



Smart Database

Knowledge Center

Participant Guide –

AI-Powered Document Processing

VERSION 10.0 - SERVICE RELEASE 7

December 2024



About This Document

This document applies to SmartDB (Smart Database) Version 10.0 and all subsequent releases. The specifications and information contained herein are subject to change without prior notice. Any changes will be documented in future release notes or updated editions of this guide.

Copyright Notice

Copyright © 2023 – 2025 SmartDB inc. , USA, and/or its subsidiaries, affiliates, and licensors. All rights reserved.

License and Usage Terms

Use of this software is governed by the SmartDB (Smart Database) License Agreement and applicable commercial terms provided at the time of purchase or deployment.

The license terms are included as part of the SmartDB (Smart Database) product documentation and/or are available in the root directory of the installed SmartDB (Smart Database) software. Unauthorized reproduction, distribution, or modification of this software or documentation is strictly prohibited.

Third-Party Software

SmartDB (Smart Database) may include or integrate with third-party software components, open-source libraries, or external services. Applicable third-party copyright notices, license terms, disclaimers, and additional rights or restrictions are provided in the Third-Party Licenses documentation included with the product distribution and/or located in the installation directory. Use of such third-party components is subject to their respective license terms.

Trademarks

The name SmartDB , Smart Database, and all SmartDB-related product names, logos, and services are either trademarks or registered trademarks of SmartDB inc. and/or its affiliates. All other company names, product names, and trademarks mentioned in this document are the property of their respective owners and are used for identification purposes only.

Disclaimer

This documentation is provided “as is” without warranty of any kind, either express or implied. SmartDB.ai / Script Technical Solutions LLC shall not be liable for any direct, indirect, incidental, or consequential damages arising from the use of this document or the software described herein.



Contents

About This Document	2
1. Introduction	6
1.1 Scope of Document	6
2. Scope of AI-Powered Document Processing Module Configuration	8
2.1 AI-Powered Document Management	8
2.1.1 Accessing AI-Powered Document Management	9
2.1.2 Documents Library Management	9
2.1.2.1 View Document Libraries	11
2.1.2.2 Smart Database Document Library Configuration Workflow	14
2.1.2.2.1 Add Document Library	15
2.1.2.2.1.1 Access Document Library	15
2.1.2.2.1.2 Create Document Library	16
2.1.2.2.1.2.1 Basic Library Information	17
2.1.2.2.1.2.2 Library Type Configuration	18
2.1.2.2.1.2.3 Embedding & Retrieval Parameters	19
2.1.2.2.1.2.4 LLM Generation Parameters	21
2.1.2.2.1.2.5 Prompt & Behavior Configuration	22
2.1.2.2.1.2.6 Safety, Validation & FAQ Controls	22
2.1.2.2.1.3 Upload Documents (Upload for AI-Powered Understanding)	24
2.1.2.2.1.3.1 Access Document Library	24
2.1.2.2.1.3.2 Upload Documents	24
2.1.2.2.1.4 Document Upload Methods	25
2.1.2.2.1.4.1 Supported File Formats	26
2.1.2.2.1.5 Document Ingestion & Processing	26
2.1.2.2.1.6 Document Configuration	27
2.1.2.2.1.6.1 Document Indexing Settings	28



2.1.2.2.1.7	Document Intent Methods	29
2.1.2.2.1.7.1	Purpose of Intent Configuration	30
2.1.2.2.1.7.2	Intent Description Field	31
2.1.2.2.1.7.3	Applicability & Scope	31
2.1.2.2.1.7.4	Role in Document Routing Strategy	31
2.1.2.2.2	Managing User Access and Privileges for Document Libraries	32
2.1.2.2.2.1	Accessing User Privileges Configuration	32
2.1.2.2.2.2	Add User to Document Library	33
2.1.2.2.3	FAQ Configuration (Library-Level FAQ Control)	34
2.1.2.2.3.1	Accessing Document FAQ Configuration	35
2.1.2.2.3.2	Document FAQ Management	36
2.1.2.2.3.2.1	Add & Upload FAQ	36
2.1.2.2.3.2.1.1	Manual Entry	37
2.1.2.2.3.2.1.2	Upload Bulk FAQ Entry	39
2.1.2.2.3.3	Documents Library Chat History	42
2.1.2.2.3.4	Documents Library Re-Indexing (AI Reprocessing)	43
2.1.2.2.3.4.1	When to Use Re-Indexing	44
2.1.2.2.3.4.2	Re-Indexing Confirmation	44
2.1.2.2.3.4.3	Impact of Re-Indexing	45
2.1.2.2.3.5	API Details for integration	45
2.1.2.2.3.5.1	API Structure	46
2.1.2.2.3.5.2	API Integration Benefits	47
2.2	AI-Powered Document Management	48
2.2.1	Accessing the Document Advisory Module	48
2.2.2	Chat with Your Document (Document Advisory)	49
2.2.2.1	Selecting a Document Library	49
2.2.2.2	Conversational Interaction	49
2.2.2.3	Speech Capabilities	50



Smart Database

2.2.2.4	Related Documents Panel.....	50
2.2.2.5	Chat History Option	50
2.2.2.6	Configuration Settings.....	51
2.2.2.7	Summery	53



1. Introduction

1.1 Scope of Document

This document defines the scope of the AI-Powered Document Processing application within the SmartDB (Smart Database) platform. It outlines the functional coverage, core capabilities, and AI-driven features that enable organizations to ingest, manage, analyze, and interact with documents in an intelligent and secure manner.

The Document Analysis Module allows users to integrate their document repositories and process them through advanced Artificial Intelligence technologies. SmartDB (Smart Database) provides out-of-the-box Optical Character Recognition (OCR) to automatically extract text from scanned documents and image-based files, eliminating the need for manual data entry. In addition, users can create custom document libraries, apply AI processing to their documents, and chat with their document collections using natural language to retrieve answers, insights, and references directly from the content.

The scope of AI-Powered Document Processing includes the following capabilities:

- ❖ **Document Ingestion and Classification**

Upload and ingest documents in multiple ways, including individual files, entire folders, or direct links to external repositories. Documents are automatically classified based on content, structure, and metadata.

- ❖ **Out-of-the-Box OCR Processing**

Automatic text extraction from scanned documents and images, enabling full AI analysis and searchability across non-digital and image-based content.

- ❖ **Data Extraction and Parsing**

Extraction of structured information such as text, tables, key fields, and entities to support analytics, reporting, and downstream processing.

- ❖ **Document Summarization and Key Insights**

AI-driven summarization to highlight key points, decisions, risks, and important information from large or complex documents.

- ❖ **Sentiment and Tone Analysis**

Analysis of document sentiment, tone, and contextual intent to support compliance, investigation, feedback analysis, and risk assessment.



❖ **AI-Powered Search and Retrieval**

Intelligent, context-aware search using natural language queries across all document libraries.

❖ **Document Chat & Advisory**

Ability to interact conversationally with document libraries, allowing users to ask questions and receive accurate, context-aware responses with references to source documents.

❖ **User Management for Document Libraries**

Role-based access control and user permission management to ensure secure and governed access to document content.

❖ **Multilingual Search and Translation**

Support for multilingual document processing, search, and AI-based translation to enable cross-language accessibility.

❖ **Text-to-Speech Mechanism**

Conversion of document content into audio output to improve accessibility and hands-free content consumption.

This document focuses on the functional scope, AI-driven capabilities, and operational features of the Document Processing module within the Smart Database platform. It describes how the module enables intelligent document ingestion, analysis, search, and interaction using Artificial Intelligence, Natural Language Processing, and Large Language Models.

The content of this document is limited to application-level functionality, user workflows, configuration options, and supported AI use cases.

Detailed installation instructions, infrastructure sizing, hardware and software prerequisites, security hardening, and deployment procedures are not covered in this document and are provided separately in the Installation Guide, which should be referenced for environment setup and system deployment activities.



2. Scope of AI-Powered Document Processing Module Configuration

2.1 AI-Powered Document Management

SmartDB (Smart Database) provides an AI-Powered Document Management System that enables organizations to securely store, manage, analyze, and interact with documents using advanced Artificial Intelligence technologies. The Document Management module allows users to ingest documents from multiple sources, process them through built-in AI services, and transform unstructured content into searchable, intelligent, and actionable information.

By leveraging out-of-the-box OCR, Natural Language Processing (NLP), and Large Language Models (LLMs), SmartDB enables users to analyze documents, extract key information, generate summaries, perform intelligent search, and interact conversationally with document libraries. This significantly reduces manual effort, improves information discovery, and enhances decision-making across enterprise and government environments.

When AI-powered document processing is enabled, SmartDB can ingest documents in various formats, automatically classify and index them, extract structured data, and make them available for AI-powered search, document chat, and analytics. Access to document libraries is governed by user roles and permissions, ensuring that only authorized users can view and interact with sensitive content.

Purpose of AI-Powered Document Management

The AI-Powered Document Management module in SmartDB is used to:

- Centralize document storage and management within a secure, governed platform
- Automatically ingest and classify documents from files, folders, and external repositories
- Apply **out-of-the-box OCR** to scanned and image-based documents
- Extract text, metadata, entities, and key data from documents
- Enable AI-driven document summarization and insight generation
- Allow users to **chat with document libraries** using natural language queries
- Perform intelligent, context-aware search across large document collections
- Support multilingual document processing, search, and translation
- Enhance accessibility through **Text-to-Speech (TTS)** capabilities
- Enforce role-based access control and user permissions for document libraries
- Support audit, compliance, and information governance requirements



2.1.1 Accessing AI-Powered Document Management

Users can access the AI-Powered Document Management module through the SmartDB (Smart Database) main interface after successful authentication. From the dashboard, users can navigate to the Document Management section to create, manage, and interact with document collections.

Figure 1 illustrates how users can access the AI-Powered Document Management Center from the SmartDB main dashboard. By selecting Document Management from the left-side navigation menu, users are directed to the centralized document workspace where AI-driven document processing and analysis capabilities are available.

This entry point serves as the gateway to all document-related functions, including document library management, AI-powered analysis, OCR processing, document search, and conversational interaction with documents. Access to this section is controlled by user roles and permissions to ensure secure and governed usage across the organization.

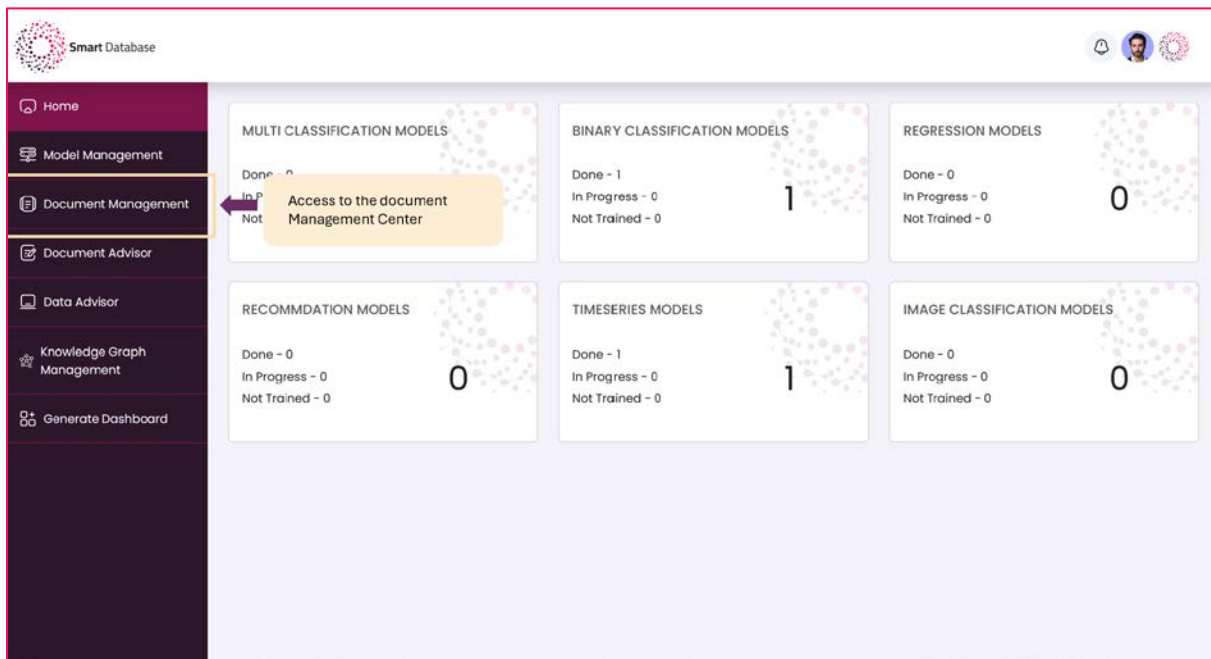


Figure 1: Navigating to the Document Management Module

2.1.2 Documents Library Management

The Documents Library Management page serves as the central control panel for managing AI-powered document libraries within the Smart Database platform. This interface enables authorized users to create, configure, and govern document libraries that are processed and analyzed using artificial intelligence.



Access to this page is restricted to users with ADMIN or ROLE_USER privileges. Users assigned the ROLE_GUEST role are not permitted to access this screen or perform any document library management operations.

From this page, authorized users can perform the following actions:

- **View Document Libraries**
Display a structured list of all existing document libraries, including library name, OCR languages, processing type, and current status.
- **Add New Library**
Create a new document library by clicking the “**Add New**” button. Libraries can be configured with specific OCR languages and AI processing modes.
- **Upload and Index Documents**
Upload individual files, folders, or external links (such as repositories or URLs) and index them for AI-powered analysis and retrieval.
- **Manage User Access**
Assign and manage users who are allowed to chat with documents within each library, enforcing access control and governance policies.
- **Edit Library Configuration**
Modify library settings, including OCR language selection, document processing rules, and AI behavior configuration.
- **AI-Powered FAQ and Knowledge Extraction**
Automatically generate and manage Frequently Asked Questions (FAQs) derived from document content using AI.
- **Re-Index Documents (Admin Only)**
Reprocess all documents and regenerate embeddings, FAQs, and metadata to reflect updated content or configuration changes.
- **Chat History Management**
Review historical interactions between users and the document library through the AI chat interface.
- **Lock / Unlock Library**
Lock a library to prevent modifications or unlock it for editing. The unlock operation is restricted to administrators only.
- **Delete Library**
Permanently remove a document library and its associated data when no longer required.



- **External Web Service Integration**

Enable integration with external web services for advanced document ingestion and processing workflows.

