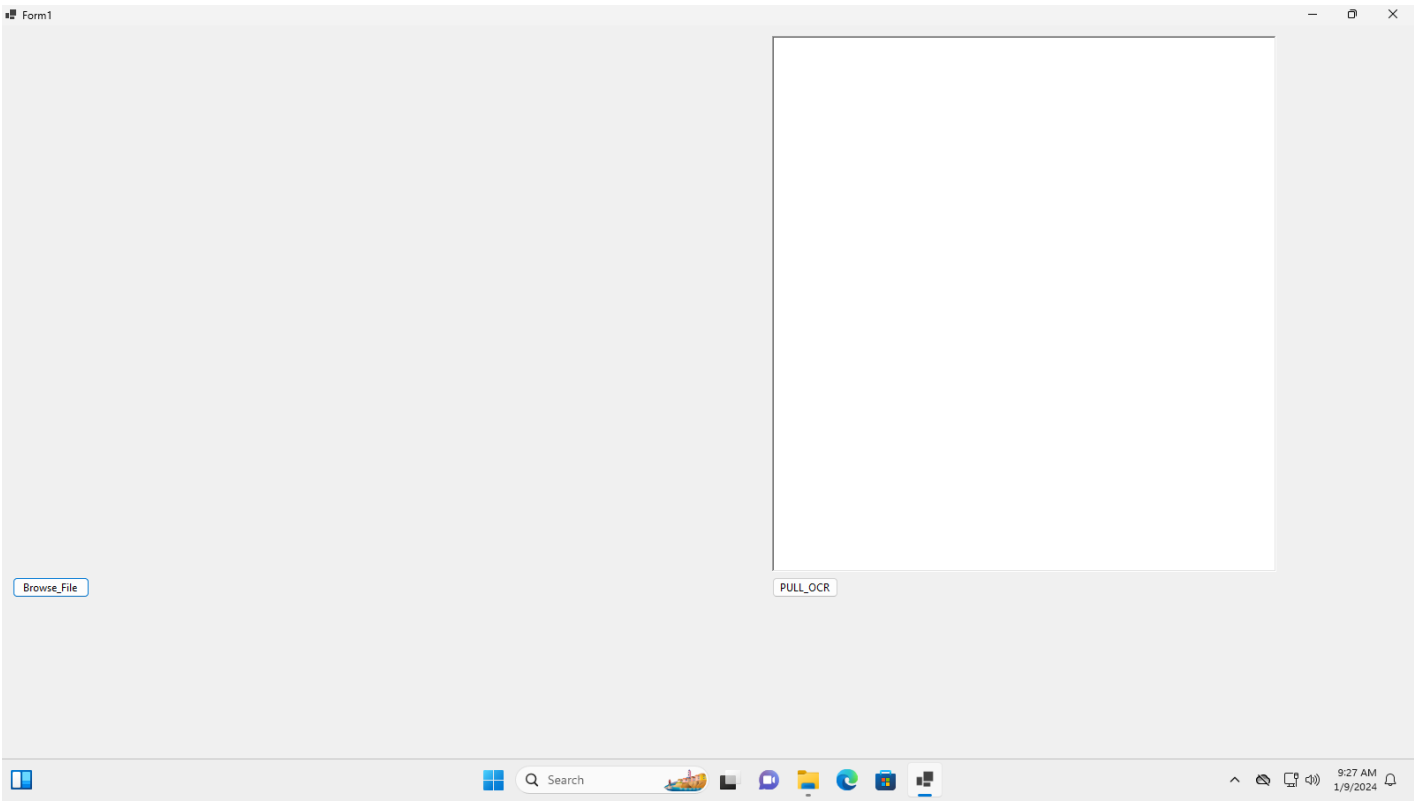
