A blue background with white text

AI-generated content may be incorrect.

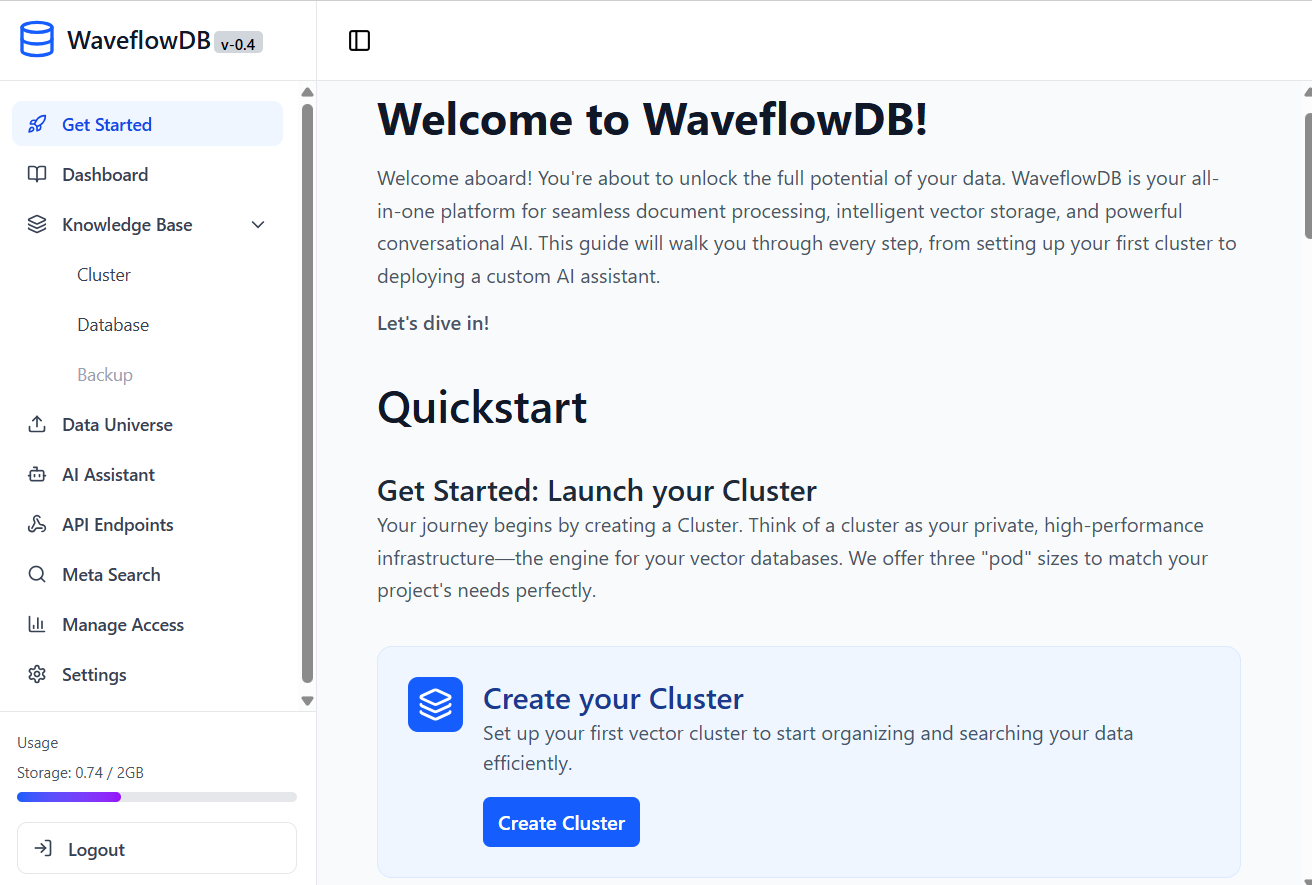
# **What is WaveflowDB?**

WaveflowDB is a **deep search full corpus vector database** which is specifically created for **building AI agents** and can be used for deep enterprise search, powerful conversational AI, seamless document processing, intelligent vector storage, etc. This comprehensive guide will walk you through every feature, configuration option, and best practice to help you maximize the potential of your data for building AI agents using WaveflowDB.



# **Understanding WaveflowDB’s Features**

Before diving into cluster creation, it's important to understand how WaveflowDB is structured:



cc

* **Get Started**: An interactive guide that directs new users through the initial setup process.
* **Dashboard:** The Dashboard acts as your mission control, empowering you to monitor your resources, track usage, and manage your account efficiently from one convenient location.
* **Cluster:** Your dedicated infrastructure layer that provides compute resources.
* **Database:** A logically isolated data container within a cluster for organizing different projects.
* **Data Universe:** Your central hub for uploading, processing, and managing the source files (PDFs, DOCX, etc.) that form the foundation of your knowledge base.
* **AI Assistant:** RAG-based conversational AI interfaces built on your data.
* **Meta Search:** Your data exploration tool. Use it to search and verify the files within your Data Universe before building an assistant.
* **Manage Access:** The team collaboration centre for inviting users and managing permissions across your workspace.
* **API Endpoint:** An authentication credential and an access control mechanism, allowing your systems to securely interact with clusters, databases, and AI assistant endpoints.

