

Tenant Workspace Configuration

Last Modified on 02/05/2025 4:41 am EST

Overview

The tenant-workspace-configuration.yaml file helps the system manage different workspaces for different document types. In the yaml you define the connection name to the file, the connection to the master data, the workspace port, and other workspace related properties.

Template

```
kind: document
metadata:
  name: extraction/v1/tenant-workspace
spec:
  configs:
    common:
      values:
        documentType:
          value: Invoice
        fileConnectionName:
          value: some_azure_bucket
        masterDataConnectionName:
          value: TestConnection
        transactionsConnectionName:
          value: fine_connection_name
        documentIntelligenceSubscriptionKey:
          value: 8eb04706b5e844a49184f554b82698f1
        documentIntelligenceEndpoint:
          value: https://open-ai-form-recognizer.cognitiveservices.azure.com/documentintelligence
    classification:
      values:
        classificationApiVersion:
          value: "2023-07-31"
    extraction:
      values:
        extractionApiVersion:
          value: "2024-11-30"
    auditLog:
      values:
        externalDBName:
          value:
        externalSchema:
          value: dbo
        externalTableName:
          value: MyTransactionsTable
    extractionGateway:
      values:
        domainName:
          value: pnmsoftlabs.com
        description:
          value: null
        specVersion:
          value: "1.0"
        type:
          value: edge-ide.v1.documentInjected
        source:
          value: /edge-ide/eaas-dev-webapi/invoices/events
        traceParent:
          value: 00-df3726d34dad4c39abdc6be8651eb75c-8b0c6a7f88c8ee69-01
    kafkaAuditLog:
```

```
values:
  topic:
    value: edge-test
  partition:
    value: null
  key:
    value: null
kafkaPublisher:
values:
  topic:
    value: edge-data-out
  partition:
    value: null
  key:
    value: null
preprocessing: #applicable only to Apflow
values:
  fileConnectionName:
    value: syscoocr
  targetPath:
    value: UAT/MAN_EMAIL_ACK
  maxFileSizeInMB:
    value: 30
  metadataroot:
    value: CoraSysco_OCR
splitHandling: #splitHandling
values:
  checkForSplit:
    value: true
  detectOnly:
    value: false
  rejectReason:
    value: Multiple invoices in single PDF
additionalCharges:
values:
  llmURL:
    value: https://llm-edge-dev-freight-inf.swedencentral.inference.ml.azure.com/score
  authorizationToken:
    value: Bearer eMgJ8slqLR9fNNCXbL27icspnJCCa8q0
eyeball:
values:
  textRecognitionEndPoint:
    value: https://open-ai-form-recognizer.cognitiveservices.azure.com/documentintelligence/documentModels/pr
ebuilt-layout:vcfg
  textRecognitionSubscriptionKey:
    value: sequence:secrets:eyeballSubscriptionKey
eyeball-client-configuration:
values:
  confidenceThreshold:
    value: 0.8
results:
values:
  adaptor:
    value: apflow #apflow or default (for webapi)
  kind:
    value: xml #(Xml/Json)
apFlow:
values:
  fileConnectionName:
    value: apflowfilesout
  targetPath:
    value: UAT/ShlomoOUT
  splittedFilesTargetPath:
    value: UAT/MAN_EMAIL_Split_Inbound
```

splitedRejectFilesTargetPath:
value: UAT/MAN_EMAIL_Split_Reject_Inbound

Parameter	Description
documentType	Name of the workspace.
fileConnectionName	The connection name of the extraction internal storage.
masterDataConnectionName	The connection name of the operational database.
documentIntelligenceSubscriptionKey	Subscription key to the DocIntel.
documentIntelligenceEndpoint	Endpoint to the DocIntel models.
domainName	
sourceFilesDirectory	The source from where the File Listener takes files.
targetFilesDirectory	The destination for APFlow files.
classification	The configuration for classification.
extraction	The configuration for extraction.
auditLog	Option to use the audit log data.
extractionGateway	Domain name structure.
kafkaAuditLog	Define Kafka connections.
kafkaPublisher	Connection to Kafka to publish full result.
preprocessing <i>(applicable only to APFlow)</i>	The configuration parameters for preprocessing, if applicable.
splitHandling	The configuration for document split.
additionalCharges	The configuration for handling additional charges. The LLM properties.
eyeball	Details of the OCR provider.
eyeball-client-configuration	Details of which fields will be displayed in the UI and how.
results	The required format. The adaptor value could be APFlow and XML or JSON for files.
apFlow	The configuration for APFlow.