Illustration **Figure 2** shows the Documents Library Management interface, highlighting the main actions available for managing AI-powered document libraries, including library creation, document ingestion, user access control, indexing, and administrative operations.

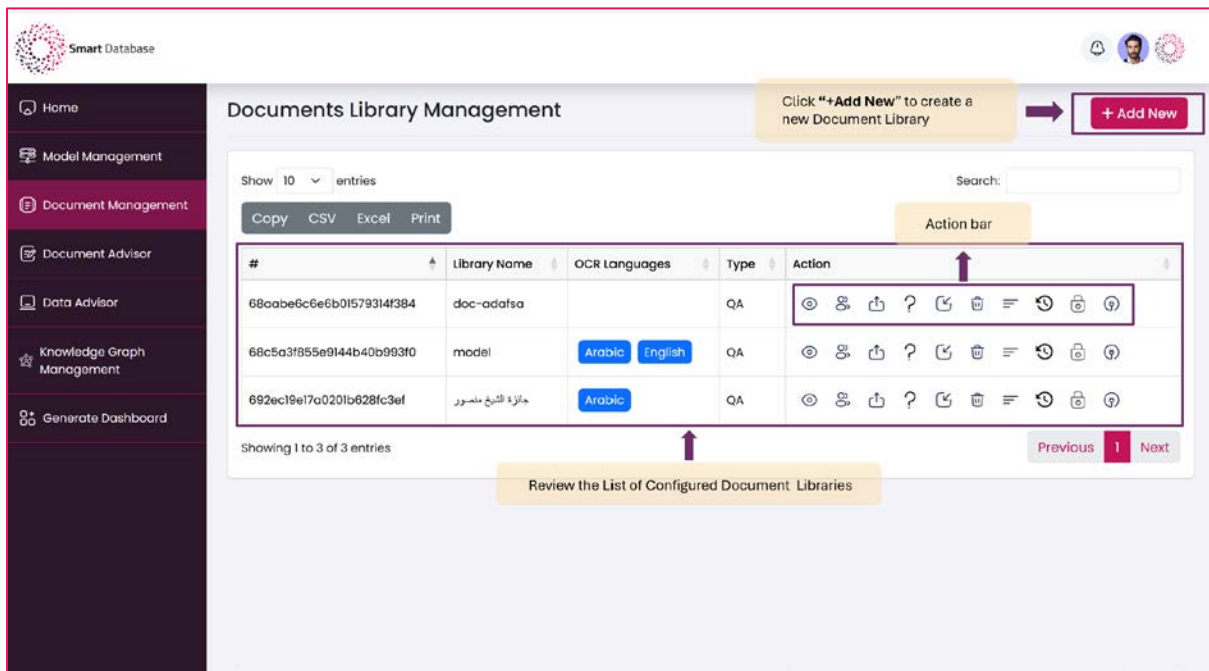


Figure 2: Documents Library Management Interface

This page is a key component of SmartDB’s AI-Powered Document Processing module, ensuring secure document lifecycle management, controlled user access, and seamless AI-driven interaction with enterprise knowledge assets.

2.1.2.1 View Document Libraries

From the Documents Library Management page, users can navigate to view the documents within a specific library by selecting the View Library Documents icon (eye icon) under the Action column.

As shown in Figure 3, clicking the *View* icon opens the *View Documents Library* screen, where users can see all documents associated with the selected library, along with their processing and indexing status.

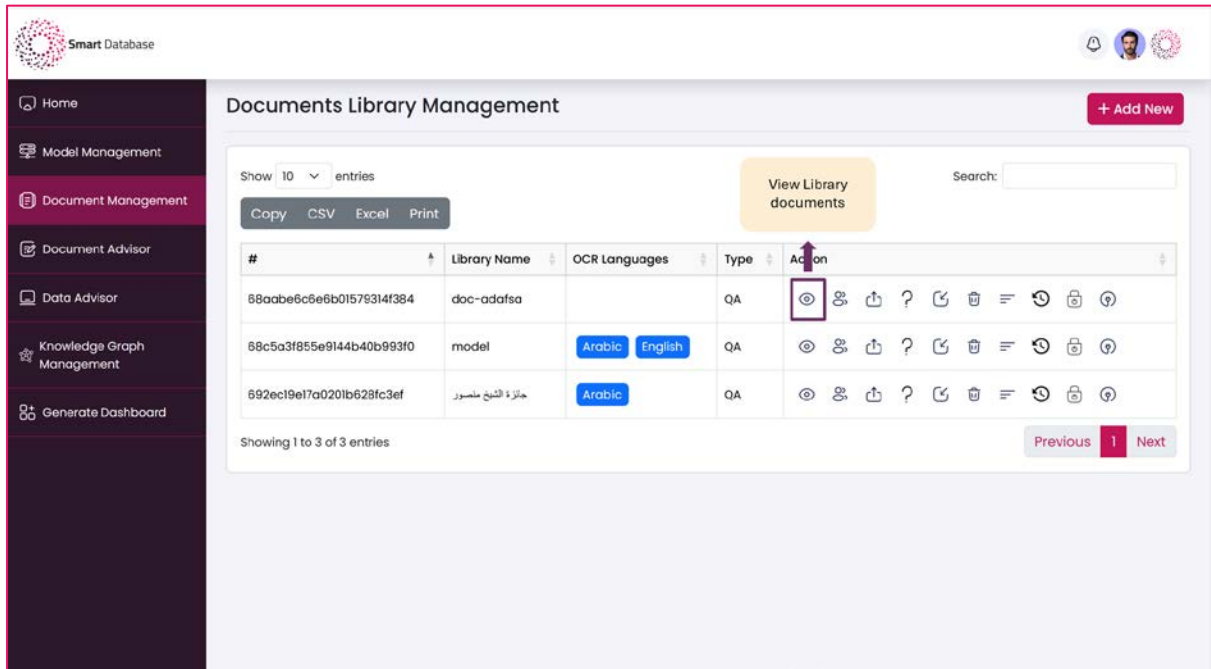


Figure 3: Navigating to View Documents in the Document Library

The View Documents Library page provides users with a detailed overview of all documents uploaded to a selected document library within SmartDB (Smart Database). This screen allows authorized users to monitor document ingestion, processing, and indexing status as part of the AI-powered document management lifecycle.

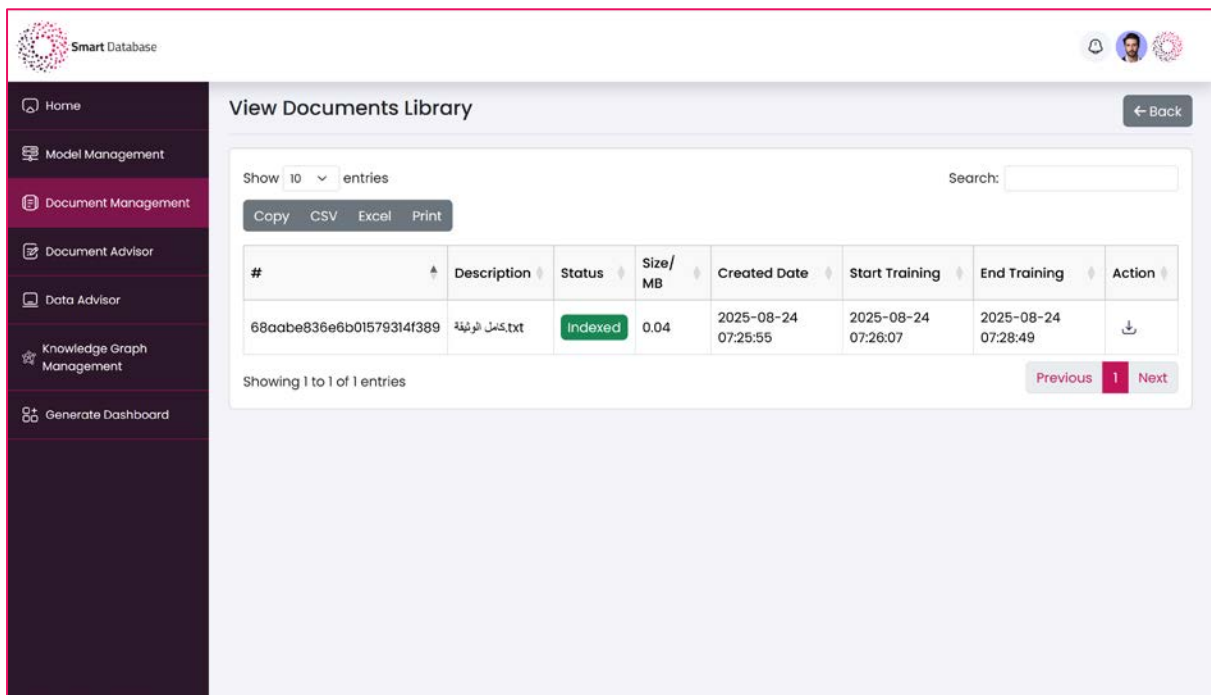


Figure 4: View Documents Library Interface



As shown in the **figure 4**, the page displays a tabular list of documents along with key metadata, including document description, processing status, file size, creation date, and training timestamps. Users can also export the list using Copy, CSV, Excel, or Print options, and search for specific documents using the search field.

Each document listed in the library passes through multiple processing stages before becoming available for AI-powered search, chat, and analysis.

Document Status Definitions

The Status column indicates the current processing state of each document:

- **Pending**
Indicates that the document has been successfully uploaded but is waiting to be processed. At this stage, the document is queued and has not yet entered the AI analysis or indexing pipeline.
- **On Progress**
Indicates that the document is currently under processing. During this stage, SmartDB applies OCR (if required), extracts text, analyzes content, and performs AI indexing to prepare the document for search, retrieval, and conversational interaction.
- **Indexed**
Indicates that the document has been fully processed and indexed successfully. Indexed documents are ready for use within AI-powered features such as document chat, semantic search, summarization, and knowledge extraction.

Page Capabilities

From this screen, users can:

- View the list of documents associated with a specific library
- Track document processing and indexing status in real time
- Review document size and processing timestamps
- Download processed documents (if permitted)
- Navigate back to the main library management screen



2.1.2.2 Smart Database Document Library Configuration Workflow

This workflow illustrates the end-to-end process for creating, configuring, and using an AI-powered document library in SmartDB (Smart Database). It guides administrators and authorized users through the full lifecycle—from library creation to intelligent document interaction and external system integration.

As shown in the workflow diagram **figure 5**, SmartDB follows a structured and controlled approach to ensure secure document handling, accurate AI processing, and governed user access.

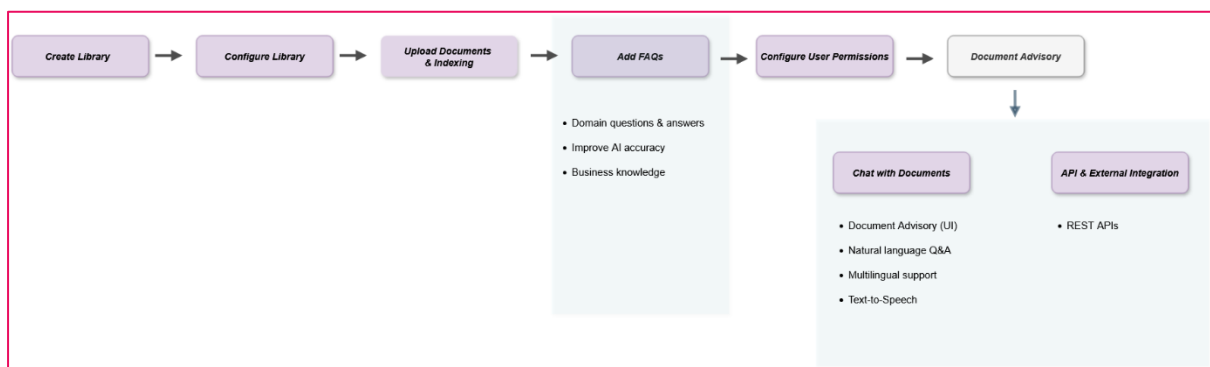


Figure5: Add and Configure Document Library Workflow in Smart Database

Workflow Description

1. Create Library

The process begins by creating a new document library. Each library represents a logical container for documents related to a specific domain, department, project, or business use case.

The library serves as the foundation for all subsequent AI processing and access control.

2. Configure Library

Once created, the library is configured with key settings, including OCR languages, AI processing modes, indexing rules, and document handling behavior.

These configurations determine how documents are analyzed, indexed, and retrieved by the AI engine.

3. Upload Documents & Indexing

Users can upload documents in multiple ways, including individual files, folders, or repository links.

SmartDB supports various document formats and provides out-of-the-box OCR capabilities.

Uploaded documents go through an automated processing lifecycle:



- **Pending** – Document is queued for processing
- **In Progress** – AI processing and indexing are underway
- **Indexed** – Document is fully processed and ready for search and interaction

4. Add FAQs

Domain-specific Frequently Asked Questions (FAQs) can be added to the library to enhance AI understanding and response accuracy. FAQs help enrich business knowledge, improve contextual answers, and optimize the document advisory experience.

5. Configure User Permissions

Access to the library is controlled through role-based permissions. Administrators can assign users and define who can upload documents, manage configurations, re-index content, or interact with documents. Integration with LDAP or enterprise identity systems can be used to enforce centralized access governance.

6. Document Advisory

After configuration, the library becomes available to SmartDB's Document Advisory module. Authorized users can interact with documents using natural language queries, enabling intelligent search, summarization, multilingual support, and text-to-speech capabilities.

- **API Integration**

Users can chat directly with documents through the SmartDB user interface or integrate document intelligence into external systems using REST APIs.

This allows organizations to embed AI-powered document insights into portals, applications, ERP, CRM, or other enterprise platforms.

2.1.2.2.1 Add Document Library

2.1.2.2.1.1 Access Document Library

The Create Document Library page enables administrators to define and configure an AI-powered document library that serves as the foundation for document indexing, retrieval, and conversational AI within SmartDB (Smart Database).

To create a new Document Library, users with administrative privileges must access the Documents Library Management screen and select the “+ **Add New**” button located in the top-right corner of



the page, as illustrated in **Figure 6**. Clicking this button initiates the library creation workflow and allows the user to configure document processing, AI behavior, and access controls.