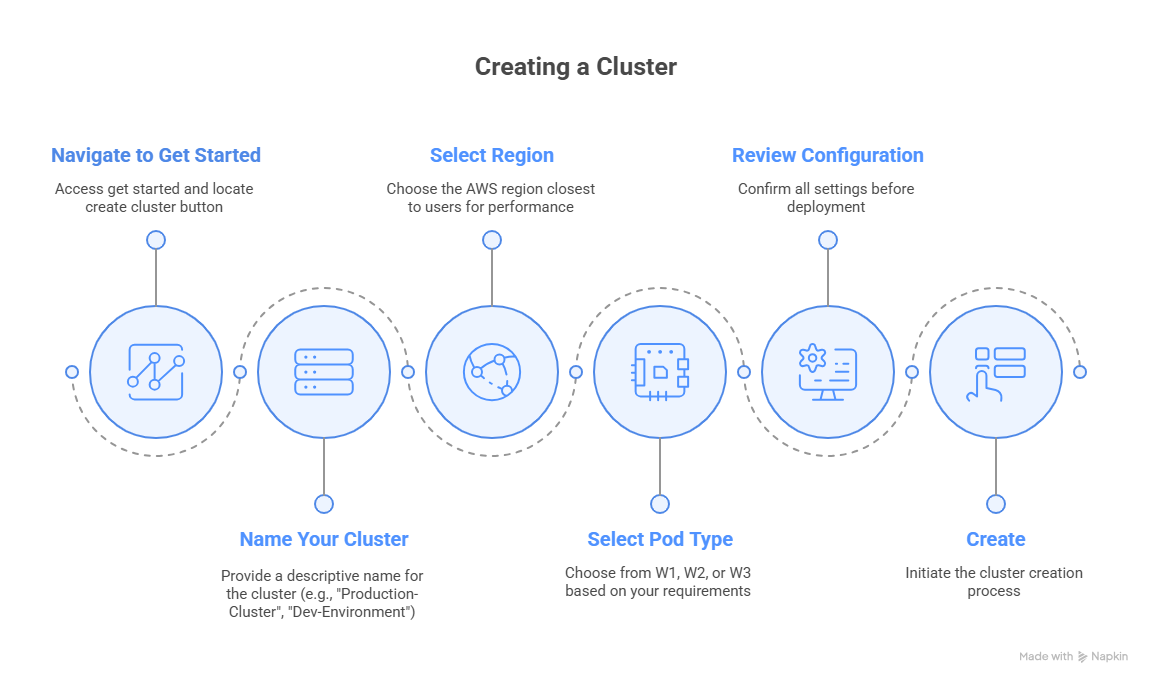
# **Get Started**

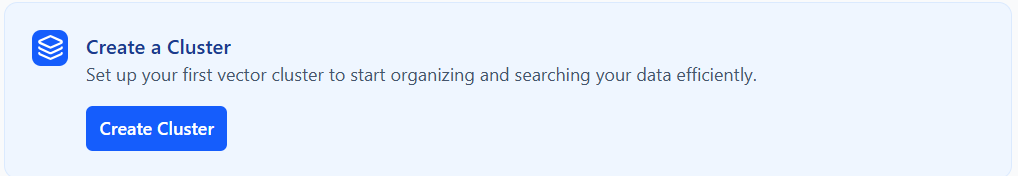
The "Get Started" page is the primary landing page for all users upon their first login to WaveflowDB. Its core purpose is to serve as an interactive, guided onboarding tool that directs new users through the essential steps required to achieve their first successful outcome: **creating a cluster**.

A diagram of a wave

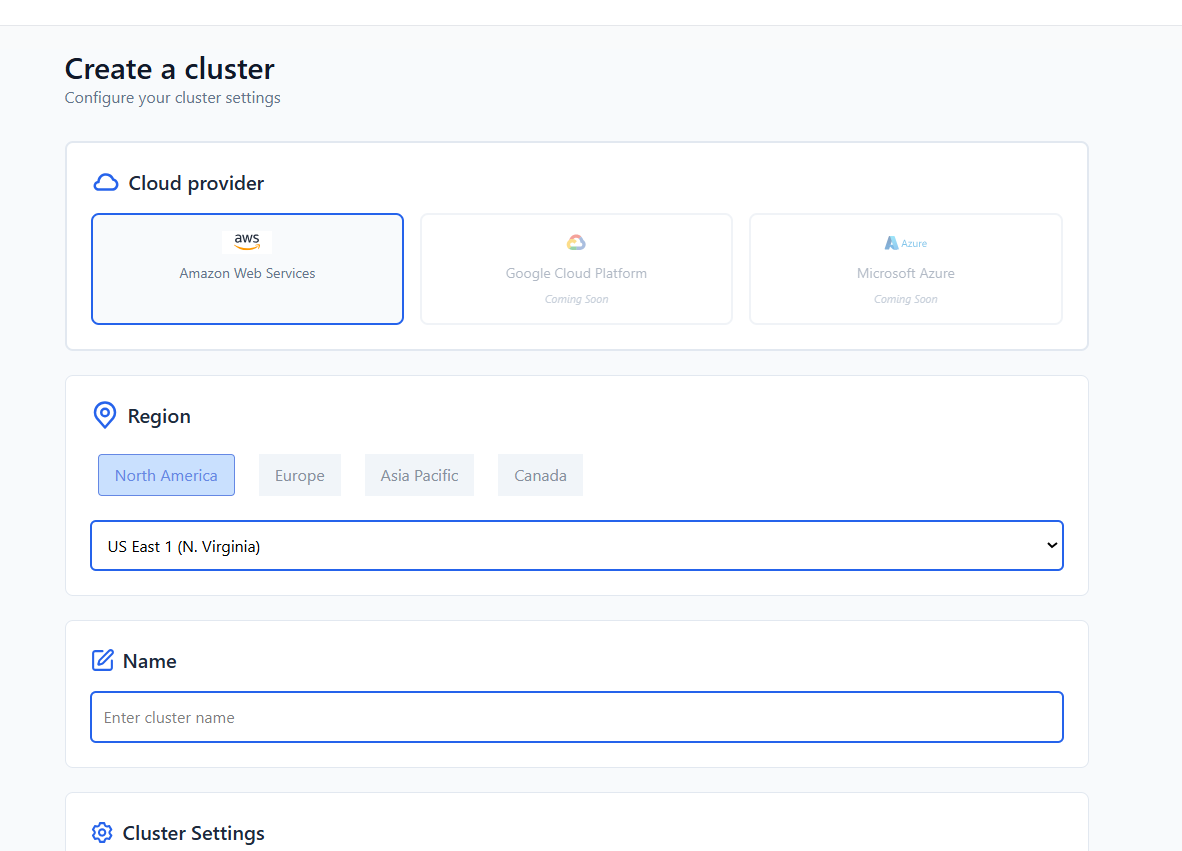
AI-generated content may be incorrect.This page is designed specifically to solve the "empty state" problem, preventing user confusion and providing a clear, actionable path from account creation to platform value. It acts as a temporary, task-oriented homepage until the user has provisioned their initial set of resources.

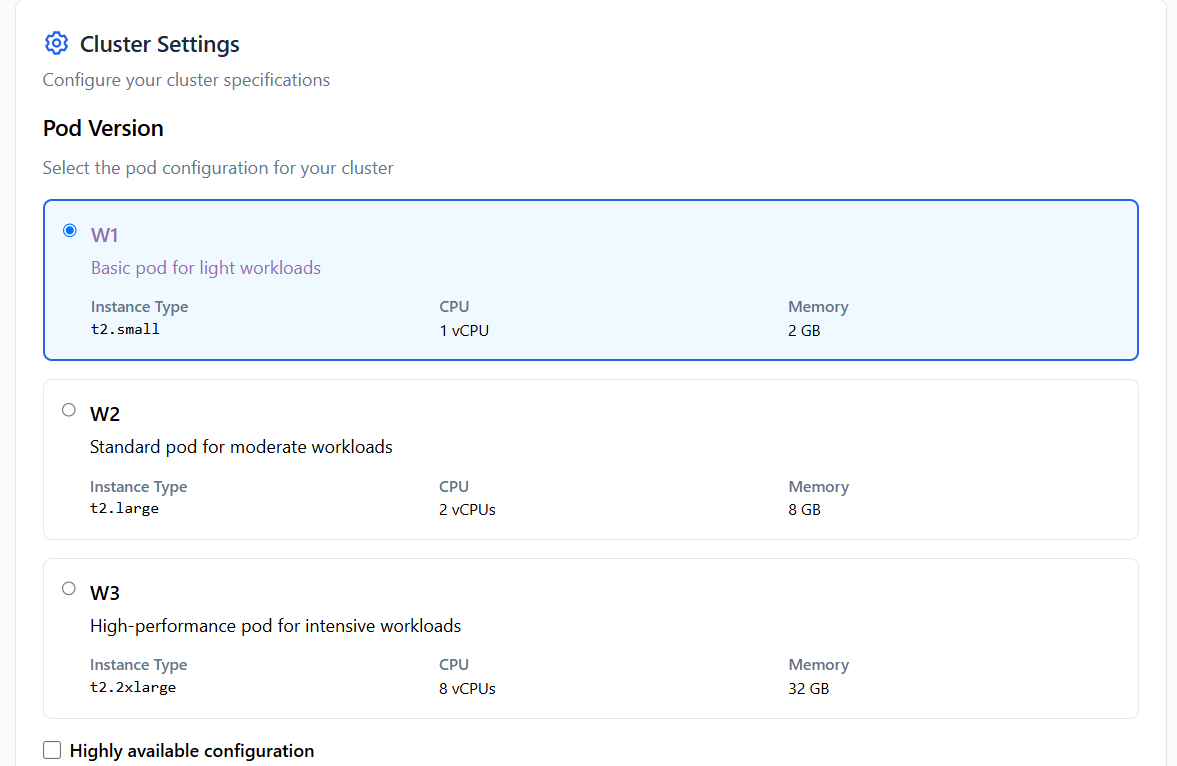
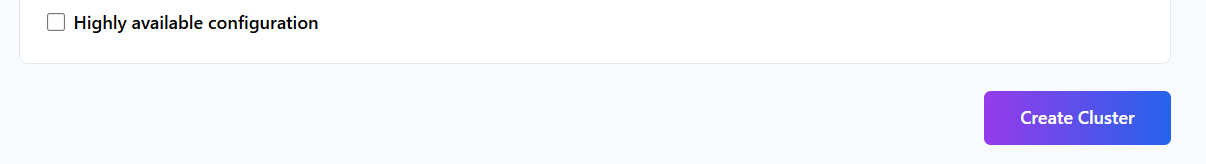
## **3.1 Cluster Creation Process**



