Cold Storage Archival and Retrieval using Azure Functions with Managed Identity

# 1. Introduction

This document describes the process of archiving old data from Azure Cosmos DB to Azure Cold Storage (Blob Storage with Cool or Archive tier), and retrieving it using Azure Functions written in Python. It also explains how Azure Managed Identity is used for secure, password-less access to Azure Blob Storage.

# 2. Architecture Overview

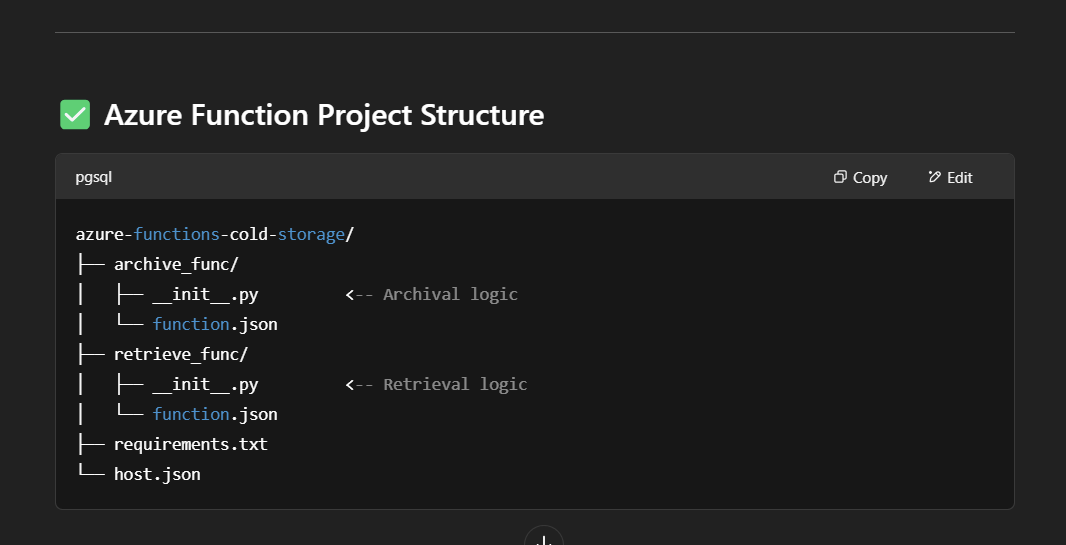
Components:

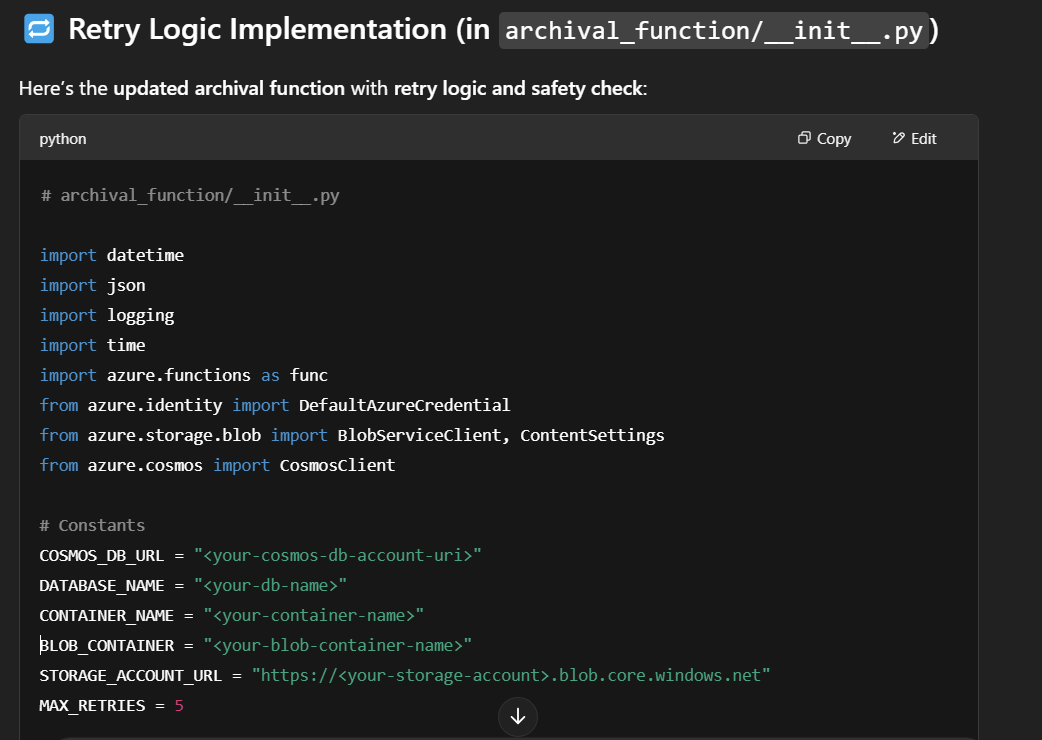
* - Azure Cosmos DB (source data)
* - Azure Blob Storage with Cool/Archive tier (cold storage)
* - Azure Function (Time-triggered for archival)
* - Azure Function (HTTP-triggered for retrieval)
* - Azure Managed Identity for secure access