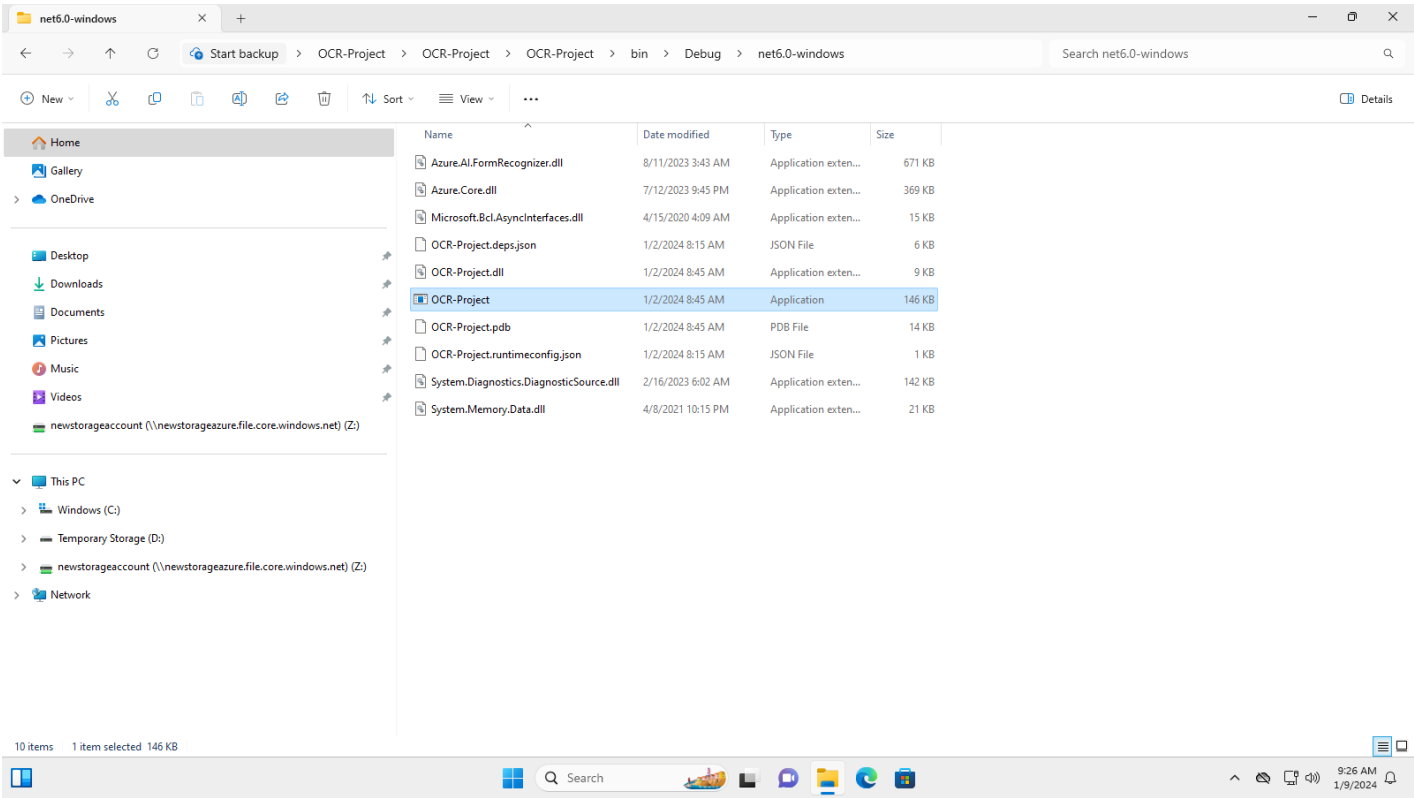
Now Download a Sample Image to Test