1. **Navigate to Dashboard**: From the main dashboard, locate the **Create Cluster** button

1. **Select Region**: Choose the AWS region closest to your users for optimal performance. Currently only one region is provided.
2. **Name Your Cluster**: Provide a memorable, descriptive name (e.g., "Production-Cluster", "Dev-Environment"). Make sure the name does not contain any special characters as well.



1. **Select Pod Type**: Choose from W1, W2, or W3 based on your requirements (see table above)
2. **Review Configuration**: Confirm your settings before deployment
3. **Create**: Click create and wait for provisioning.

# **Knowledge Base**

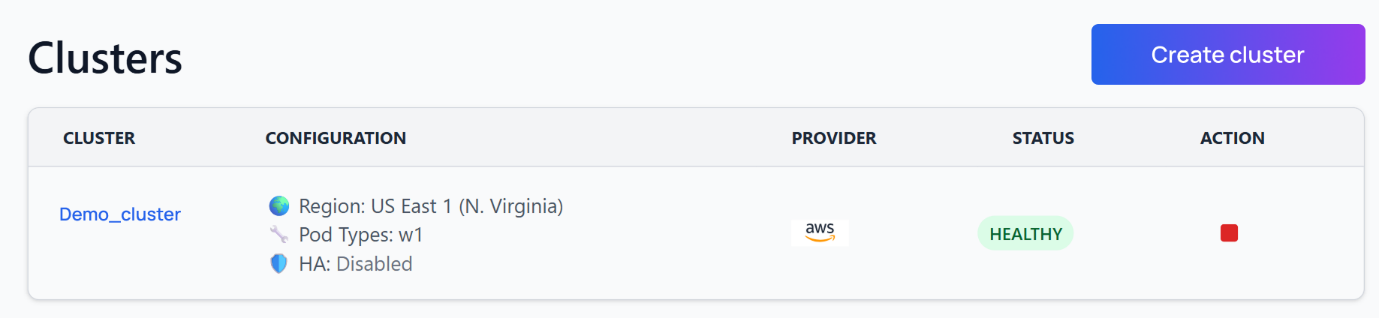
## **4.1 Cluster**

A **Cluster** is your private, high-performance infrastructure that serves as the foundation for all your vector databases. Think of it as your dedicated compute environment that processes, stores, and serves your vectorized data.

### 4.1.1 Getting to Cluster After Creation

Created clusters are shown in the Cluster page under Knowledge Base along with features like configuration, provider, status and action.

**‘Action’** has pause and resume functionality allowing temporary break to the cluster when not in use.

****

Use the **create cluster** button to create a new cluster.

### 4.1.2 Cluster Pod Types and Specifications

Choose the right pod based on your workload requirements:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Pod Type | Instance Type | CPU | Memory | Recommended Use Cases |
| W1 | t2.small | 1 vCPU | 2 GB | • Development and testing  • Small document collections (<1000 files)  • Proof of concepts  • Learning and experimentation |
| W2 | t2.large | 2 vCPUs | 8 GB | • Production applications  • Medium document collections (1000-10,000 files)  • Multiple concurrent users (10-50)  • Regular business operations |
| W3 | t2.2xlarge | 8 vCPUs | 32 GB | • Enterprise applications  • Large document collections (10,000+ files)  • High concurrent usage (50+ users)  • Mission-critical deployments  • Real-time applications |

### 4.1.3 Cloud Provider Support

|  |  |  |
| --- | --- | --- |
| Provider | Status | Notes |
| AWS | ✅ Available | Full support with all pod types |
| Google Cloud (GCP) | 🔄 Coming Soon |  |
| Microsoft Azure | 🔄 Coming Soon |  |

### 4.1.4 Cluster Management for Best Practices

* **Naming Convention**: Use clear, descriptive names that indicate purpose and environment
* **Resource Planning**: Start with W1 for testing, scale up as needed
* **Region Selection**: Choose regions close to your primary user base

## **4.2 Database**

A **Database** is an isolated container within your cluster that organizes your data by project, department, or use case. Multiple databases can exist within a single cluster, each with its own configuration and access controls.

### 4.2.1 Embedding Models

|  |  |  |  |
| --- | --- | --- | --- |
| Model | Status | Dimensions | Best For |
| all-MiniLM-L6-v2 | ✅ Available | 384 | • General purpose text  • Fast processing  • Balanced performance |
| sentence-transformers/all-mpnet-base-v2 | 🔄 Coming Soon | 768 | • High-quality embeddings  • Complex documents |
| text-embedding-ada-002 | 🔄 Coming Soon | 1536 | • OpenAI compatibility  • Premium quality |

### 4.2.2 Vector Configuration

|  |  |  |  |
| --- | --- | --- | --- |
| Setting | Current Options | Coming Soon | Description |
| Vector Type | Dense | Sparse, Hybrid | Type of vector representation |
| Distance Metric | Cosine | Dot Product, Euclidean | How similarity is calculated |
| Dimensions | 384 | 512, 768, 1024, 2048 | Vector size (higher = more precise) |