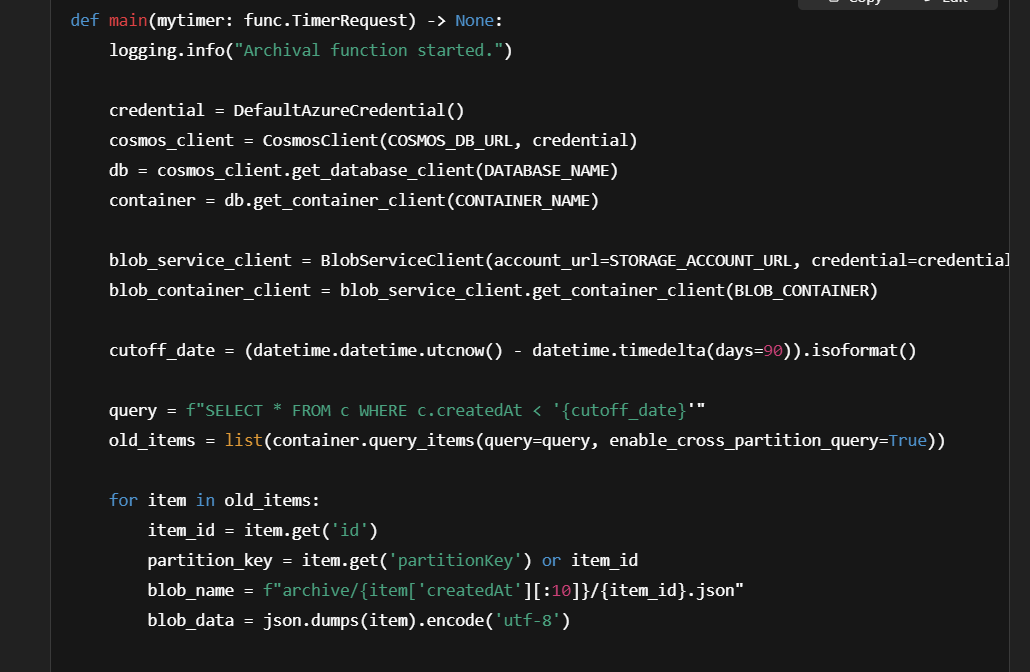
# 3. Time-Triggered Azure Function (Archival)