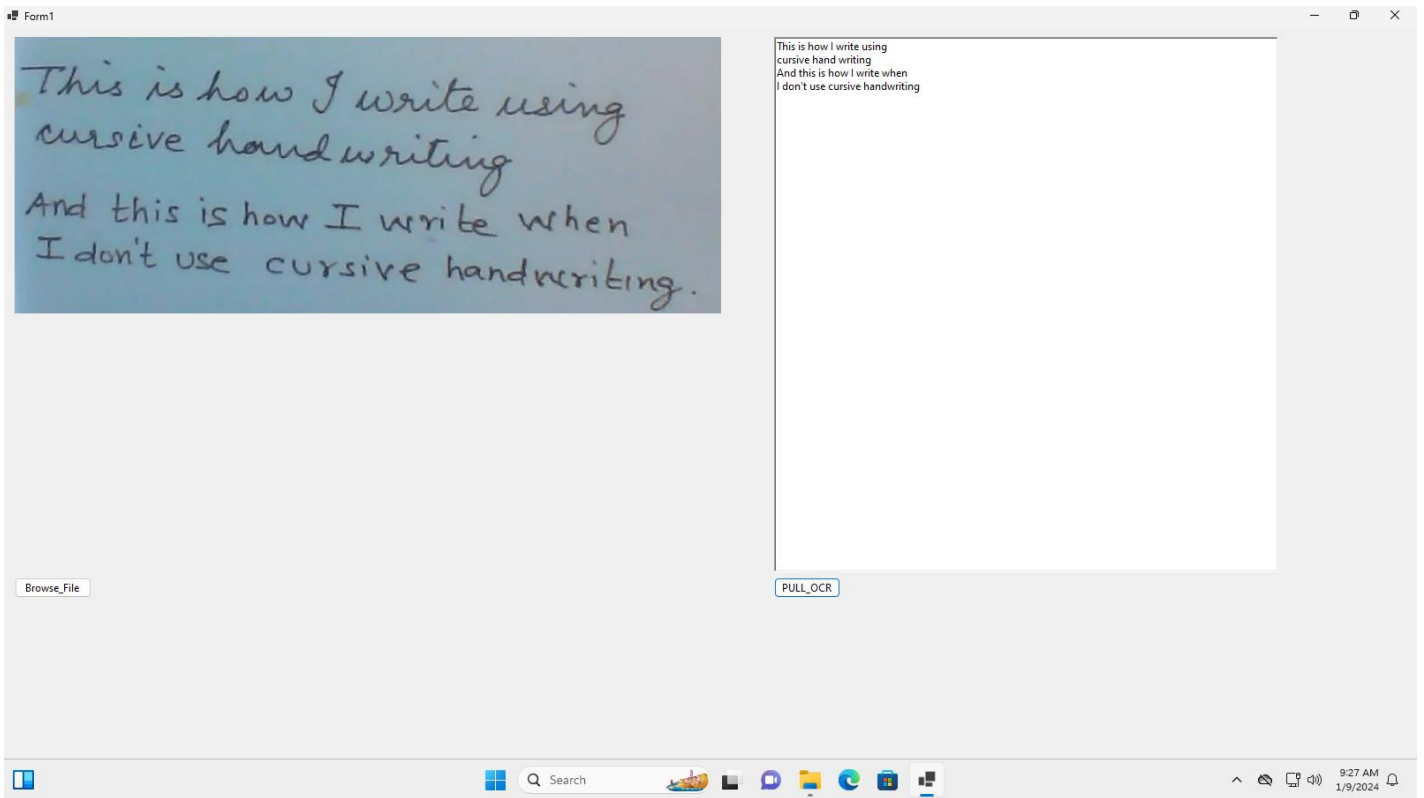
Copy the OCR Folder on VM from Local Machine