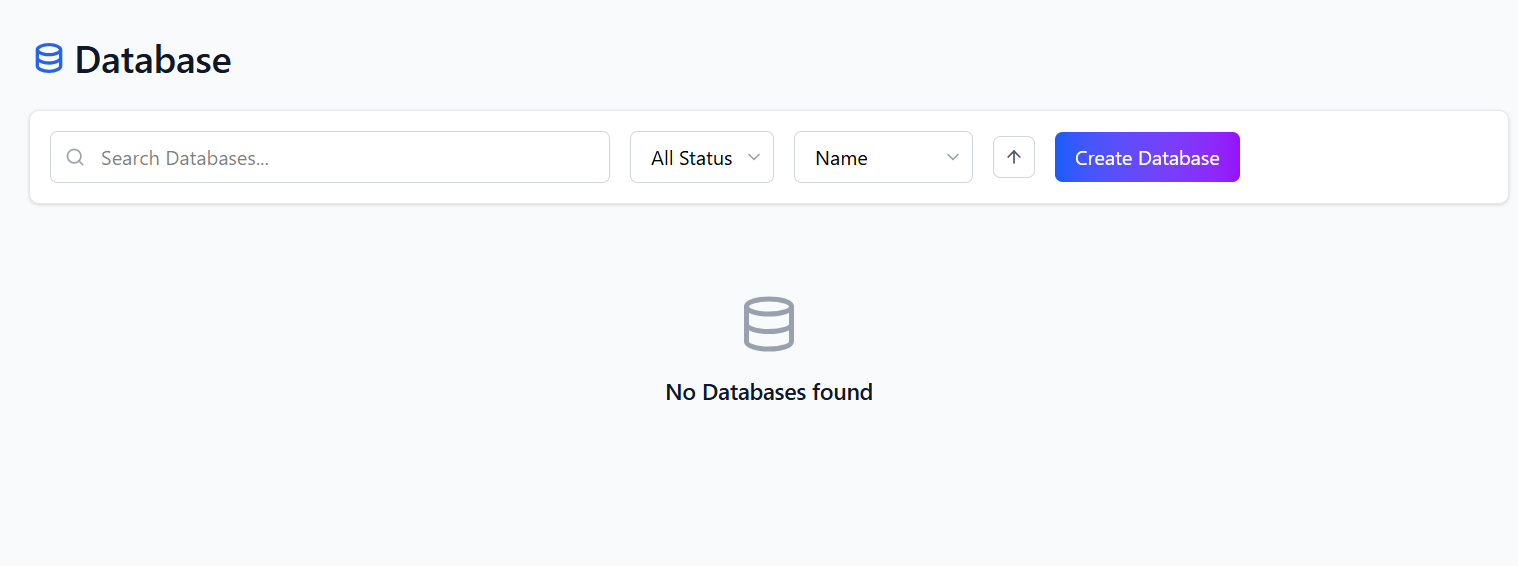
### 4.2.3 Database Creation Process

**Note:** A database can only be created in an existing cluster.

A diagram of a database creation process

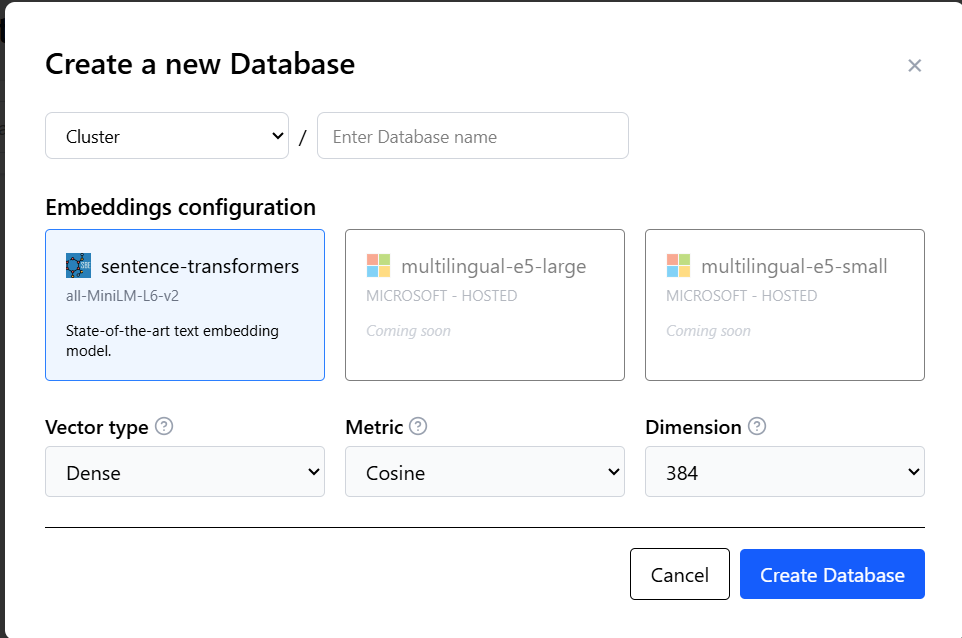
AI-generated content may be incorrect.

1. **Click Create Database**: Navigate to the database creation interface

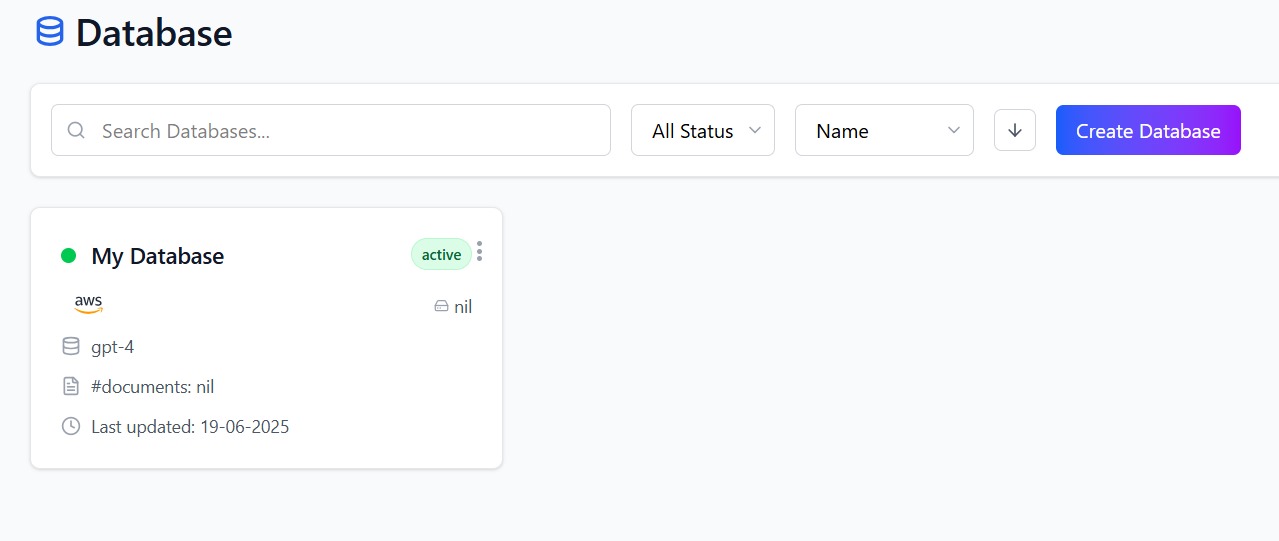


1. **Select Cluster**: Choose the cluster where you want to create the database
2. **Configure Settings**:

* **Database Name**: Choose a unique, descriptive name
* **Embedding Model**: Select from available options (current v0.4 has all-MiniLM-L6-v2)
* **Vector Type**: Choose Dense (only current option, spare vector type is part of road map)
* **Distance Metric**: Select Cosine (only current option, Euclidean distance measure is part of roadmap)
* **Dimensions**: Set to 384 (only current option, other higher dimensions are a part of the roadmap)



1. **Review and Create**: Confirm your configuration and create the database



### 4.2.4 Database Organization Strategies

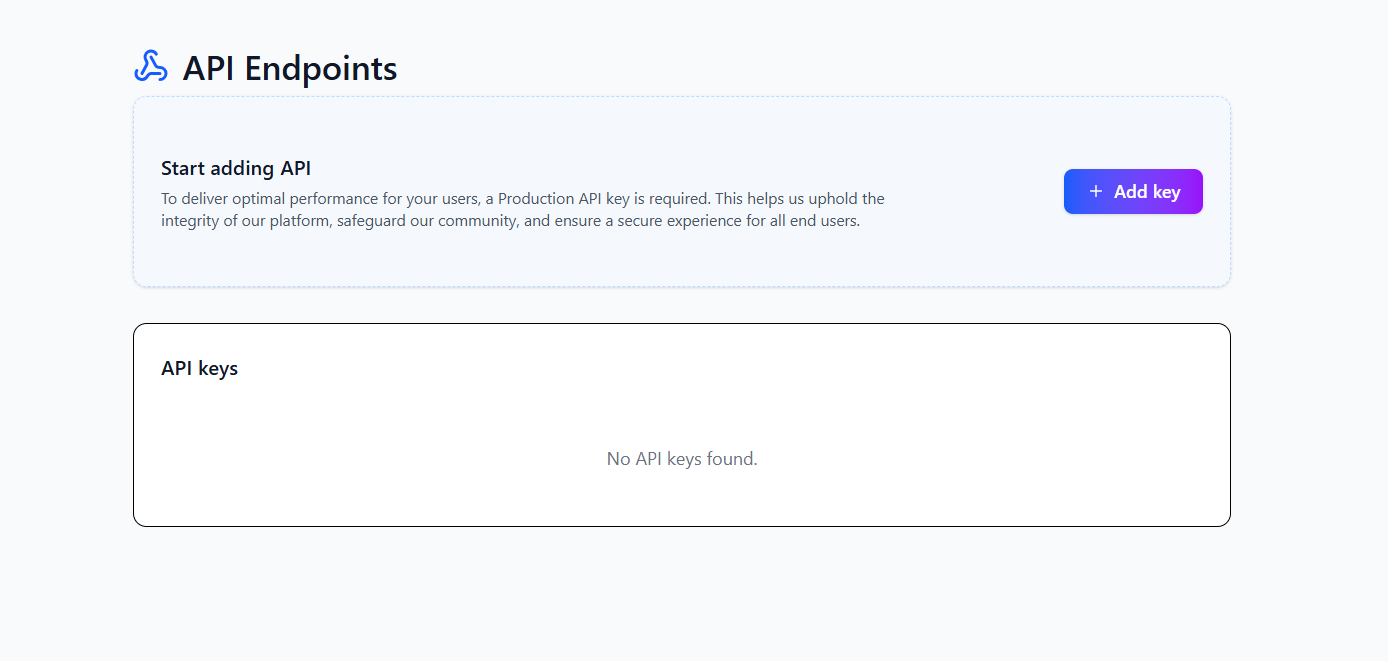
|  |  |  |
| --- | --- | --- |
| Strategy | Functionality | Benefits |
| By Name | Sorts the list alphabetically from A-Z | Quick navigation |
| By Last Updated | Sorts the list showing most recently modified resources. | Focus on Active Work |
| By Document | Sorts the list based on the number of documents uploaded | Resource Management |
| By Status | Filters the list to show databases with certain statuses like ‘active’ only. | Troubleshooting |

# **API Endpoints**

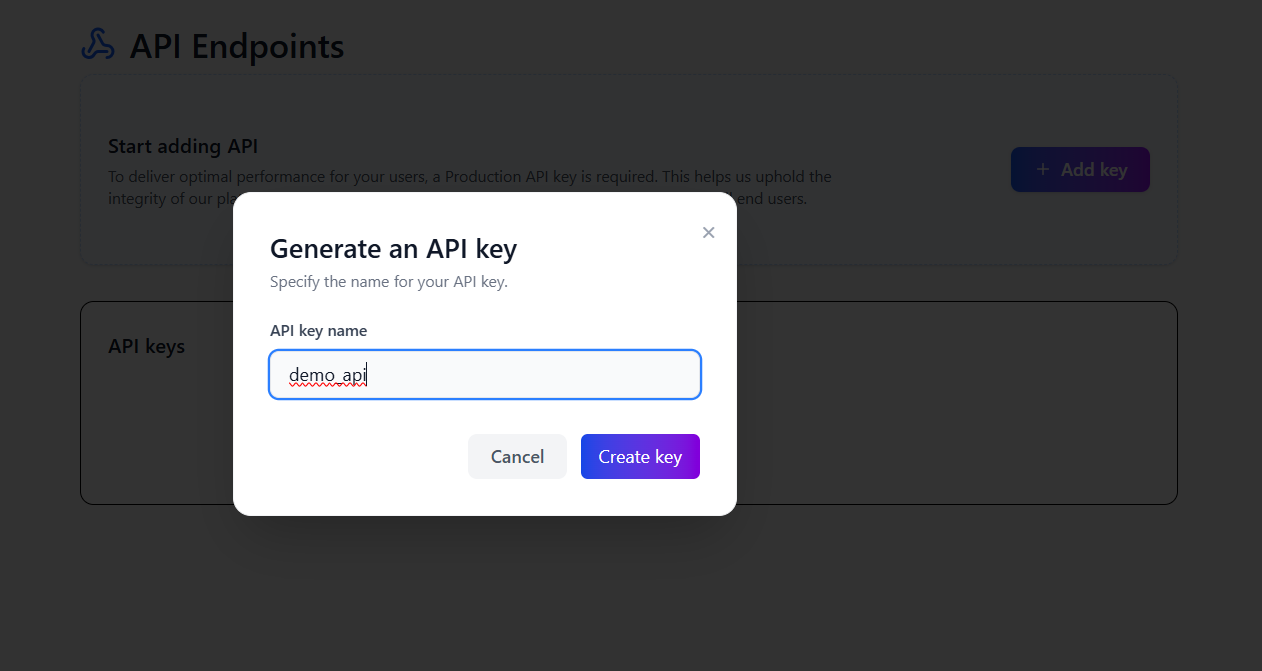
An API key in WaveflowDB is a unique, secret token issued to your account that governs access to your resources. It functions as both a critical **authentication credential** and a powerful **access control mechanism**, enabling your applications to securely interact with your clusters, databases, and AI Assistants.