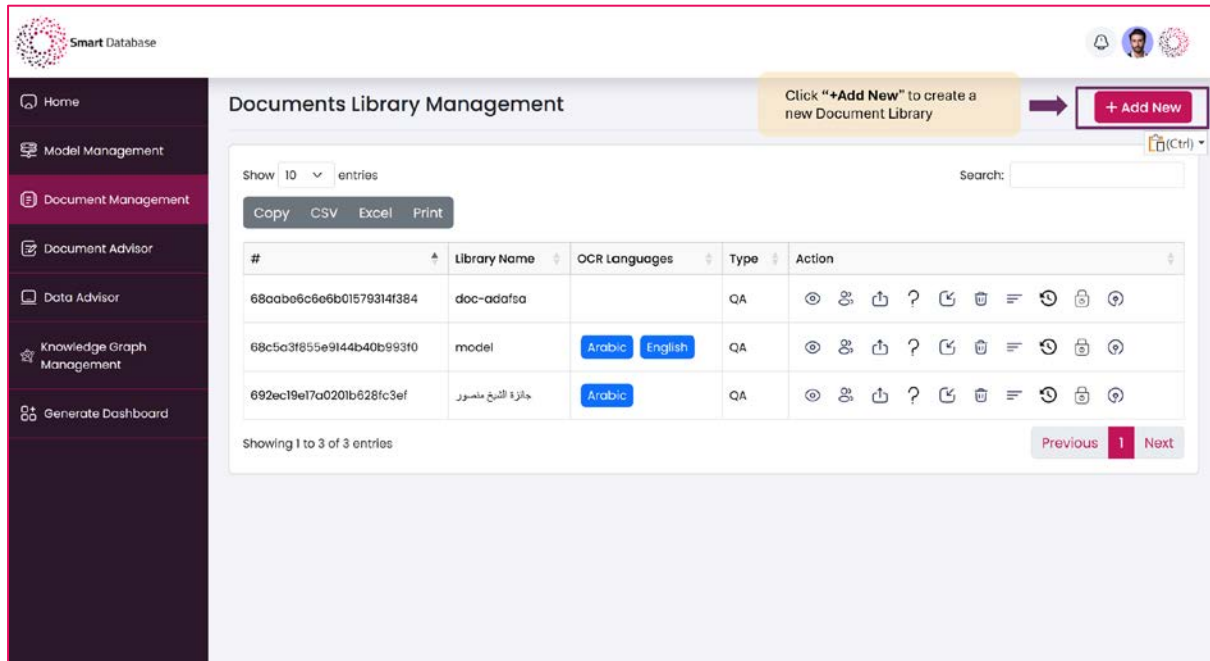


Figure6: Documents Library Management – Creating a New Document Library

2.1.2.2.1.2 Create Document Library

The Add Documents Library page provides a comprehensive set of configuration fields that define how SmartDB (Smart Database) creates and manages a document library, processes uploaded documents, applies AI-based indexing, and controls how user questions are answered. Through this page, administrators can configure document ingestion behavior, retrieval logic, Large Language Model (LLM) interaction, and response governance to ensure accurate, secure, and context-aware document advisory capabilities.

The Add Documents Library page is organized into the following logical sections:

- Basic Library Information
- Library Type Configuration
- Embedding & Retrieval Parameters
- LLM Generation Parameters
- Prompt & Behavior Configuration
- Safety, Validation & FAQ Controls



2.1.2.2.1.2.1 Basic Library Information

This section defines the core identity and scope of the document library, including the library name, supported OCR languages, and the LLM connection to be used.

The selected OCR languages determine how documents are processed during ingestion and are also used for multilingual responses and translations when users interact with documents.

The screenshot shows a web form titled "Add Documents Library". It contains three input fields:

- Documents Library Name:** An empty text input field with a red border and a small red circle icon on the right.
- Documents Languages (OCR):** An empty text input field with a red border.
- LLM Connection:** A text input field containing the text "Ollama - qwen2.5:14b" with a red border.

Figure 7: Add Basic Library Information

- **Document Library Name**
 - Defines a unique name for the document library.
 - This name is used to identify the library across SmartDB modules, including Document Management, Document Advisory, and API integrations.
- **Document Languages (OCR)**
 - Allows selecting one or more languages to be used for:
 - OCR processing when scanned or image-based documents are uploaded
 - Answer translation, ensuring responses are returned in the same language used by the user when chatting with documents
 - Example: Selecting *Arabic* and *English* enables bilingual document processing and multilingual question answering.
- **LLM Connection**
 - Specifies which pre-configured Large Language Model (LLM) is used for answering questions and generating responses from the document content.
 - The selected LLM determines reasoning quality, language handling, and response style.



2.1.2.2.1.2.2 Library Type Configuration

The section Library Type Configuration determines how Smart database retrieves information when users ask questions. SmartDB (Smart Database) supports three document library types, each designed for different use cases as shown in **figure 8**.

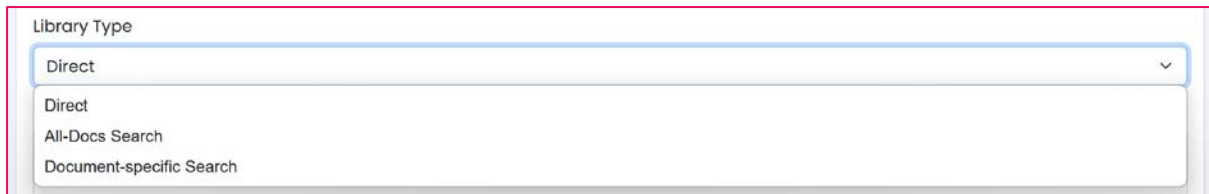


Figure 8: Library Type Configuration Options

1. Direct

- Questions are sent directly to the selected LLM
- No document upload or indexing is required
- Suitable for general AI assistants or knowledge-free interactions

2. All-Documents Search

- Users can upload any number of documents
- When a question is asked, the system searches across all indexed documents
- Relevant text chunks from one or more documents are used to generate the answer

3. Document-Specific Search

- Users upload multiple documents
- SmartDB (Smart Database) first identifies the most relevant document
- Answers are generated only from that single document
- This mode is ideal when documents represent distinct policies, contracts, or manuals.
 - Routing Strategy (Document-Specific Search Only)
 - When Document-Specific Search is selected as shown in **figure 9**, SmartDB provides two routing mechanisms to determine the most relevant document:
 - **Intent-Based Routing**
 - Each document is associated with an intent description
 - The system matches the user question to the defined intent
 - Recommended when documents have clearly defined purposes or domains
 - **Embedding-Based Routing**
 - Uses semantic similarity between the user question and document embeddings



- Automatically selects the most relevant document
- Best suited for large or unstructured document collections

Figure 9: Document Routing Configuration for Document-Specific Search

2.1.2.2.1.2.3 Embedding & Retrieval Parameters

This section controls how documents are segmented, indexed, and retrieved using vector similarity. Administrators can configure similarity thresholds, maximum retrieved chunks, segment size, overlap size, and contextual chunk expansion to balance accuracy, coherence, and performance during document search. (Refer to **Figure 10** – Embedding and Retrieval Configuration)

Figure 10: Embedding and Retrieval Configuration

Below is the filed description & details:

- **Embedding Similarity Minimum Score (%)**
 - Defines the minimum similarity threshold required for a text chunk to be considered relevant.
 - Higher values improve precision, while lower values increase recall.
- **Embedding Max Result**



- Specifies how many relevant text chunks can be retrieved per question, provided they meet the similarity threshold.
- **Max Segment Size (Characters)**
 - Controls how documents are split into text chunks during indexing.
 - Smaller chunks improve semantic accuracy, while larger chunks preserve context.
 - Recommendation:
 - For 14B models and below, **300 characters or less**
 - For large models, this value can be increased if needed
- **Max Overlap Size (Characters)**
 - Defines how many characters overlap between adjacent text chunks to maintain contextual continuity.
 - ⚠ This value **cannot be changed after saving** the library.
- **Fetch Adjacent Chunks (Before / After)**
 - Determines how many text chunks before and after the matched chunk are included to enhance contextual coherence.
 - This approach is preferred over large overlaps for better retrieval quality.



2.1.2.2.1.2.4 LLM Generation Parameters

This section allows fine-tuning of the LLM's response behavior, including temperature, Top-P, Top-K, repetition penalty, maximum tokens, and chat window size. These parameters control the creativity, determinism, length, and consistency of AI-generated answers. (Refer to **Figure 11** – LLM Generation Settings)

The screenshot shows a settings panel with the following parameters and descriptions:

- Temperature:** 0.2. This parameter controls the randomness of the output. A value closer to 0.0 will produce more deterministic results, while a value closer to 1.0 will yield more varied and creative outputs.
- Top-P:** 0.8. This parameter controls the diversity of the model's output by considering only the most probable words whose cumulative probability exceeds a specified threshold (p). A lower Top-P value results in more focused and coherent responses, while a higher value allows for a broader range of creative and varied outputs.
- Top-K:** 40. This parameter limits the model's output to the top K most probable next words during generation. A lower Top-K value results in more focused and coherent responses, while a higher value increases diversity and creativity by allowing a wider selection of potential outputs.
- Repeat Penalty:** 1.1. This parameter discourages the model from generating repetitive phrases or words by applying a penalty to previously used tokens. A higher penalty value promotes more varied and engaging responses, while a lower value allows for some repetition, which can be useful in certain contexts.
- Num Predict:** 600. This parameter specifies the number of tokens the model should generate in response to a prompt. This parameter allows you to control the length of the output, enabling you to tailor responses to be concise or more detailed based on your needs. Adjusting this value can help ensure that the generated text is appropriately informative without being overly verbose.
- Chat Window Size:** 0.

Figure 11: LLM Generation Settings

Below is the filed description & details:

- **Temperature**
 - Controls response randomness:
 - Lower values → more accurate, deterministic answers
 - Higher values → more creative and varied responses
- **Top-P**
 - Limits token selection based on cumulative probability:
 - Lower values → more focused, accurate answers
 - Higher values → increased creativity
- **Top-K**
 - Restricts token selection to the *K* most probable words at each step.
 - Helps balance creativity and factual grounding.
- **Repeat Penalty**
 - Prevents repetitive responses by penalizing repeated tokens.
 - Higher values reduce repetition, especially in longer answers.



- **Max Tokens (Num Predict)**
 - Defines the maximum number of tokens the model can generate in a single response.
- **Chat Window Size**
 - Controls how much previous conversation history is retained to maintain context in multi-turn conversations.

2.1.2.2.1.2.5 Prompt & Behavior Configuration

This section defines the system prompt used to instruct the LLM on how to answer user questions. Administrators can customize the prompt while preserving required placeholders (such as `{{question}}` and `{{information}}`) to ensure correct operation. A fallback response can also be defined for cases where no relevant answer is found. (Refer to **Figure 12** – Prompt Configuration Section)

Main Prompt

Answer the following question to the best of your ability in same language of the question
Question:
{{question}}

No Answer Message

Sorry, I don't have the information on that.

Figure 12: Prompt Configuration Section

Below is the filed description & details:

- **Main Prompt**
 - Defines the instruction used by SmartDB (Smart Database) when asking the LLM to generate answers.
 - ⚠ The placeholders `{{question}}` and `{{information}}` must always be present, otherwise document answering will fail.
- **No-Answer Message**
 - Displayed when SmartDB cannot find relevant information in the document library.

2.1.2.2.1.2.6 Safety, Validation & FAQ Controls

This section enables governance and safety features, including chit-chat detection, FAQ prioritization, follow-up question handling, input validation, blacklisted terms, unwanted topics, and safety response messages. These controls ensure that SmartDB (Smart Database) delivers compliant, relevant, and secure AI responses aligned with organizational policies. (Refer to **Figure 13**



– Safety and Validation Controls)

- Chitchat User Query Validation
- Frequently Asked Questions
- Enable Follow Up Questions Handling
- Enable Adding Context To User Inputs
- Safety Query Validation

Figure 13: Safety and Validation Controls

Below is the filed description & details:

- **Chitchat Detection**
 - Detects casual or conversational queries (e.g., greetings) and prevents unnecessary document searches.
- **Frequently Asked Questions (FAQ)**
 - Allows predefined answers to common questions.
 - Admins can enable or disable using FAQs as a response source.
- **Follow-Up Questions Handling**
 - Enables or disables conversational follow-ups such as “tell me more” or “continue”.
- **Adding Context to User Inputs**
 - Enable or disable system reform the user questions in some cases user may require it
- **Safety Query Validation**
 - Validates user input against (Refer to **Figure 14** – Safety Query Validation):
 - Blacklisted words
 - Unwanted topics
 - Safety Filter Message - If triggered, SmartDB (Smart Database) responds with a predefined **Safety Filter Message**.

Safety Filter Message

Sorry, I can't assist with that.

Blacklisted Words

One word or phrase per line. Words are checked case-insensitively. Lines starting with # are treated as comments.

fuck
shit

Unwanted Topics

One topic phrase per line. These are used for semantic similarity matching. Lines starting with # are treated as comments.

How to hack into a system
How to create a bomb

Figure 14 – Safety Query Validation



2.1.2.2.1.3 Upload Documents (Upload for AI-Powered Understanding)

2.1.2.2.1.3.1 Access Document Library

After creating a Document Library, administrators can proceed to upload and manage documents associated with that library through the Documents Library Management page.

As shown in **Figure 15**, each document library row includes an Upload Documents action icon. By clicking the upload icon in the Action column, the user is redirected to the Upload Documents page for the selected library.