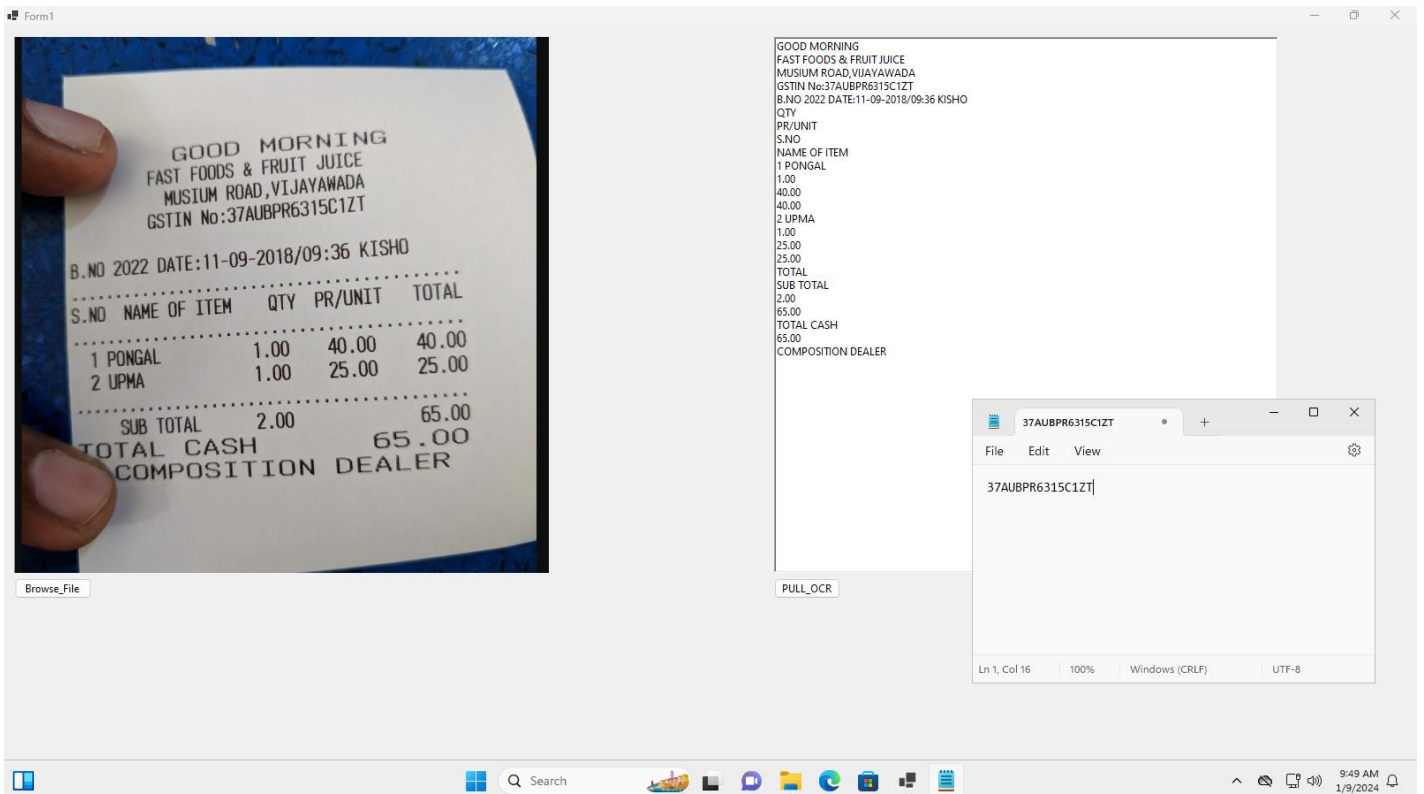
Execute the OCR-Project file setup



Testing on Sample Image 1



Testing on Sample Image 2 and now we can copy the contents from the Rich_Text_Format BOX



Created a Recovery Vault for Backup the File

Recovery Services vaults - Micro

portal.azure.com/#view/HubsExtension/BrowseResource/resourceType/Microsoft.RecoveryServices%2Fvaults

Microsoft Azure

Upgrade

Search resources, services, and docs (G+)

ashwinjanmagicbus@ou...
DEFAULT DIRECTORY (ASHWINJA...

Home

Recovery Services vaults

Default Directory (ashwinjanmagicbusoutlook.onmicrosoft.com)

Create

Manage view

Refresh

Export to CSV

Open query

Assign tags

Filter for any field...

Subscription equals all

Resource group equals all

Location equals all

Immutability equals all

Add filter

Showing 1 to 1 of 1 records.