## **5.1 Creating API Key**

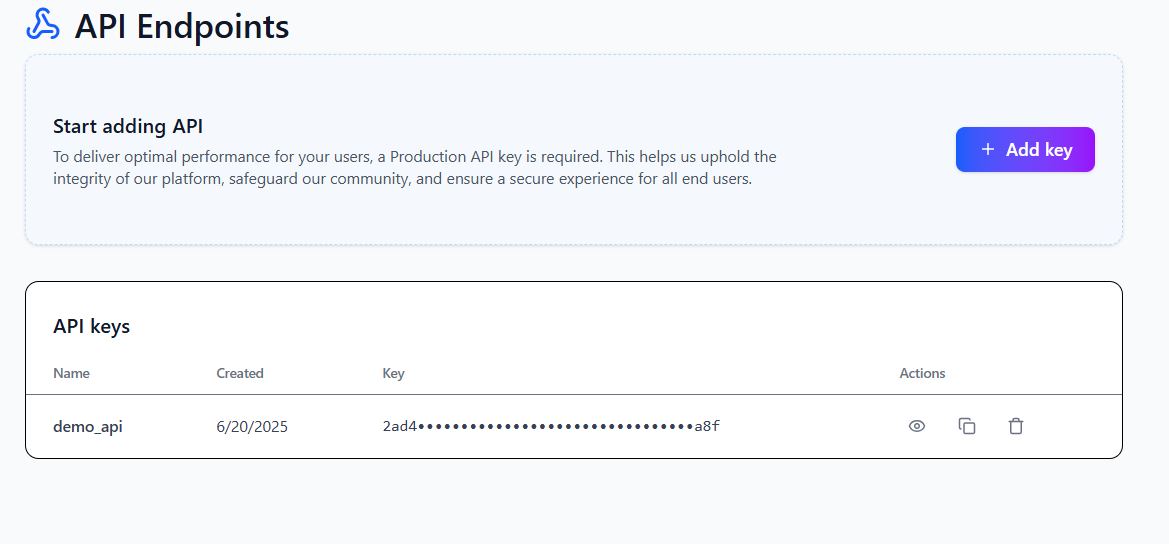
* 1. Navigate to the **API Endpoints** page
  2. Click on **add API** key



* 1. Give an appropriate name and press on **create key**.



* 1. An **API key** is created in the API endpoints screen.



## **5.2 API Key Capabilities**

|  |  |
| --- | --- |
| Capability | Functionality |
| Upload | Access API endpoints to upload files |
| Fetch | Access API endpoints to ask a query and get relevant data |
| View | Access API endpoints to view the existing data |

## **5.3 API Endpoint Features (Coming Soon)**

**Resource-Specific Keys**

* You will be able to bind an API key to a specific Database or Cluster. This provides granular control, perfect for multi-tenant applications where you need to guarantee strict data isolation between different projects or customers.

**Activity Monitoring**

* The API key dashboard will display a "Last Used" timestamp for each key, helping you identify and safely remove old or inactive keys from your account.

**Usage Quotas Per Key**

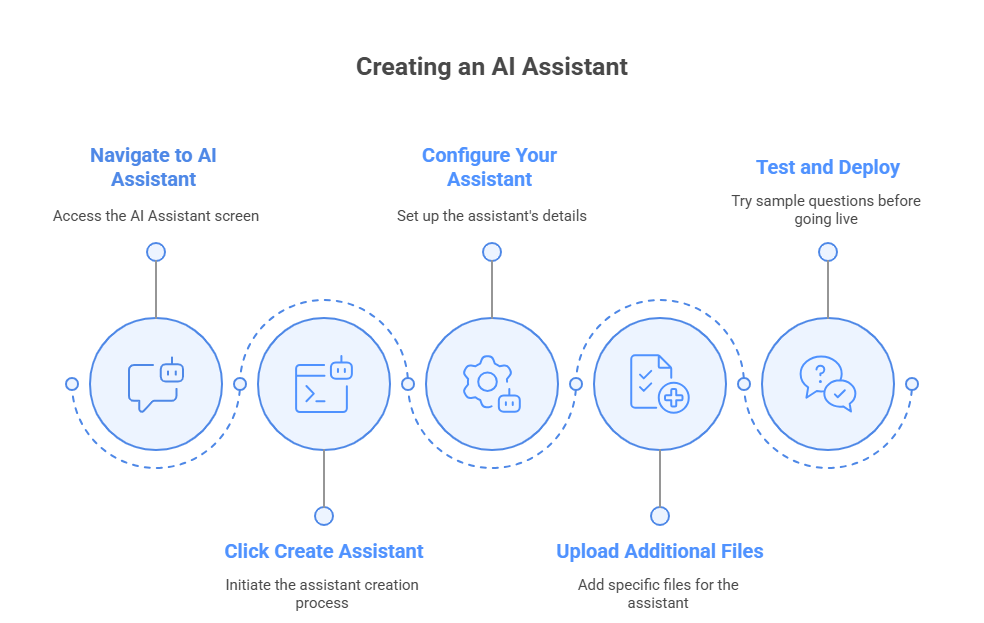
* You will be able to assign specific rate limits or usage quotas (e.g., max queries per day) to individual API keys. This feature will provide fine-grained cost control and help prevent abuse.

# **AI Assistant**

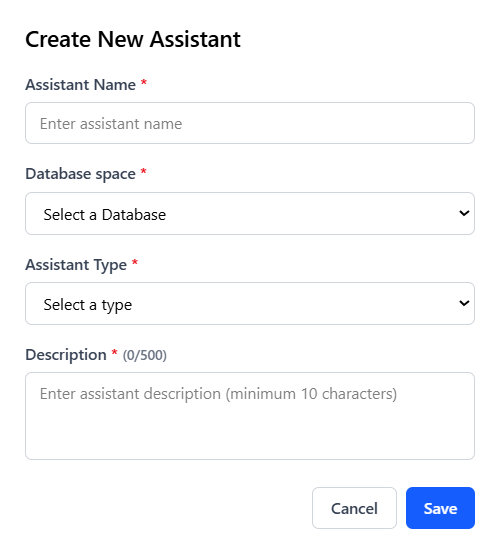
An **AI Assistant** is a RAG-based (Retrieval-Augmented Generation) conversational interface that can intelligently answer questions, provide summaries, and engage in natural language conversations based on the documents you've uploaded to your databases.