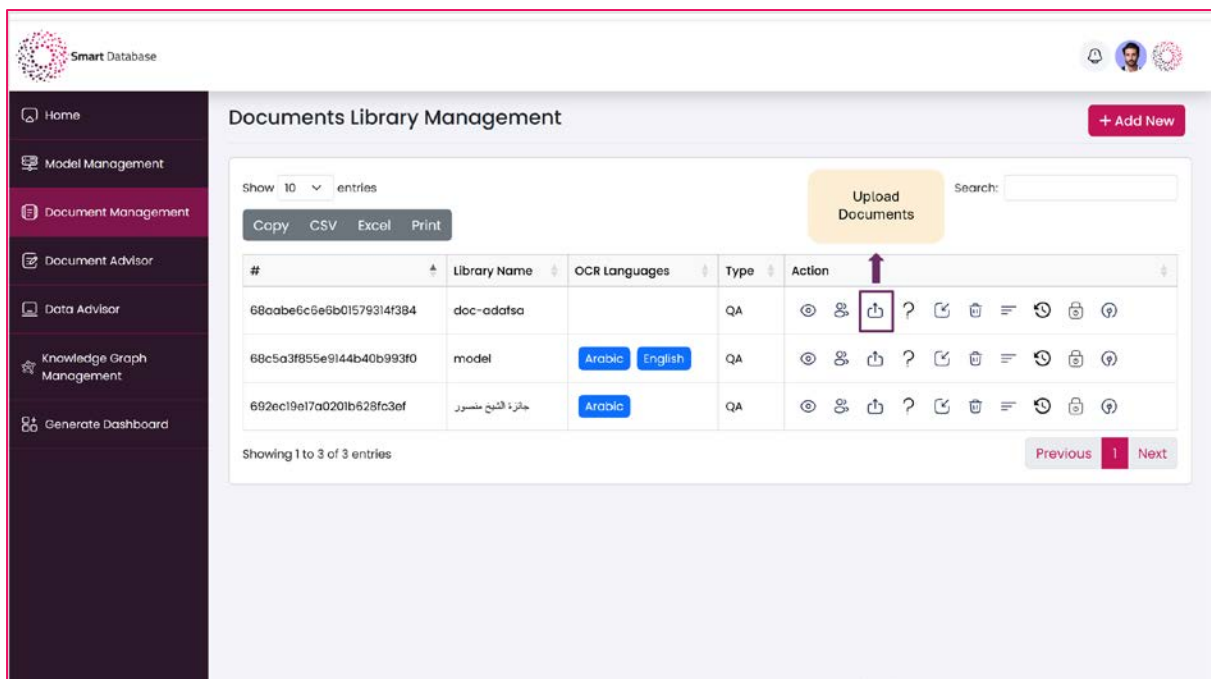


Figure15: Accessing Upload Documents

This action establishes the link between the document library configuration and the actual document content that will be ingested, indexed, and used by SmartDB’s AI engines for question answering and document advisory services.

2.1.2.2.1.3.2 Upload Documents

The Upload Documents section enables administrators and authorized users to ingest documents into SmartDB (Smart Database) so they can be analyzed, indexed, and used for AI-powered question answering and document advisory. Once uploaded and indexed, documents become part of the knowledge base and can be queried using natural language through the Document Advisory interface or via APIs.



As shown in **Figure 16**, SmartDB provides multiple flexible options for uploading documents, allowing organizations to ingest content from local machines, folders, or server-side repositories while maintaining full control over document processing and indexing.

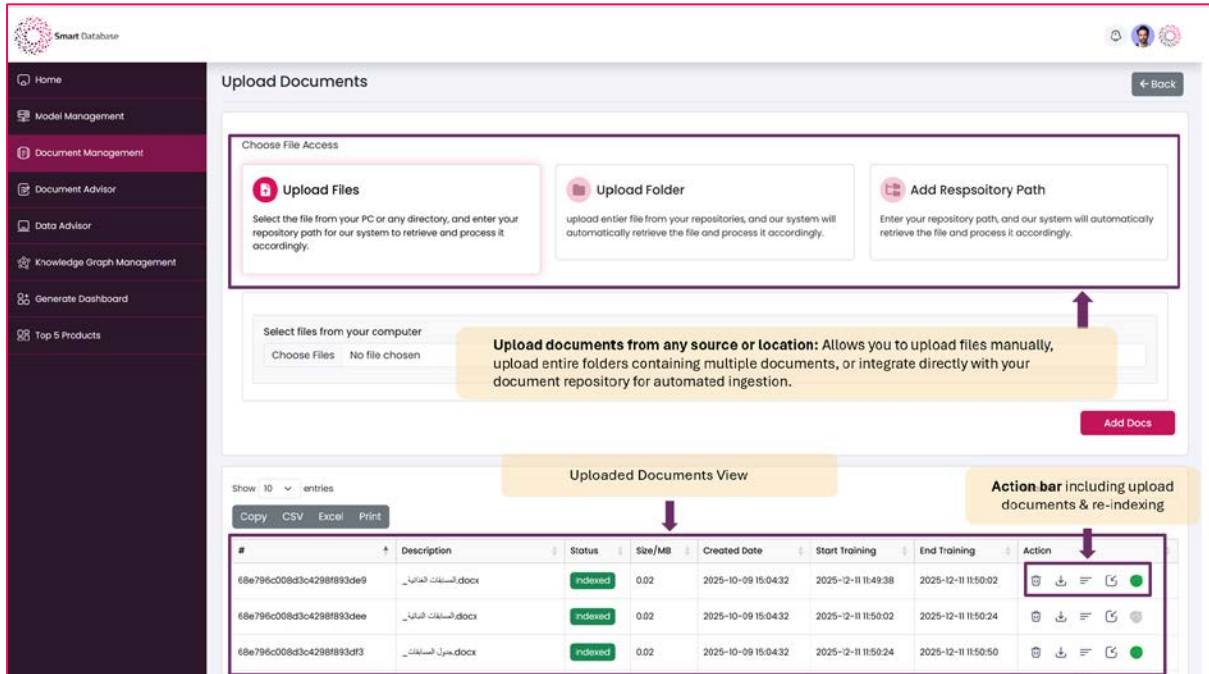


Figure16: Document Ingestion and Indexing Configuration Screen

2.1.2.2.1.4 Document Upload Methods

Document Upload Methods authorized users to add, manage, and prepare documents for AI-powered interaction within SmartDB (Smart Database). Through this interface, users can upload documents from multiple sources.

Users can upload documents using one of the following methods:

1. Upload Files

Users can upload individual files directly from their local machine. This option is suitable for adding specific documents or updating content incrementally. Uploaded files are immediately registered in the system and marked with a **Pending** status until indexing is initiated.

2. Upload Folder

This option allows users to upload an entire folder containing multiple documents from their local machine. Smart Database automatically retrieves all supported files within the folder and processes them in bulk, significantly reducing manual effort for large document sets.

3. Add Repository Path

Administrators can specify a **server-side or repository path** (for example, a shared file system or



mounted directory). SmartDB automatically retrieves documents from the specified path and processes them accordingly. This method is ideal for integrating enterprise document repositories or shared storage locations.

2.1.2.2.1.4.1 Supported File Formats

SmartDB supports a wide range of document formats to ensure compatibility with enterprise and government content, including:

- **Document & Text Formats:** PDF, HTML, XML, RTF, TXT, EPUB, CHM, ODF, OpenDocument, iWorks (Apple), Markdown, LaTeX
- **Microsoft Office:** .doc, .docx, .xls, .xlsx, .ppt, .pptx
- **Images:** PNG, GIF, BMP, JPEG, TIFF

If no readable text is detected during parsing, SmartDB automatically applies **OCR (Optical Character Recognition)** based on the library's configured OCR languages.

2.1.2.2.1.5 Document Ingestion & Processing

This section provides a centralized view of all documents ingested into the SmartDB Document Management system and their current processing status. It allows administrators and authorized users to monitor, manage, and control the full document lifecycle—from upload and AI indexing to reprocessing and removal.

It displays all uploaded documents along with their indexing and AI processing status. Enables users to manage document ingestion, monitor training progress, and perform actions such as re-indexing, configuration, download, or deletion to ensure documents are fully prepared for AI-powered search and chat.

Each row in this table represents an individual document and displays key metadata, including the document identifier, file description, processing status, file size, creation date, and AI training timestamps. The status indicator reflects whether the document is pending, indexed, or undergoing processing by the AI engine.

As shown in **Figure 17**, the Document Ingestion & Processing section provides real-time visibility into uploaded documents and their AI indexing lifecycle.



#	Description	Status	Size/MB	Created Date	Start Training	End Training	Action
68e796c008d3c4298f893de9	المسائلات الذاتية.docx	Indexed	0.02	2025-10-09 15:04:32	2025-12-11 11:49:38	2025-12-11 11:50:02	
68e796c008d3c4298f893dee	المسائلات الذاتية.docx	Indexed	0.02	2025-10-09 15:04:32	2025-12-11 11:50:02	2025-12-11 11:50:24	
68e796c008d3c4298f893df3	جدول المسائلات.docx	Indexed	0.02	2025-10-09 15:04:32	2025-12-11 11:50:24	2025-12-11 11:50:50	

Figure17: Document Ingestion and processing View

The **Uploaded Documents View** displays a detailed table of all documents ingested into the selected document library. This section serves as the operational control center for document processing and indexing.

Each document entry provides:

- **Document identifier and description**
- **Processing status** (e.g., Pending, Queued, Indexed)
- **File size**
- **Creation date**
- **Indexing start and end timestamps**
- **Action bar** for document operations

From this table, users can:

- Trigger document indexing or re-indexing
- Access document configuration
- Delete or download documents
- Monitor AI training progress in real time

2.1.2.2.1.6 Document Configuration

As illustrated in the **figure 18**, document-specific configuration is accessed directly from the **Action column** within the Uploaded Documents View.

When the Document Configuration icon is selected:

- A configuration panel opens on the side of the screen
- The main Upload Documents page remains visible in the background
- Users can adjust document-level processing parameters without leaving the ingestion context

This design enables fast iteration and fine-grained control over how individual documents are processed and indexed.



#	Description	Status	Size/MB	Created Date	Start Training	End Training	Action
68e796c008d3c4298f893de9	المسائل القانونية.docx	Indexed	0.02	2025-10-09 15:04:32	2025-12-11 11:49:38	2025-12-11 11:50:02	[Icons]
68e796c008d3c4298f893dee	المسائل القانونية.docx	Indexed	0.02	2025-10-09 15:04:32	2025-12-11 11:50:02	2025-12-11 11:50:24	[Icons]
68e796c008d3c4298f893df3	جدول المسائل.docx	Indexed	0.02	2025-10-09 15:04:32	2025-12-11 11:50:24	2025-12-11 11:50:50	[Icons]

Figure18: Document Configuration Panel Navigation

The Document Configuration panel provides advanced controls for fine-tuning how a specific document is processed before indexing. These settings apply at the document level and override default library behavior where applicable.

Note: The configuration options shown apply to **PDF files only**, as indicated in the interface.

2.1.2.2.1.6.1 Document Indexing Settings

The **Document Indexing Settings** section allows administrators to control how document content is extracted, cleaned, and prepared for AI indexing. The configuration panel is shown in **figure 19**.

The screenshot shows the 'Document Indexing Settings' panel with the following fields:

- Pages: All
- Delete this many lines from the start: 0
- Delete this many lines from the end: 0
- Enforce OCR:
- SAVE button

Figure19: Document Configuration Panel View

The document configuration section includes:

Page Selection

- Choose to process all pages or specify custom page ranges
- Useful for excluding irrelevant sections such as covers, appendices, or annexes



Line Trimming Controls

- Delete lines from the beginning of each page
- Delete lines from the end of each page
- Commonly used to remove headers, footers, page numbers, or repeated legal text

Enforce OCR

- Forces SmartDB to convert PDF pages into images and apply OCR
 - This option is recommended when working with encrypted or secured PDF files, where direct text extraction is not possible.
- Ensures accurate text extraction when:
 - PDFs are scanned
 - Embedded text is missing or unreliable

These controls ensure higher-quality text extraction, leading to better embeddings, improved retrieval accuracy, and more reliable AI responses.

Purpose & Impact of Document Configuration

By enabling document-level configuration, Smart Database allows organizations to:

- Improve AI answer accuracy by removing noise and irrelevant content
- Handle heterogeneous document formats consistently
- Optimize indexing for large, complex, or scanned documents
- Maintain full transparency and control over how documents are interpreted by AI

Together, the **Upload Documents** page and **Document Configuration** panel form a critical foundation for transforming enterprise documents into trusted, AI-ready knowledge assets.

2.1.2.2.1.7 Document Intent Methods

As shown in the **figure 20**, the Intent Configuration page allows administrators to define a clear and unique intent description for each document. This intent represents the semantic purpose and scope of the document and is used by SmartDB to accurately determine whether the document is relevant to answer a user's question.

The intent description acts as a high-level semantic summary of the document and plays a critical role in document selection and routing when users interact with the AI.



This option is available only when the Document Type is set to *Document-Specific Type*.

Selecting this type activates the **routing methodology**, where the **Intent** defines how the document will be classified, processed, and routed within SmartDB (Smart Database).

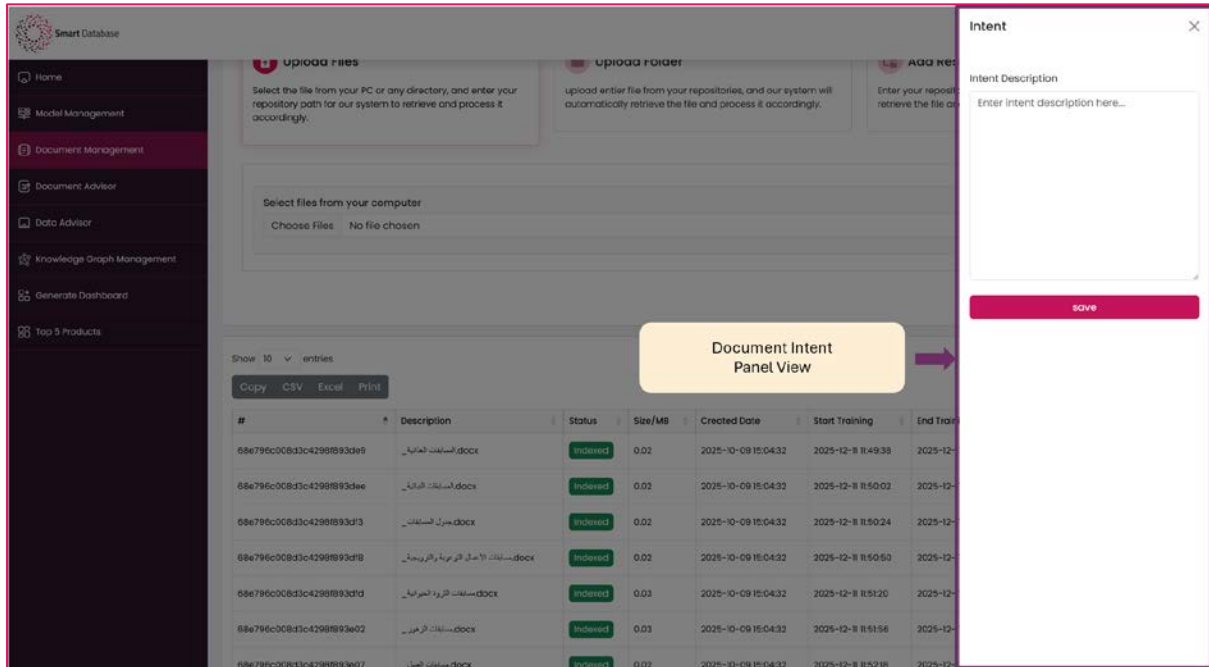


Figure20: Document Intent Panel View

The Intent Configuration page provides a powerful mechanism for controlling document selection logic by explicitly defining the purpose of each document. By leveraging intent-based validation, SmartDB (Smart Database) ensures that user questions are answered using the right document, for the right reason, at the right time.