No grouping

List view

Name	Type	Immutability	Resource group	Location	Subscription
vault-lqzjs4v1	Recovery Services vault	Disabled	OCR	Central India	Free Trial

< Previous

Page 1 of 1

Next >

Give feedback

Windows Taskbar

System Tray

Backup is working on Daily regular basis

Backup Jobs - Microsoft Azure

portal.azure.com/#view/Microsoft_Azure_DataProtection/V1JobsListBlade/vaultId/%2Fsubscriptions%2F3ca8fc99-4b10-4d5c-9fad-489553f3420c%2FresourceGroups%2FOCR%2Fproviders%2FMicr...

Microsoft Azure

Upgrade

Search resources, services, and docs (G+)

ashwinjanmagicbus@ou...
DEFAULT DIRECTORY (ASHWINJA...

Home

Recovery Services vaults

vault-lqzjs4v1

Backup items

Backup Items (Azure Storage (Azure Files))

newstorageaccount

Backup Jobs

Choose columns

Filter

Export jobs

Refresh

Feedback

Filtered by: Item Type - All, Operation - All, Status - All, Start Time - 1/8/2024, 3:22:01 PM, End Time - 1/9/2024, 3:22:01 PM

For backups, try our new Backup Center. It offers Azure Backup customers a unified view of Recovery Services Vaults used for backup in Azure. It also provides improved sorting and filtering along with new governance capabilities. Click here to get the new experience.

All data fetched from the service.

Filter items ...

Workload name	Operation	Status	Type	Start time	Total Duration	Details
newstorageaccount (newstorageazure)	Backup	Completed	Azure Storage	1/9/2024, 4:02:21 AM	00:00:04	View details
newstorageaccount (newstorageazure)	Backup	Completed	Azure Storage	1/9/2024, 12:02:04 AM	00:01:07	View details

< Previous

Page 1 of 1

Next >

Windows Taskbar

System Tray

- 24 -

Chapter 5: Benefits

1. Scalability and Availability:

- By leveraging Azure services like Availability Sets and Virtual Network, your OCR application can achieve high scalability and availability. Availability Sets ensure that your application remains accessible even if one of the virtual machines fails, providing enhanced fault tolerance. Additionally, Virtual Network facilitates secure communication between components, creating a robust and scalable infrastructure.

2. Cost-Effective Storage:

- The integration of Azure Storage Account allows you to efficiently manage and store large volumes of data generated by OCR processing. Azure Storage provides scalable and cost-effective storage options, ensuring that you only pay for the storage you use. This can lead to significant cost savings over traditional storage solutions.

3. Document Processing with AI:

- Utilizing Azure Document Intelligence services enhances your OCR capabilities by incorporating artificial intelligence (AI) for document understanding. This can result in improved accuracy and efficiency in extracting information from documents. Azure Document Intelligence provides features like entity recognition, key phrase extraction, and language detection, contributing to more sophisticated document processing.

4. Data Backup and Recovery:

- The integration of Recovery Services Vault ensures the safety and availability of your critical data. This Azure service provides automated backup and recovery options, enabling you to easily recover from data loss or system failures. This enhances the overall reliability of your OCR application and safeguards against potential data disasters.

5. Secure Communication and Compliance:

- Leveraging Virtual Network not only aids in scalability but also enhances the security of your OCR application. With Virtual Network, you can create isolated and secure communication channels between different components of your application. This helps in maintaining data privacy and can contribute to compliance with industry-specific regulations regarding data protection and security.

Chapter 6: Conclusion

OCR project leverages ASP.NET C# and Azure services, including Availability Sets and Virtual Networks for robust infrastructure. Storage Accounts ensure efficient data management, while Document Intelligence facilitates accurate optical character recognition. The inclusion of a Recovery Services Vault ensures data backup and disaster recovery capabilities. The project demonstrates a comprehensive solution for document processing, benefiting from Azure's scalability and reliability. Overall, your implementation seamlessly integrates key Azure services to deliver a powerful OCR system with enhanced security and resilience.