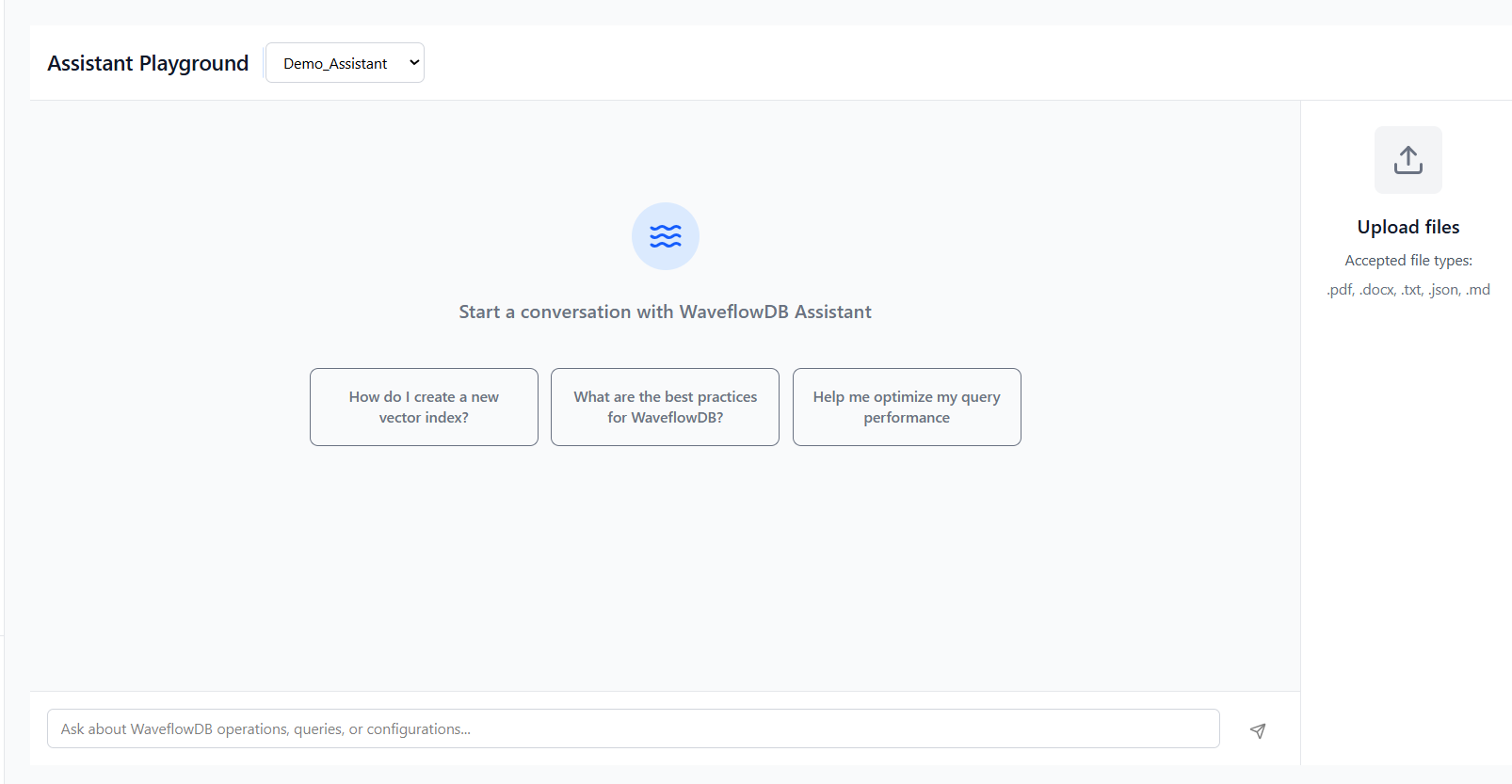
## **6.1 Key Components**

|  |  |  |
| --- | --- | --- |
| Component | Description | Purpose |
| Knowledge Base | Your uploaded documents | Source of truth for answers |
| Retrieval System | Vector search engine | Finds relevant content |
| Language Model | AI conversation engine | Generates human-like responses |

****

## **6.2 Create your AI Assistant**

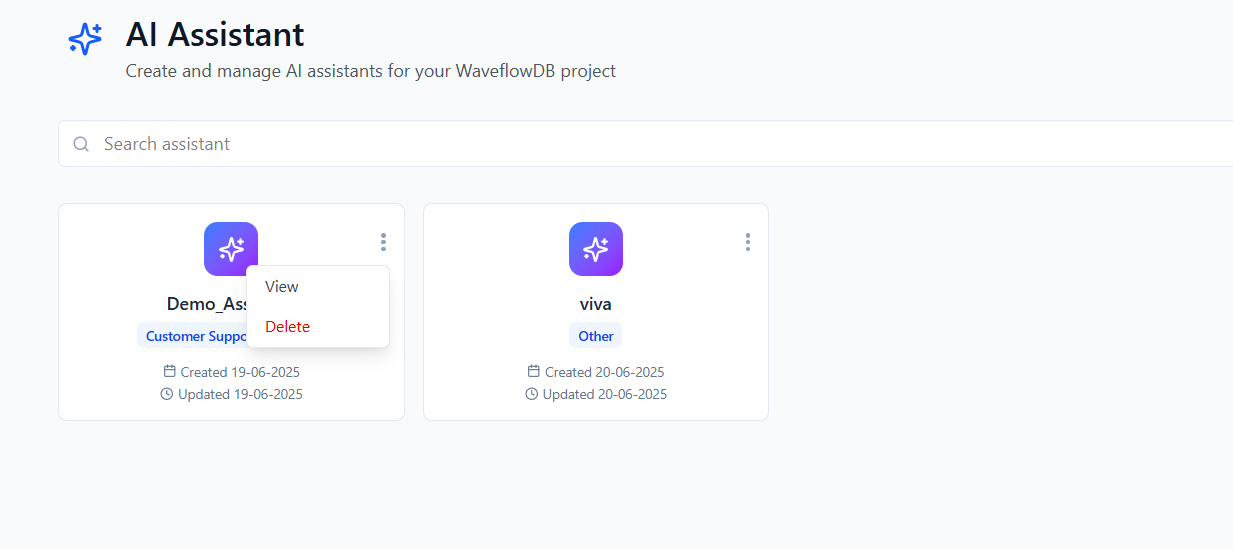
1. **Navigate to AI Assistant**: Go to the AI Assistant screen
2. **Click Create Assistant**: Start the creation process
3. **Configure Your Assistant**:
   * **Name**: Give it a memorable, descriptive name
   * **Assistant Type**: Define its primary function (e.g., "Customer Support Assistant, Sales Enablement Assistant")
   * **Description**: Explain what it does
   * **Database Selection**: Choose which database it can access



1. **Upload Additional Files** (Optional): Add specific files for this assistant
2. **Test and Deploy**: Try sample questions before going live
3. **AI Assistant page** shows the list of AI agents created.



1. **Click on view** to go to the playground to interact with your agent.



## **6.3 Assistant Configuration Examples**

|  |  |  |  |
| --- | --- | --- | --- |
| Use Case | Name | Assistant Type | Description |
| Customer Support | SupportBot | Customer Service Assistant | Answers questions about products, policies, and procedures based on support documentation |
| Employee Onboarding | OnboardingGuide | HR Assistant | Helps new employees navigate company policies, procedures, and benefits information |
| API Documentation | DevHelper | Developer Assistant | Provides code examples, API references, and technical guidance from documentation |
| Research Assistant | ResearchAI | Research Analyst | Synthesizes information from research papers and reports to answer complex queries |

## **6.4 Best Practices for AI Assistants**

|  |  |  |
| --- | --- | --- |
| Practice | Benefit | Implementation |
| Clear Role Definition | Better focused responses | Be specific about the assistant's expertise area |
| Comprehensive Knowledge Base | More accurate answers | Include all relevant documents |
| Regular Testing | Improved user experience | Test with common user questions |
| Iterative Improvement | Enhanced performance | Add new documents based on user needs |

# **Data Universe**

The **Data Universe** is your centralized command centre for managing and uploading files. It provides a robust interface for bringing your documents into WaveflowDB, where they're automatically processed, vectorized, and made searchable.