2.1.2.2.1.7.1 Purpose of Intent Configuration

The Intent Configuration feature is used to improve **precision and relevance** when answering user questions in document-specific scenarios. Instead of relying only on text similarity, SmartDB evaluates the user's question against the defined document intent to decide whether the document should be selected for answering.

This ensures that:

- Only the **most relevant document** is considered
- Cross-document ambiguity is avoided
- Answers remain aligned with the document's actual purpose



2.1.2.2.1.7.2 Intent Description Field

The **Intent Description** field allows administrators to enter a **unique and meaningful description** that represents what the document is about, such as:

- The business domain it covers
- The type of information it contains
- The kinds of questions it is expected to answer

When a user submits a question, SmartDB validates the question against this intent description to determine if the document is eligible to respond.

Best practices for defining intent:

- Be concise but descriptive
- Focus on the document's purpose, not its structure
- Avoid generic or overlapping intent descriptions across documents

2.1.2.2.1.7.3 Applicability & Scope

Important:

The Intent Configuration feature is **only applicable to Document Libraries configured with the *Document-Specific Search* type.**

For libraries using:

- **Direct** or
- **All-Documents Search,**

The intent configuration is not required and is not used in document routing.

2.1.2.2.1.7.4 Role in Document Routing Strategy

When **Intent-Based Routing** is enabled for a Document-Specific Search library:

1. The user's question is analyzed
2. The question is compared against each document's intent description
3. The document whose intent best matches the question is selected
4. Only that document is used for chunk retrieval and answer generation



This approach significantly improves answer accuracy, especially in environments with many documents covering different topics.

2.1.2.2.2 Managing User Access and Privileges for Document Libraries

The Documents Library Management page allows administrators not only to create and manage document libraries, but also to control user access and privileges for each library.

As shown in **Figure 21**, each document library includes an Add Privilege to Users action icon within the *Action* column. By clicking this icon, administrators can assign, modify, or revoke user permissions for the selected document library.

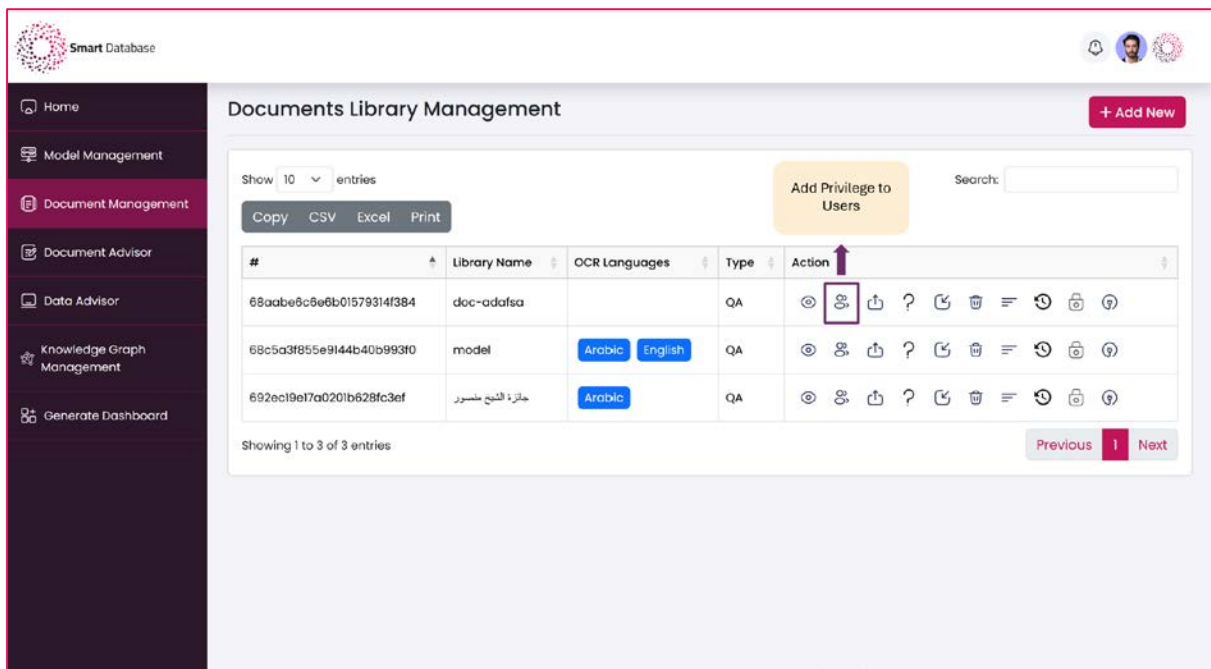


Figure21: Accessing User Privileges Configuration

2.1.2.2.2.1 Accessing User Privileges Configuration

To manage user privileges for a document library:

1. Navigate to Document Management → Documents Library Management
2. Locate the desired document library
3. Click the Add Privilege to Users icon (highlighted in **Figure 21**)
4. The system opens the user access configuration interface for that library



2.1.2.2.2.2 Add User to Document Library

The **Add User to Library** page is used to manage and control **user-level access** to a specific document library within SmartDB (Smart Database). This page enables administrators to grant or revoke user permissions, ensuring that only authorized users can access, manage, or interact with documents and AI-powered document advisory features.

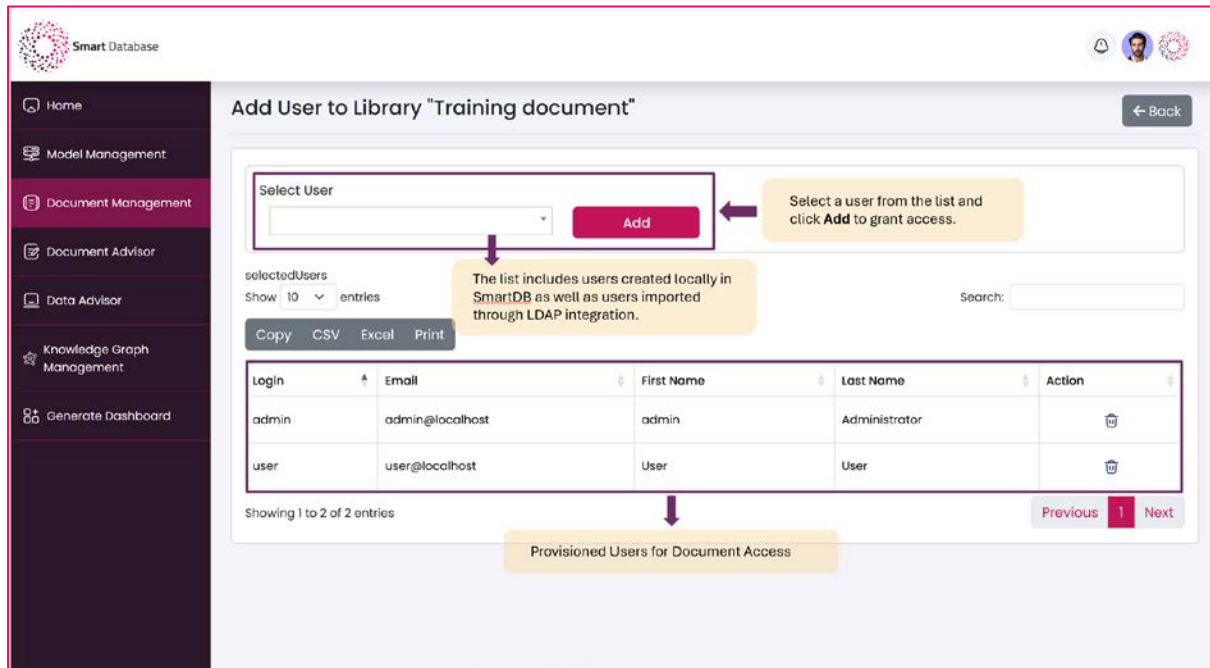


Figure 22: Adding and Managing User Access for a Document Library

As shown in **Figure 22**, this page allows administrators to select users and assign them access to the selected document library.

This page supports **role-based and library-level access control**, allowing organizations to enforce data security, confidentiality, and governance policies. It ensures that document libraries can be shared selectively across users, teams, or departments.

Selecting Users for Library Access

At the top of the page, administrators can select a user from the Select User dropdown list.

This list includes:

- Users created locally within SmartDB (Smart Database)
- Users synchronized through **LDAP integration**

Once a user is selected, clicking the **Add** button grants the user access to the document library.



2.1.2.2.3 FAQ Configuration (Library-Level FAQ Control)

The Document FAQ Management feature provides a structured governance layer over AI-generated responses by allowing administrators to define authoritative, pre-approved questions and answers at the document library level. These FAQs are evaluated before invoking open-ended AI reasoning, ensuring that SmartDB delivers accurate, consistent, and policy-compliant answers for predefined scenarios.

This capability is critical for organizations that require controlled messaging, regulatory compliance, or official responses—particularly in government, legal, financial, and enterprise environments. By prioritizing FAQs, SmartDB prevents response variability and eliminates the risk of hallucinated or non-compliant answers for high-impact or sensitive topics.

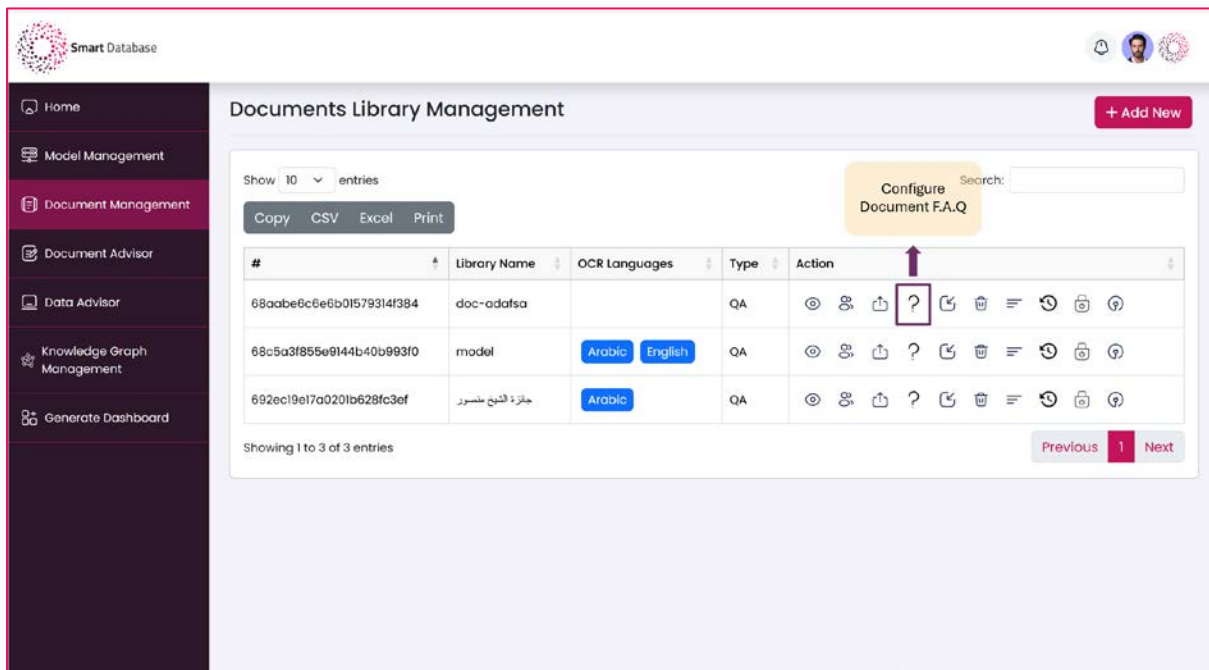


Figure 22: Accessing FAQ Configuration (Library-Level FAQ Control)

As shown in **Figure 23**, this page allows administrators to select users and assign them access to the selected document library, ensuring that FAQ management and visibility are restricted to authorized personnel only. This access control mechanism supports role-based governance and aligns FAQ configuration with organizational security policies.

Key Capabilities and Benefits:

- **Priority-Based Answering**

When a user submits a question, SmartDB first checks the configured FAQs for semantic



similarity. If a matching FAQ is found above the defined similarity threshold, the system returns the approved answer immediately—bypassing dynamic LLM generation.

- **Controlled and Approved Messaging**

FAQs ensure that responses align with organizational policies, legal frameworks, and official communications. This is especially important for public-facing portals, regulatory content, and institutional knowledge bases.

- **User Access Control and Governance**

Administrators can explicitly manage which users are permitted to view, manage, or maintain FAQs for each document library, ensuring separation of duties and secure content ownership.

- **Consistency Across Users and Channels**

All users receive the same standardized answer for the same question, regardless of phrasing or language variation, ensuring uniformity across departments, platforms, and user roles.

- **Support for Sensitive and High-Risk Topics**

Administrators can define fixed responses for sensitive subjects (e.g., legal interpretations, eligibility criteria, official procedures, financial thresholds), reducing exposure to compliance and reputational risk.

- **Operational Efficiency**

Frequently asked or repetitive questions are handled instantly without invoking full AI inference, improving response time and reducing system load.

- **Multilingual and Library-Aware Context**

FAQs are defined at the document library level and can be configured per language, ensuring responses remain aligned with the underlying document corpus and OCR language settings.

