1. run\_management Schema

Tracks all document-level processing steps.

1. entity\_extraction\_runs
   1. Tracks AI extraction for each segregation\_id.

Final Payload Contracts

1. Text Extraction (Backend → Service)

Input

{

"segregation\_id": "SEG\_CITI\_ENG2024\_2024\_Q1\_CRE\_Sherwood",

"doc\_id": "DOC001",

"sas\_url": "https://<blob>.core.windows.net/container/blobpath?token=xyz",

"meta": {

"tenant\_id": "CITI",

"engagement\_id": "ENG2024",

"industry\_type": "CRE",

"obligor\_name": "Sherwood Village",

"year": 2024,

"quarter": "Q1"

}

}

Output

{

"doc\_id": "DOC001",

"status": "success",

"message": "Text extracted and stored successfully"

}

1. Vector Indexing (Backend → Service)

Input

{

"segregation\_id": "SEG\_CITI\_ENG2024\_2024\_Q1\_CRE\_Sherwood",

"doc\_id": "DOC001",

"meta": {

"tenant\_id": "CITI",

"engagement\_id": "ENG2024",

"industry\_type": "CRE",

"obligor\_name": "Sherwood Village",

"year": 2024,

"quarter": "Q1"

}

}

Output

{

"doc\_id": "DOC001",

"status": "indexed",

"vector\_id": "VEC123456"

}

1. AI Entity Extraction (Backend → Service)

Input

{

"extraction\_run\_id": "EXR003",

"segregation\_id": "SEG\_CITI\_ENG2024\_2024\_Q1\_CRE\_Sherwood"

}

Output

{

"extraction\_run\_id": "EXR003",

"segregation\_id": "SEG\_CITI\_ENG2024\_2024\_Q1\_CRE\_Sherwood",

"status": "completed",

"entities": [

{ "field": "Borrower Name", "value": "Sherwood Village" },

{ "field": "Loan Amount", "value": "$1,000,000" },

{ "field": "Interest Rate", "value": "6.5%" }

],

"source\_documents": ["DOC001", "DOC002"]

}

1. Update Entity Extraction Results (Frontend → Backend)

Input:

{

"segregation\_id": "SEG\_CITI\_CRE\_Q1\_2024\_OBL001",

"extraction\_run\_id": "EXTRACT\_002",

"updated\_entities": [

{ "field": "Loan Term", "value": "12 Years" }

]

}

Output

{

"message": "Update successful",

"archived\_version": {

"archive\_id": "ARCHIVE\_002",

"archived\_on": "2024-03-30T10:00:00Z"

}

}