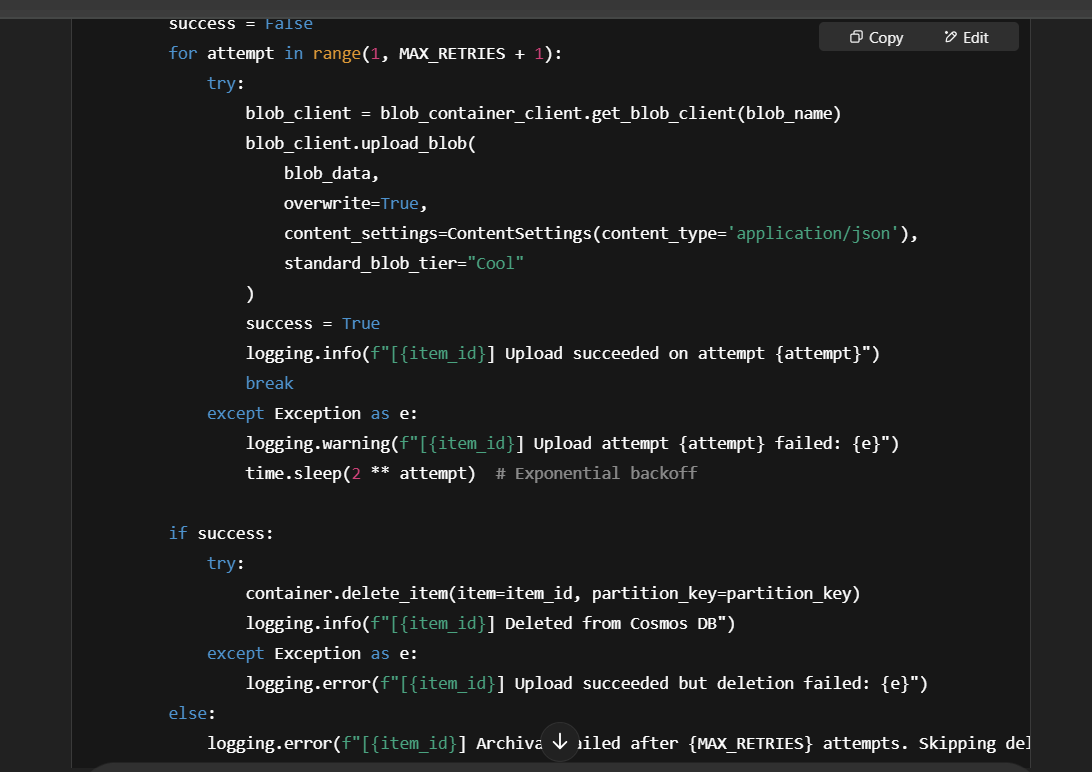
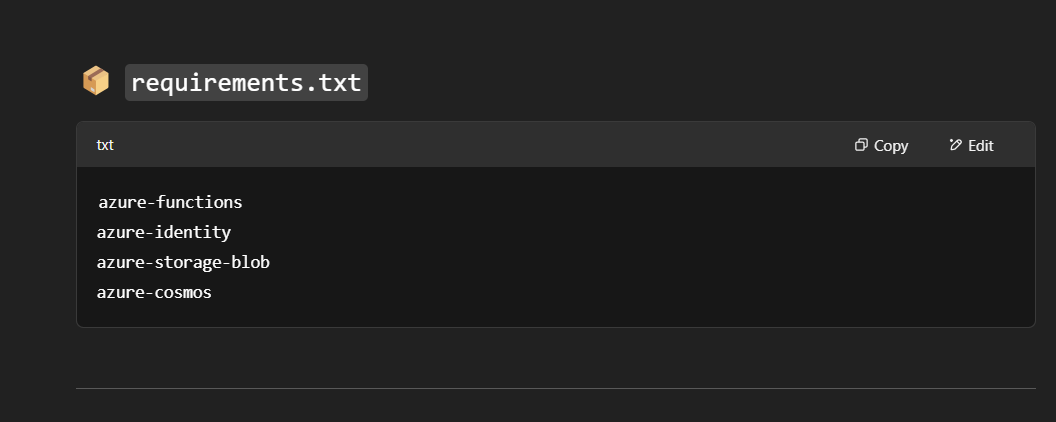
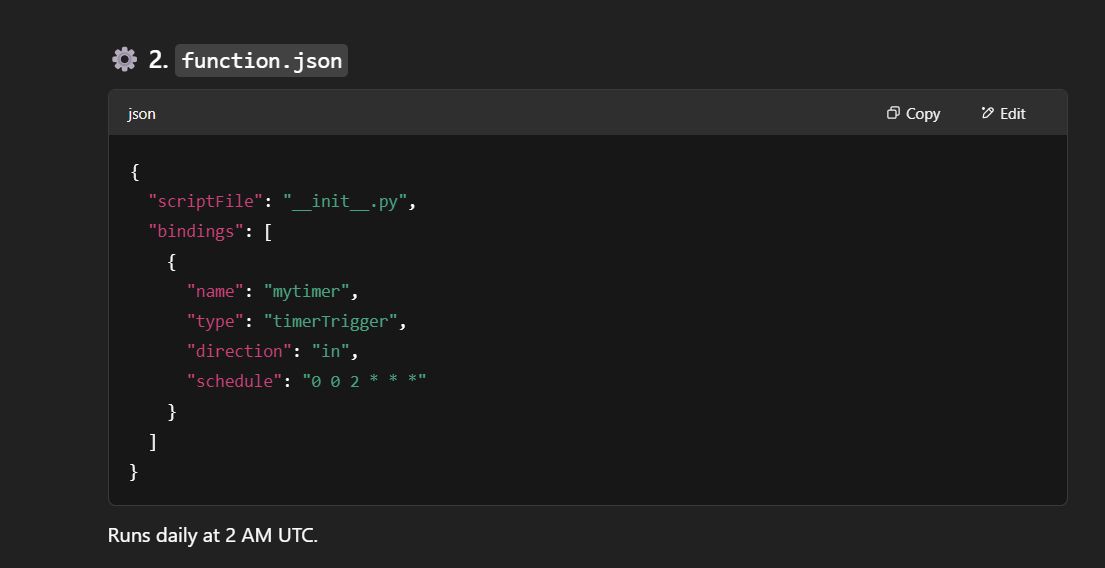
This function is triggered on a schedule (e.g., daily) to move data older than 3 months from Cosmos DB to Azure Blob Storage in the Cool/Archive tier.