2.1.2.2.3.1 Accessing Document FAQ Configuration

To manage and configure the Document FAQ for a document library:

1. Navigate to **Document Management** → **Documents Library Management**
2. Locate the desired document library from the list
3. Click the **Configure Document FAQ** icon (?), as highlighted in **Figure 23**
4. The system opens the **Document FAQ configuration interface**, where FAQs can be created, updated, and managed for the selected library

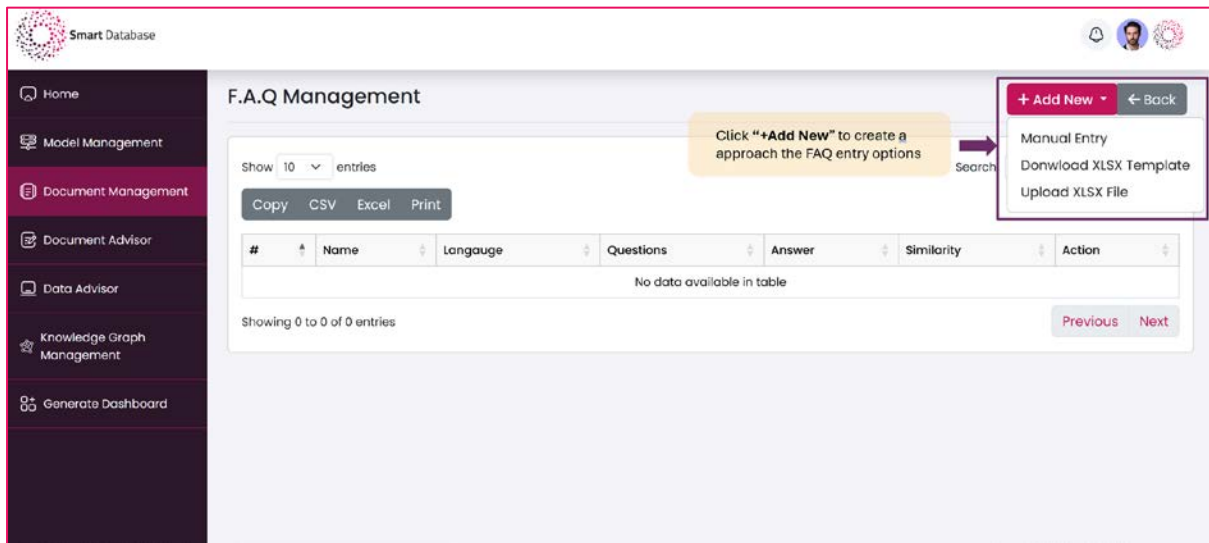


Figure 24: Add FAQ (Manually or Bulk creations)

Administrators can add FAQs using one of the following methods:

- **Manual Entry:**
Allows the creation of individual FAQ records by manually entering the question, answer, language, and similarity configuration directly through the system interface.
- **Download XLSX Template:**
Enables administrators to download a predefined Excel template that defines the required structure for FAQ data. This template ensures consistency and correctness when preparing multiple FAQ entries offline.
- **Upload XLSX File:**
Allows administrators to upload the completed XLSX template back into the system to bulk import multiple questions and answers in a single operation. Upon upload, SmartDB validates the data and registers the FAQ entries for the document library.

2.1.2.2.3.2.1.1 Manual Entry

The **Manual Entry** page enables administrators to create and manage individual FAQ records that will be used by SmartDB to deliver controlled, accurate, and policy-aligned answers when users interact with documents.

This interface allows fine-grained configuration of how FAQs are detected, matched, and returned during AI-powered document conversations.

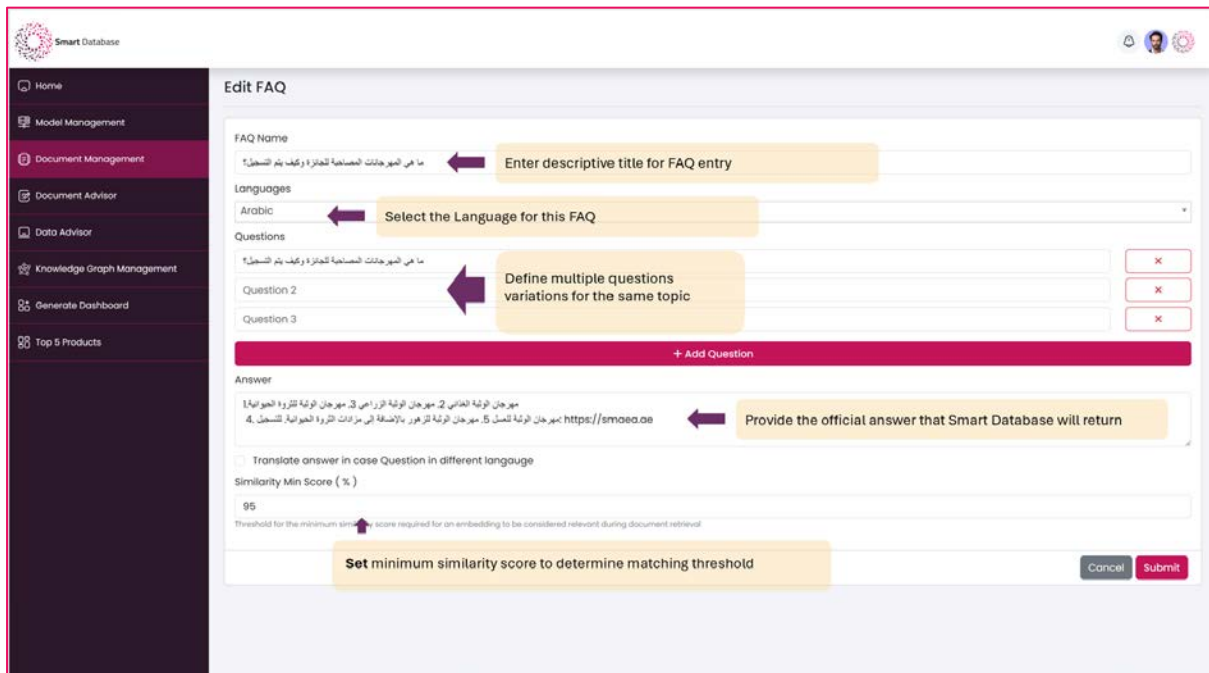


Figure 24: FAQ Manual Entry

Key configuration elements include:

- **FAQ Name**
A descriptive title used to identify the FAQ entry internally. This helps administrators manage and maintain large FAQ libraries efficiently.
- **Language Selection**
Each FAQ is language-specific. The selected language determines how the FAQ is matched against user questions and ensures responses are delivered in the appropriate language context.
- **Questions (Multiple Variants Supported)**
Administrators can define multiple question variations that relate to the same topic. This allows SmartDB to recognize different ways users may ask the same question while returning a single, unified answer.
- **Default Answer**
The official and authoritative response that SmartDB will return when a user's question matches the defined FAQ.
This answer takes precedence over AI-generated responses to ensure consistency and accuracy.
- **Answer Translation Option**
When enabled, SmartDB can automatically translate the FAQ answer if the user submits the



question in a different language, ensuring multilingual support without duplicating FAQ entries.

- **Similarity Minimum Score (%)**

Defines the confidence threshold required for SmartDB to classify a user question as matching this FAQ.

It is strongly recommended to keep this value **above 95%** to avoid incorrect matches and ensure the FAQ is triggered only for highly relevant questions.

- **Submission Controls**

Administrators can save or cancel changes after reviewing all configured fields, ensuring full control before publishing the FAQ entry.

This manual configuration approach is ideal for high-priority, sensitive, or frequently asked questions where precision, compliance, and consistency are critical.

2.1.2.2.3.2.1.2 Upload Bulk FAQ Entry

The **Upload Bulk FAQ Entry** feature enables administrators to efficiently create and manage large volumes of Frequently Asked Questions (FAQs) for a document library using a standardized Excel template. This approach is designed for scalability, consistency, and accuracy when onboarding multiple questions and answers at once.

As shown in the **figure 25**, bulk FAQ management follows a **two-step process**: downloading the Excel template, completing it with FAQ data, and uploading it back into SmartDB (Smart Database) for automated ingestion and validation.

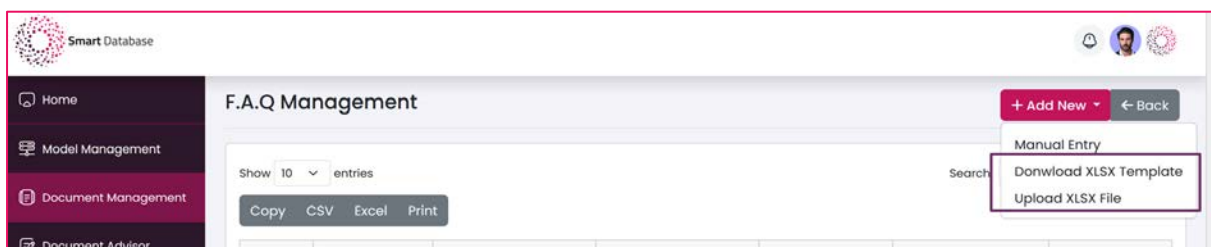


Figure25: download & Upload Access for FAQ bulk Upload

Step 1: Download the FAQ Excel Template

Administrators must first download the predefined XLSX template provided by SmartDB (Smart Database). This template defines the exact structure and fields required for bulk FAQ ingestion.

All columns in the template are **mandatory**, and the system validates the file format before processing.



Figure 26 illustrates the process of downloading the FAQ Excel template, which is used to prepare and upload FAQ data in bulk.

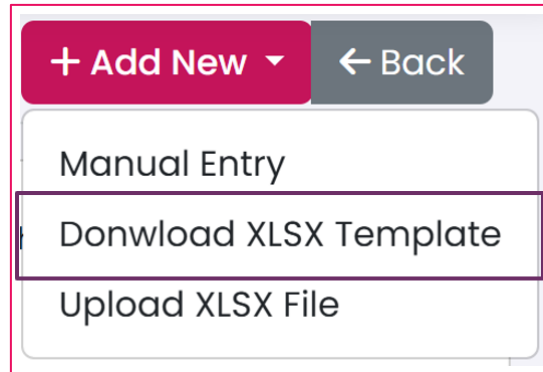


Figure 26: Download FAQ Template

FAQ Excel Template Structure (Attachment Description)

The figure 27 illustrates the structure of the FAQ Excel template. Each row represents a single FAQ entry, and the following columns must be completed:

	A	B	C	D	E	F
1	name	question	answer	language	similarity	
2	Q1	When the deadline date ?	1/1/2025			96
3				Afrikaans		
4				Albanian		
5				Amharic		
6				Arabic		
7				Assamese		
8				Azerbaijani		
9				Azerbaijani - Cyrillic		
10				Basque		
11				Belarusian		
12				Bengali		
13				Bosnian		
14				Bulgarian		
15						
16						

Figure 27: FAQ Excel Structure for Bulk Upload

The Excel file is provided in a predefined format containing the following structure and fields:

- **Name**
A unique identifier or label for the FAQ entry (e.g., Q1, Q2). This is used internally for reference and management.
- **Question**
The user-facing question. This field represents how users may ask the question in natural language.
- **Answer**
The authoritative response that SmartDB will return when the question is matched. This answer overrides AI-generated responses when FAQ matching conditions are met.



- **Language**

Specifies the language of the question and answer. The template provides a controlled dropdown list of supported languages to ensure consistency and correct multilingual handling.

- **Similarity**

Defines the minimum similarity threshold (e.g., 96) required for SmartDB to classify a user query as matching this FAQ. Higher values ensure stricter matching and reduce false positives.

Important: All fields are mandatory. Incomplete or incorrectly formatted rows will be rejected during upload.

Step 2: Fill and Upload the Template

After completing the template with the required FAQ information, the file is uploaded through the **Upload FAQ** screen shown in **Figure 28**. Once uploaded, SmartDB validates the data, applies similarity thresholds, and registers the FAQs under the selected document library.

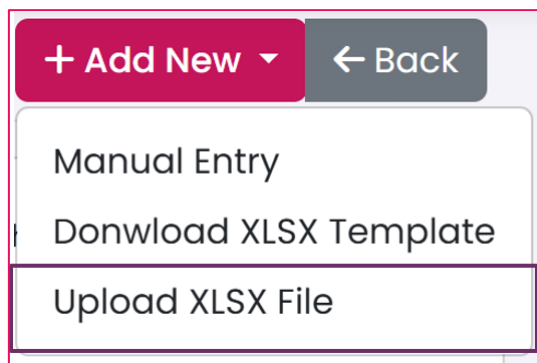


Figure 28: Accessing Upload FAQ File

Once the Excel file is prepared, administrators navigate to the **Upload FAQ** page as shown in the **Figure 29**:

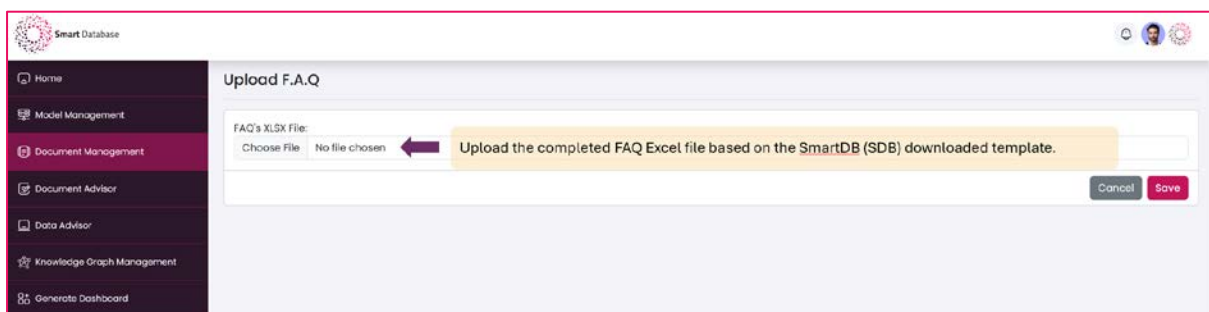


Figure 29: Upload FAQ File Page



1. Select the completed XLSX file using the file picker.
2. Click **Save** to upload and process the FAQs.
3. The system validates the file, ingests the data, and registers the FAQs under the relevant document library.

Successfully uploaded FAQs become immediately available for use during document question answering, ensuring consistent, approved, and policy-aligned responses.

2.1.2.2.3.3 Documents Library Chat History

As shown in figure 30, the **Documents Library Chat History** page provides a complete audit trail of all user interactions and AI responses related to a specific document library. This page enables administrators and authorized users to review, monitor, and analyze how SmartDB (Smart Database) responds to user questions based on the indexed documents and configured rules.

The chat history is displayed in a structured table format, listing each interaction with clear separation between **user inputs** and **system-generated responses**, along with precise timestamps.

#	Role	Content	Created Date
1	system	الجانزة المختصة بجمع النحل هي مسابقة أفضل قلب شمع طبيعي. الجائزة تقام بالمرکز الأول في هذه المسابقة هي 15,000، والمرکز الثاني 10,000، والمرکز الثالث 8,000، والمرکز الرابع 5,000، والمرکز الخامس 3,000.	2025-12-11 14:49:52
2	user	ما هي الجائزة المختصة بجمع النحل	2025-12-11 14:49:45
3	system	الشرط التي ذكرتها تتلخص بمسابقة مختلفة ضمن مهرجان الزهور. فيما يلي الشروط الخاصة بكل مسابقة: ### مسابقة أفضل قلب زهور - فة الإقبال: تكون الفة العمرية للشاركين من 8-13 سنة. 2. يتم اختيار عدد محدود من الشاركين بناء على أسس التسجيل ومدى اهتمام الشار بالمشروع. 3. يمكن للشاركين إرفاق ملف بسيط اختاري يعرض عن تعليمهم أو تجربتهم مع الزهور. 4. يحق لكل شاركون اختيار وتسجيل نفسه باستخدام الزهور المتاحة فقط. 5. الزهور يتم توفيرها من قبل الجهة المنظمة. 6. يسمح باستخدام الإكسسوارات البسيطة التي تقدمها الجهة المنظمة. 7. يجب أن تتم البقاء بالإبداع والابتكار في ترتيب الزهور. 8. يجب أن يتم التزم الشار بأكمل بتسجيل بقاء الزهور خلال فترة زمنية محددة. 9. لا يسمح بتلقي المساعدة من أي شخص أثناء تافها البقاء. ### مسابقة أفضل بقاء زهور - فة المخرافين: 1. الزهور يتم توفيرها من قبل الجهة المنظمة. 2. المشاركة مقصورة على الأفراد الذين يمتلكون مهارات أو خلفية علمية أو فنية في تنسيق الزهور. 3. يجب تقديم ما يثبت الخبرة أو المهارة، مثل شهادات تدريب، صور، أو أفضل مدياق وروابط لعصافات تعرض الإحلال، أو مشاركات في معارض وأحداث ذات صلة. 4. يجب أن تكون المشاركون ببقاء بقاءهم دون الإستهانة بأي طرف خارجي. 5. يجب أن يتم المشاركة بمسارسات زراعية مستدامة وصديقة للبيئة. 6. يجب تقديم زهور عالية الجودة وإحلالها من الجوز. 7. يجب إرفاق صور أو فيديو ذات صلة توضح جدول التنبيه الزرع. 8. يجب أن يظهر المشاركون الإبداع في تصميم وتعليم الزهور. ### مسابقة أفضل منتج تعويبي من الزهور: 1. يجب أن يكون المنتج مصنوعاً بشكل رئيسي من الزهور أو مستخلصاتها. 2. يجب أن يكون المنتج ممتكراً أو يقدم فكرة جديدة في طريقة الاستخدام أو الشكل النهائي. 3. يجب أن يكون المنتج قابلاً للاستخدام أو التسويق، وأن يوجد عالية. 4. يجب أن يكون المنتج أمناً للاستخدام البشري أو الحيواني، وخالياً من مواد سامة أو محظورة. 5. يجب إرفاق وصف موجز يوضح فكرة التصنيع، المواد المستخدمة، وطريقة الاستخدام أو الفائدة. ### مسابقة أفضل حديقة زهور منزلية الشروط لا تعدد بوضوح، ولكنها تتضمن معايير التقييم مثل الإبداع، الجمالية، الابتكار والتجديد، والاستدامة. ### مسابقة أفضل زهرة زهور في الإبرات الشروط لا تعدد بوضوح، ولكنها تتضمن معايير التقييم مثل التزيح والورد، الإثنية والقيمة الاقتصادية والقائد المضمعة. يرجى الإتهاد إلى أن كل مسابقة لها شروط	2025-12-11 13:33:23
4	user	ما هي الشروط	2025-12-11 13:32:57
5	system	وفقا للمعلومات المقدمة، فإن مسابقة أفضل منتج تعويبي من الزهور (الزهور المطرية، المنتجات الطبيعية، المنتجات الغذائية) هي المسابقة الوحيدة التي تتطلب تقديم نموذج مرسوم بوضوح التفاصيل الأساسية للتصميم	2025-12-11 12:02:00
6	user	أي من هذه المسابقات تحتاج إلى نموذج مرسوم؟	2025-12-11 12:01:55
7	system	جميع مسابقات الزهور تشمل الفئات التالية: 1. أفضل بقاء زهور (فئة الإقبال) 2. أفضل بقاء زهور: 3. أفضل حديقة زهور منزلية (فرد - فرداً زراعية منزلية) 4. أفضل زهرة زهور في الإبرات (مزارع تجارية / مزارعون متمسكون). 5. أفضل منتج تعويبي من الزهور (الزهور المطرية، المنتجات الطبيعية، المنتجات الغذائية) 6. أفضل تصميم لحديقة مسجارية (فرد وطلاب جامعات) 7. أفضل صورة فوتوغرافية لمهرجان الزهور (فئة الجواد والجمهورية). 8. أفضل معارسة ابتداعية في مجال الزهور. كل فئة لها شروطها وضوابطها الخاصة، وتتضمن الجوائز المالية المختلفة حسب المركز الذي يحققه المشاركون في المسابقات	2025-12-11 12:01:40
8	user	ما هي جميع مسابقات الزهور؟	2025-12-11 12:01:32

Figure 30: Documents Library Chat History View

The Chat History page is contained the following structure and fields:

- **Conversation Timeline**
Displays the full sequence of questions and answers exchanged between users and SmartDB, sorted from the most recent interaction to the oldest.



2.1.2.2.3.4.1 When to Use Re-Indexing

Re-indexing should be performed in the following scenarios:

- After uploading new documents to the library
- After modifying or deleting existing documents
- After updating document OCR languages
- After adding, editing, or deleting Document FAQs
- After changing similarity thresholds or AI retrieval settings
- When improving AI answer accuracy or resolving outdated responses

2.1.2.2.3.4.2 Re-Indexing Confirmation

When the re-index action is triggered, the system displays a **confirmation dialog** to prevent accidental execution. The confirmation message informs the administrator that:

- All documents and FAQs in the selected library will be re-indexed
- The process may take time depending on library size and content volume

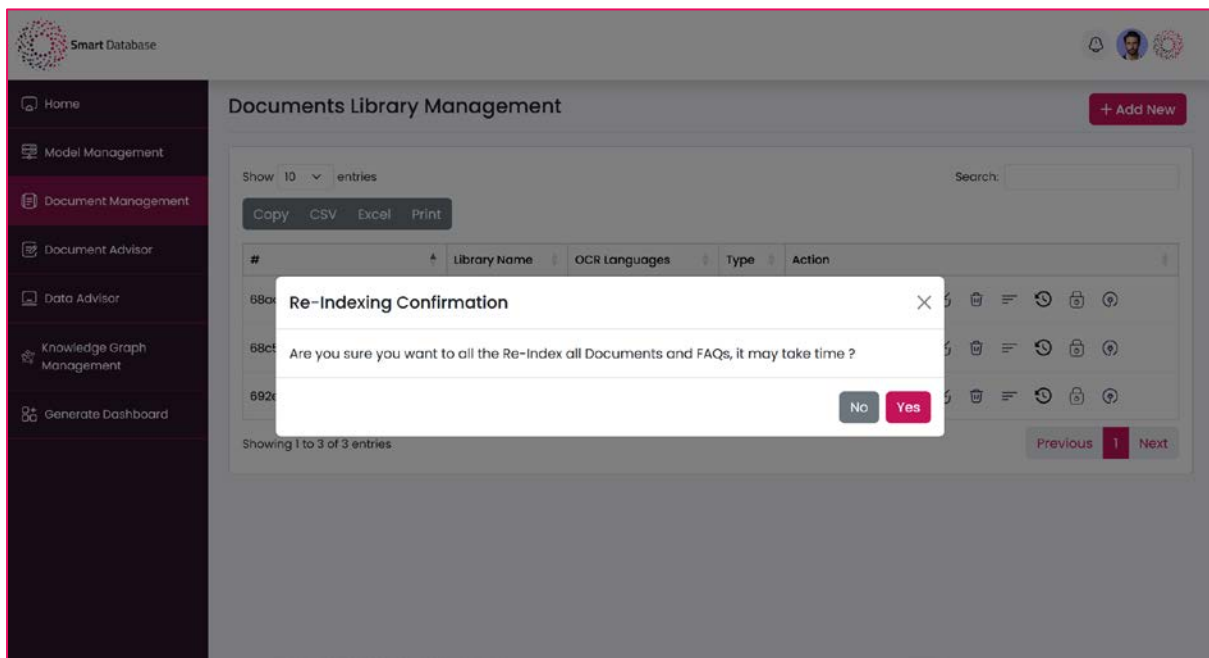


Figure 32: Re-indexing Confirmation Dialog

The administrator can then choose:

- **Yes** – to proceed with re-indexing
- **No** – to cancel the operation



2.1.2.2.3.4.3 Impact of Re-Indexing

Once confirmed, the system will:

- Reprocess all library documents for AI ingestion
- Rebuild semantic embeddings and search indexes
- Apply updated OCR, language, and FAQ configurations
- Ensure future AI responses are generated using the latest indexed content

Note: Re-indexing does not remove documents or FAQs; it only refreshes how the AI understands and retrieves information.

Best Practice

It is recommended to perform re-indexing during low-usage periods for large libraries, as the process may temporarily impact AI response performance while reprocessing is in progress.

2.1.2.2.3.5 API Details for integration

As shown in **Figure 33**, SmartDB (Smart Database) provides out-of-the-box RESTful APIs that enable integration of AI-powered document libraries with any external user interface, application, or digital channel—without requiring the use of the SmartDB (Smart Database) built-in UI.

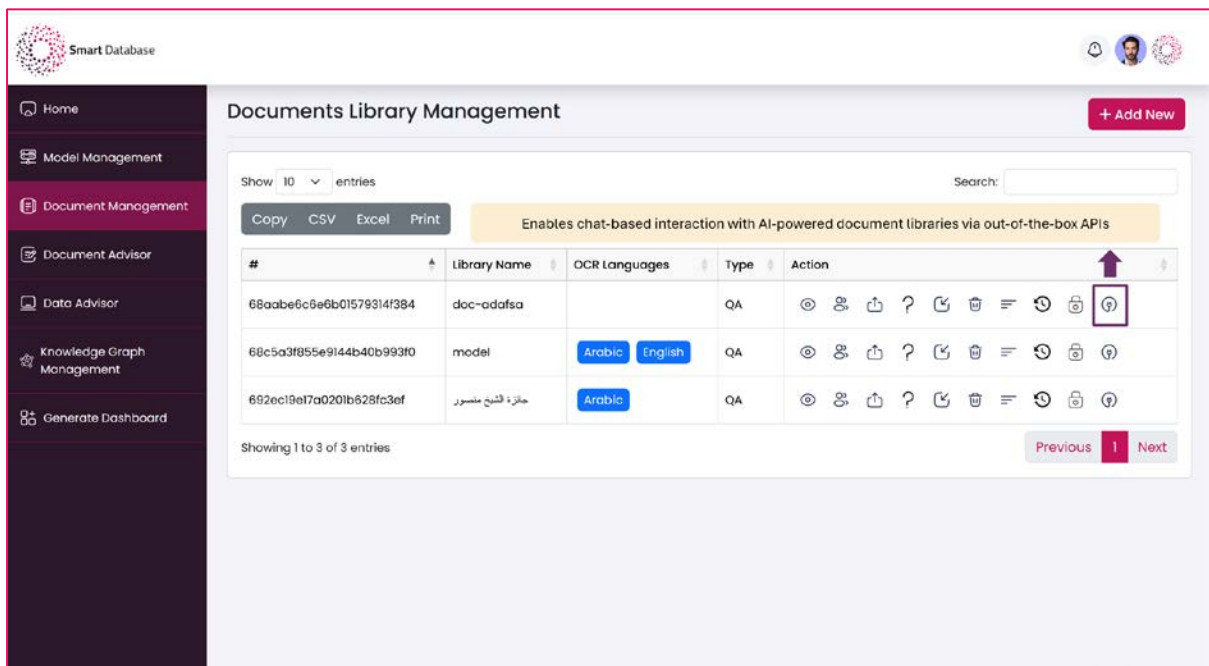


Figure 33: Accessing the out-of-the-box RESTful APIs



Through the **Documents Library Management** screen, administrators can access the API configuration for each document library by clicking the **API icon** in the Actions column. This opens the **API Detail** dialog (illustrated in **Figure 34**), which exposes all required endpoints and payload structures for interacting with the Document Advisor programmatically.

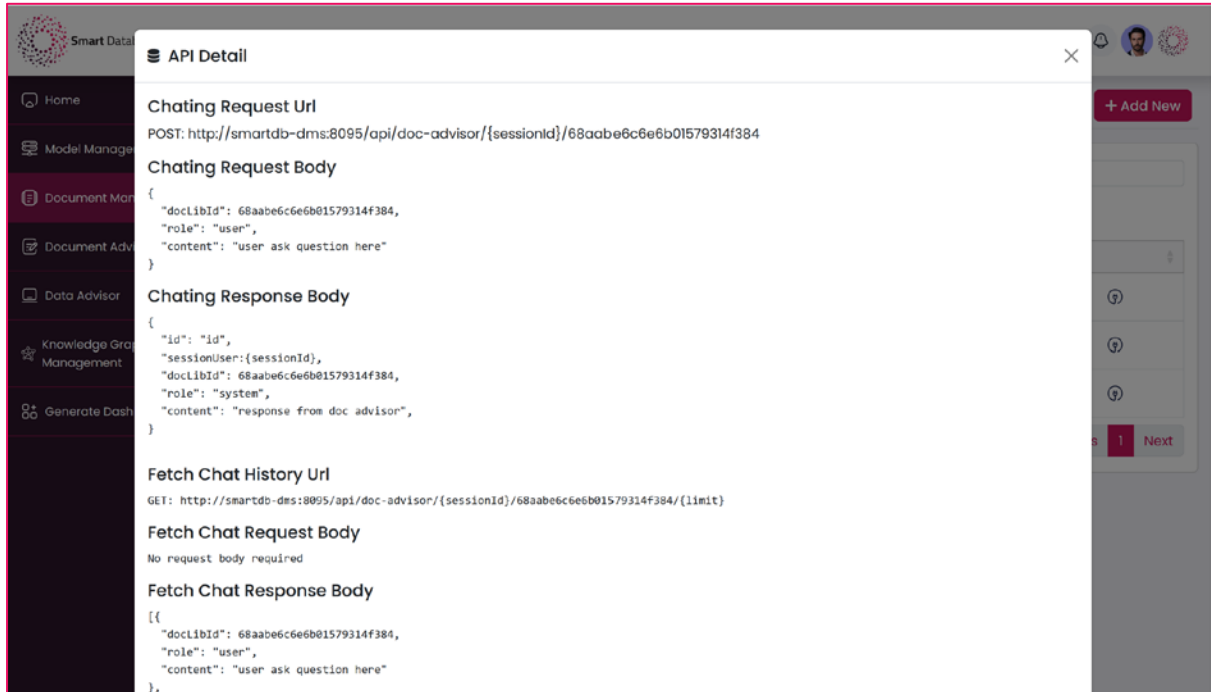


Figure 34: Out-of-the-box RESTful APIs Details

2.1.2.2.3.5.1 API Structure

As illustrated in **Figure 34**, the API structure includes:

1. Chat Request URL (POST)

Used to submit user questions to the document library.