Sample code:



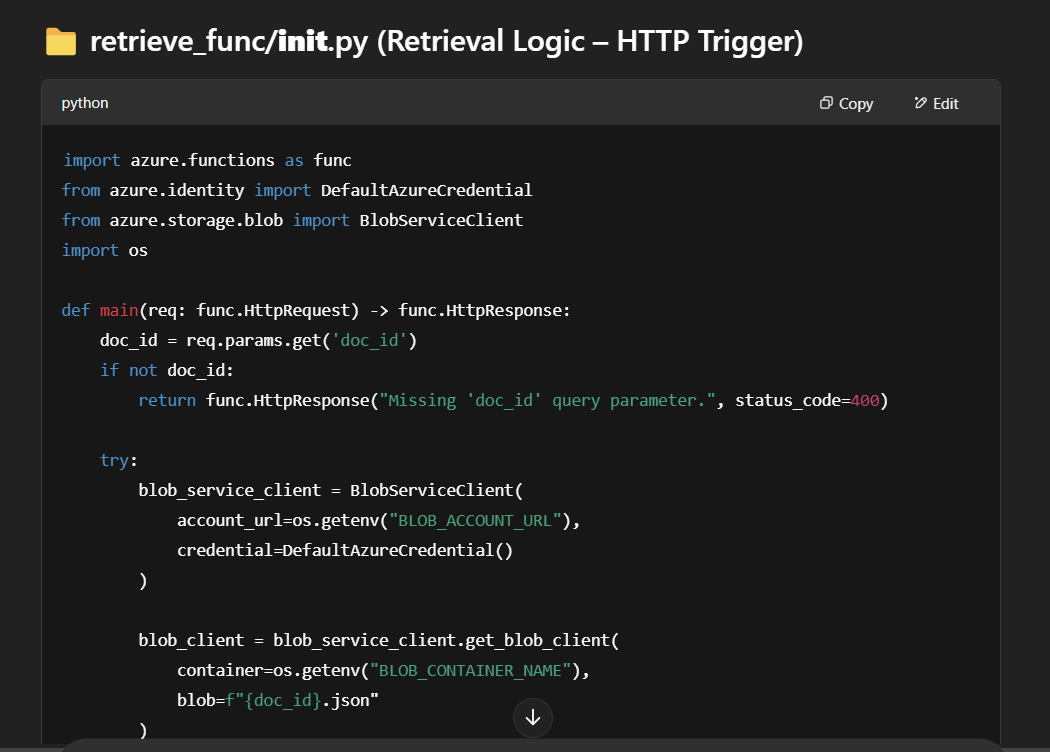


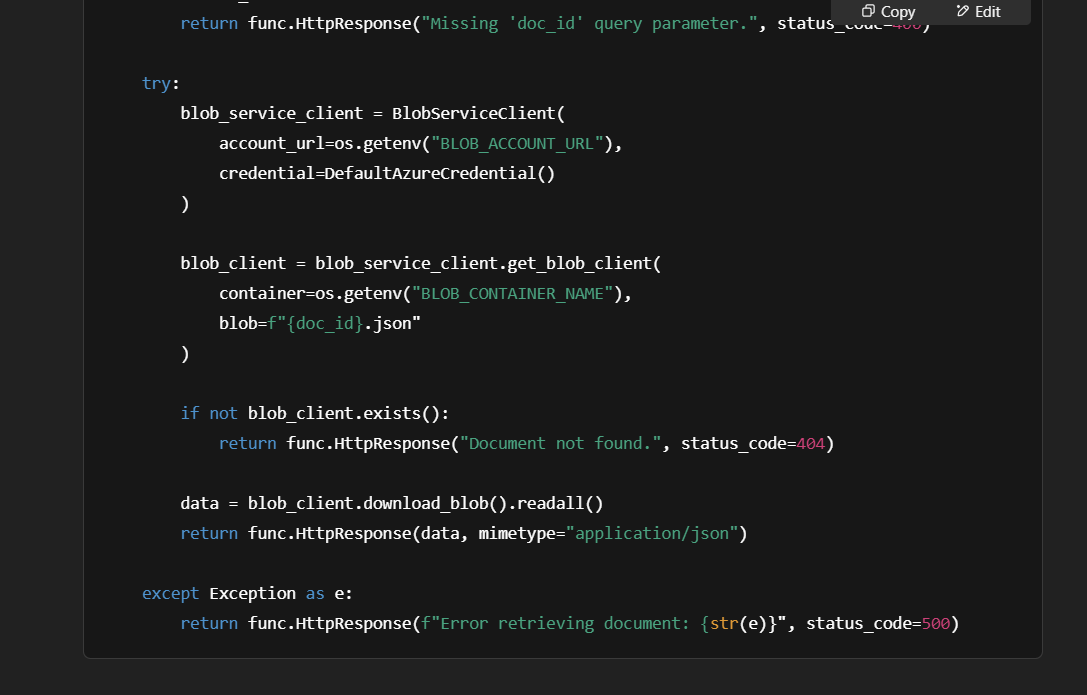


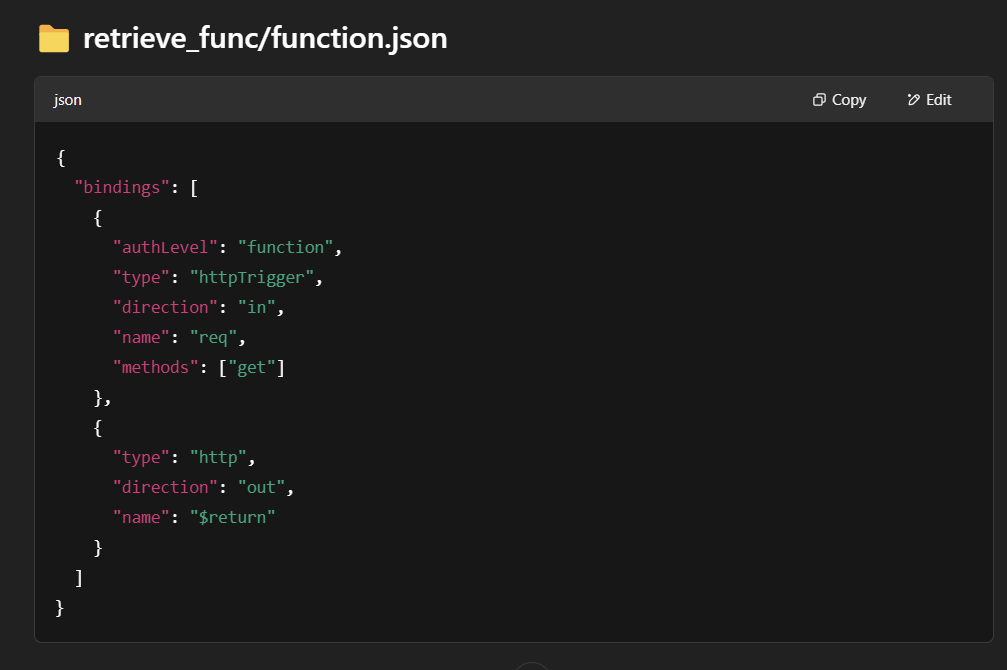
  
 

# 4. HTTP-Triggered Azure Function (Retrieval)

This function retrieves a specific document from cold storage by its ID.







# 5. Azure Managed Identity Setup

1. Enable Managed Identity in your Function App (Identity > System Assigned > On).

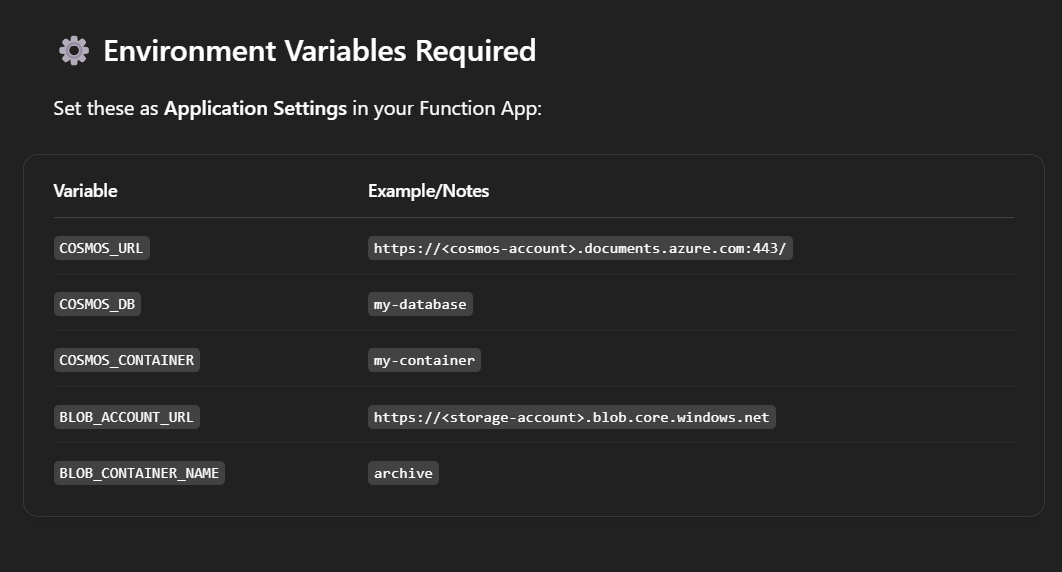
2. Assign RBAC role (Storage Blob Data Contributor) to your function on the Blob Storage:

* Command:

az role assignment create --assignee <client-id> --role "Storage Blob Data Contributor" --scope "/subscriptions/<sub-id>/resourceGroups/<rg>/providers/Microsoft.Storage/storageAccounts/<storage-account>"

# 6. Environment Configuration

Required environment variables in Azure Function App:



# 7. Key Points to Remember

* - Always use DefaultAzureCredential for Managed Identity access.
* - Schedule archival using a time trigger with CRON expression.
* - Use Cosmos DB TTL or timestamp checks for identifying old records.
* - Use Cool or Archive tier for cost-efficient cold storage.
* - Secure access via RBAC instead of connection strings.