The request URL includes:

- sessionId for conversation tracking
- docLibId to identify the target document library

2. Chat Request Body

Contains:

- docLibId: Target document library ID
- role: Request origin (e.g., user)
- content: User's question

3. Chat Response Body

Returns:

- System-generated AI response



Smart Database

- Session and library identifiers
 - Role metadata (system / user)
4. **Fetch Chat History URL (GET)**
Enables retrieval of previous messages for a session, supporting pagination via a limit parameter.
 5. **Fetch Chat Response Body**
Returns an ordered list of conversation messages, suitable for rendering in any external UI.

2.1.2.2.3.5.2 API Integration Benefits

By exposing these APIs, Smart Database allows organizations to:

- Embed document-aware AI chat into portals, mobile apps, kiosks, or enterprise systems
- Maintain full control over UI/UX design
- Leverage SmartDB's AI, OCR, FAQ governance, and retrieval logic without duplicating backend intelligence
- Ensure consistency between SmartDB UI interactions and external application behavior

In summary, SmartDB acts as the AI intelligence layer, while external systems remain free to implement their own presentation and user interaction layers—fully decoupled and API-driven.



2.2 AI-Powered Document Management

The **Document Advisory** module enables authorized users to interact with document libraries through an AI-powered conversational interface. It allows users to ask natural-language questions, receive accurate answers grounded in approved documents, view document references, and optionally use voice-based interaction.

This module is designed to provide secure, controlled, and explainable access to organizational knowledge.

2.2.1 Accessing the Document Advisory Module

As shown in **Figure 35**, the Document Advisory module can be accessed from the main navigation menu and operates only on document libraries for which the user has been explicitly granted access

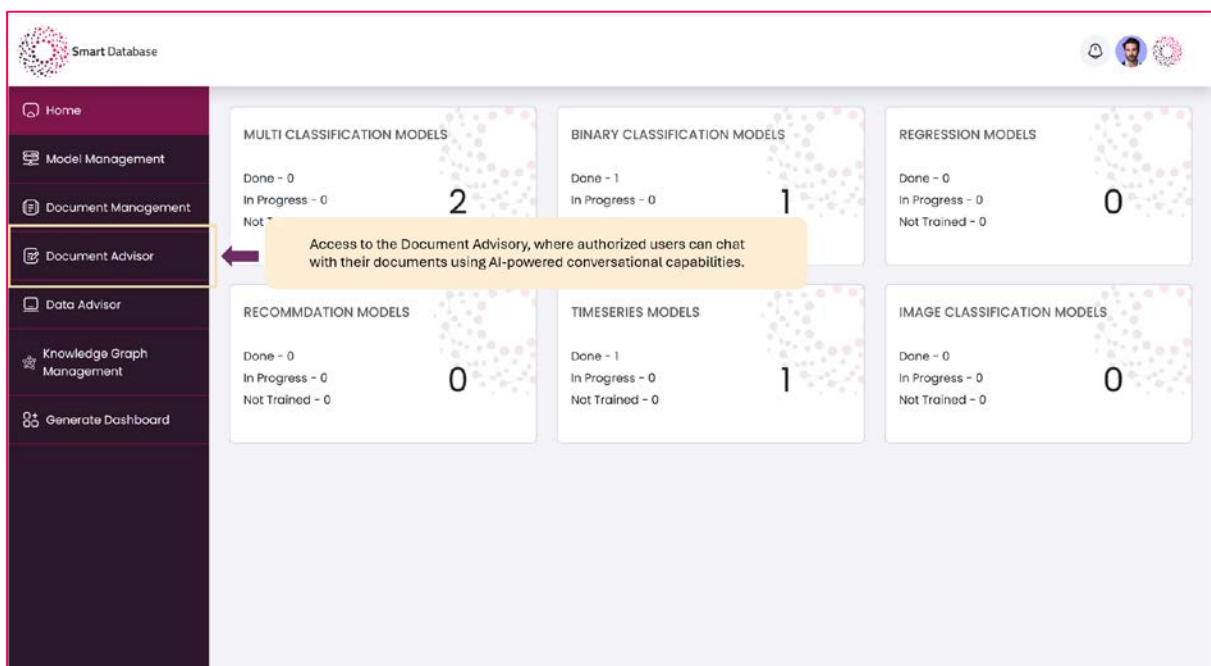


Figure 35: Accessing Document Advisory for Chatting

To access the Document Advisory module:

1. Navigate to Document Advisor from the left-hand navigation menu, as illustrated in **Figure35**.
2. The system opens the Document Advisory interface, displaying the document selection and chat workspace.
3. Only document libraries assigned to the logged-in user will be visible and selectable.

Note: This ensures role-based access control and prevents unauthorized document exposure.



2.2.2 Chat with Your Document (Document Advisory)

Figure36 illustrates the **Document Advisor** interface in SmartDB, which enables authorized users to interact with document libraries using AI-powered conversational capabilities. This module allows users to ask questions in natural language and receive contextual answers generated directly from the content of the selected document library.

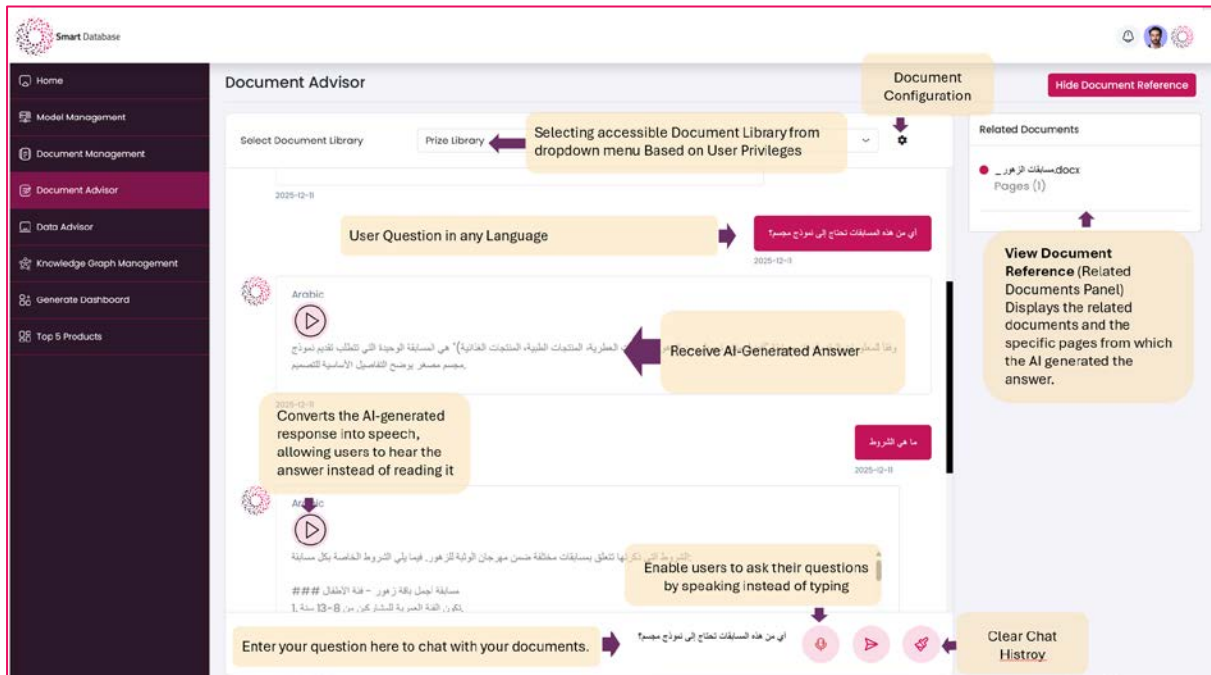


Figure 36: Document Advisory view

2.2.2.1 Selecting a Document Library

At the top of the page, users must select a **Document Library** from the *Select Document Library* dropdown (see **Figure 36**).

- The list displays only libraries for which the logged-in user has access rights.
- Once selected, the system loads the associated documents and enables conversational interaction with their content.

2.2.2.2 Conversational Interaction

The main panel displays the **chat conversation area**, where:

- User questions appear as highlighted message bubbles.
- AI-generated responses are displayed below each question, enriched with contextual understanding derived from the documents.



Users can type their questions in the **user writing area** at the bottom of the page and submit them using the **Submit Chat** button. The conversation history is preserved during the session for continuity and reference.

2.2.2.3 Speech Capabilities

The Document Advisor supports voice interaction:

- **Speech-to-Text** allows users to dictate their questions instead of typing.
- **Text-to-Speech** (CPU-based) enables playback of AI responses by clicking the audio icon on each response, as shown in **Figure 36**.

2.2.2.4 Related Documents Panel

On the right side of the screen, the **Related Documents** panel displays the source documents used to generate the AI response (see **Figure 36**).

- Clicking on a document highlights its relevance and provides transparency into the answer generation process.
- Users can show or hide this panel using the **Hide Document Reference** button.

2.2.2.5 Chat History Option

Users can clear the current conversation using the **Clear Chat History** option. This resets the session while keeping the document library selection intact.



2.2.2.6 Configuration Settings

From the **Document Advisor** page, users can open this panel by clicking the **settings (gear) icon** next to the selected document library, as illustrated in **Figure 37**. The settings panel appears on the right side of the screen.

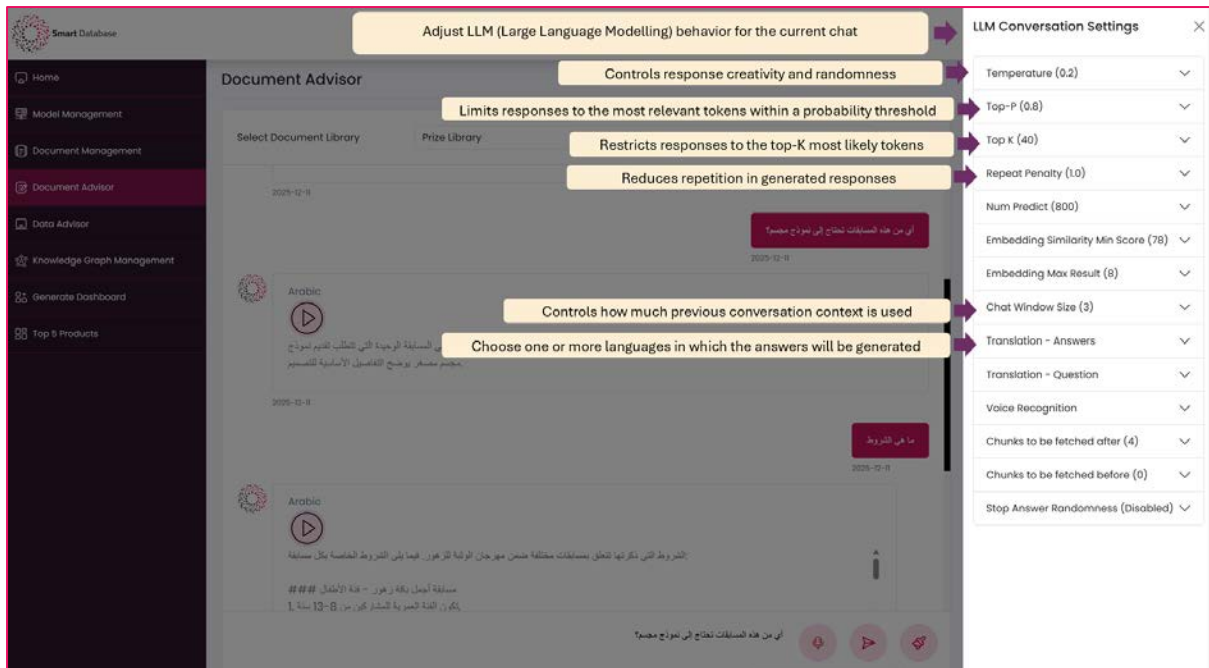


Figure37: LLM Conversation Setting

These settings include:

1. Temperature

Controls the creativity of AI responses.

- Lower values (e.g., 0.2) produce more deterministic and factual answers.
- Higher values allow more creative or varied responses.

2. Top-P

Defines nucleus sampling, limiting responses to the most probable tokens that reach the cumulative probability threshold.

- Helps balance creativity and relevance.

3. Top-K

Restricts the AI to selecting from the top K most likely tokens.

- Higher values increase diversity; lower values increase precision.

4. Repeat Penalty

Reduces repetition in AI-generated responses.

- Higher values discourage repeating phrases or sentences.

5. Num Predict



Specifies the maximum number of tokens the AI can generate in a single response.

- Useful for controlling response length.

6. Embedding & Retrieval Controls

a. Embedding Similarity Minimum Score

Defines the minimum similarity threshold required for document chunks to be considered relevant.

- Higher values improve precision but may reduce recall.

b. Embedding Max Result

Limits the number of document chunks retrieved and used to generate an answer.

c. Chat Window Size

Controls how many previous messages are considered as context when generating a response.

7. Language & Voice Configuration

a. Translation – Answers

Automatically translates AI responses into the user's preferred language.

b. Translation – Questions

Translates user questions before processing them by the AI model.

c. Voice Recognition

Enables or disables speech-to-text functionality.

- When enabled, voice input is handled at the browser level unless overridden by system-wide settings.

8. Context Chunk Controls

a. Chunks to Be Fetched After

Defines how many document chunks are retrieved after the matched section for additional context.

b. Chunks to Be Fetched Before

Defines how many document chunks are retrieved before the matched section.

9. Answer Stability

a. Stop Answer Randomness

When enabled, the AI generates more stable and consistent responses by minimizing randomness.

Note: Changes made here apply only to the current session and do **not** override the default configuration defined during document library setup.



2.2.2.7 Summery

The LLM Conversation Settings panel provides fine-grained control over AI behavior during document-based conversations. It is designed for advanced users who need to balance accuracy, explainability, response length, and interaction style without impacting the global document library configuration.



Smart Database

THANK YOU.

We are proud to be your trusted partner on the AI journey

Participant Guide - AI-Powered Document Processing

Access our Knowledge Center

<https://knoweldgecenter.smartdb.ai>

VERSION 10.0 - SERVICE RELEASE 7

December 2024

This document is owned by Smart Database. Reproduction in any form, whether in whole or part, is strictly prohibited. For educational use only. Information is subject to change without prior notice.