## **7.1 Supported File Formats**

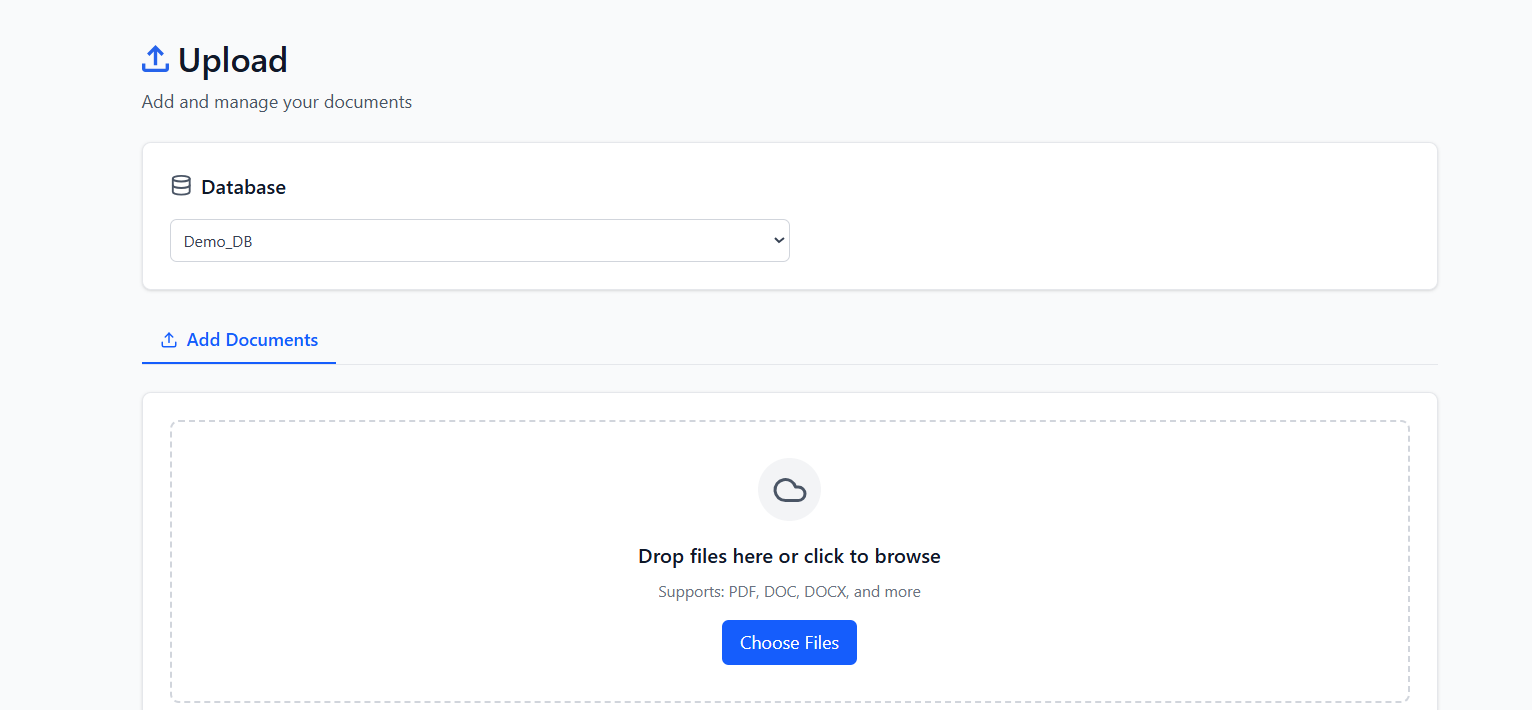
|  |  |
| --- | --- |
| Format | Extension |
| CSV | .csv |
| Python | .py |
| JSON | .json |
| Word Documents | .docx |
| PDF | .pdf |

## **7.2 File Upload Process**

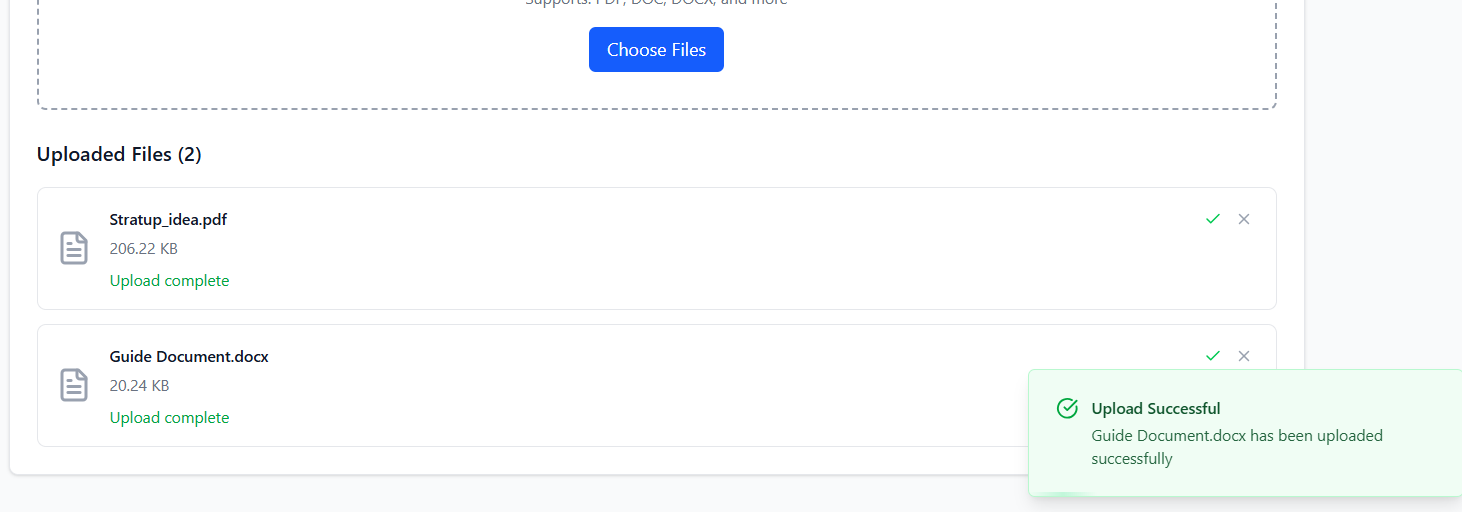
A diagram of data upload process

AI-generated content may be incorrect.

1. **Navigate to Data Universe**: Access the Data Universe tab from the main navigation
2. **Select Target Database**: Use the dropdown to choose which database will store your files

****

1. **Upload Files**:
   * **Drag & Drop**: Simply drag files into the upload zone
   * **Browse**: Click to select files from your computer
   * **Batch Upload**: Select multiple files simultaneously

****

1. **Confirmation**: Receive notification when upload and processing are complete

A diagram of a funnel

AI-generated content may be incorrect.**7.3 File Processing Pipeline**

When you upload files, WaveflowDB automatically:

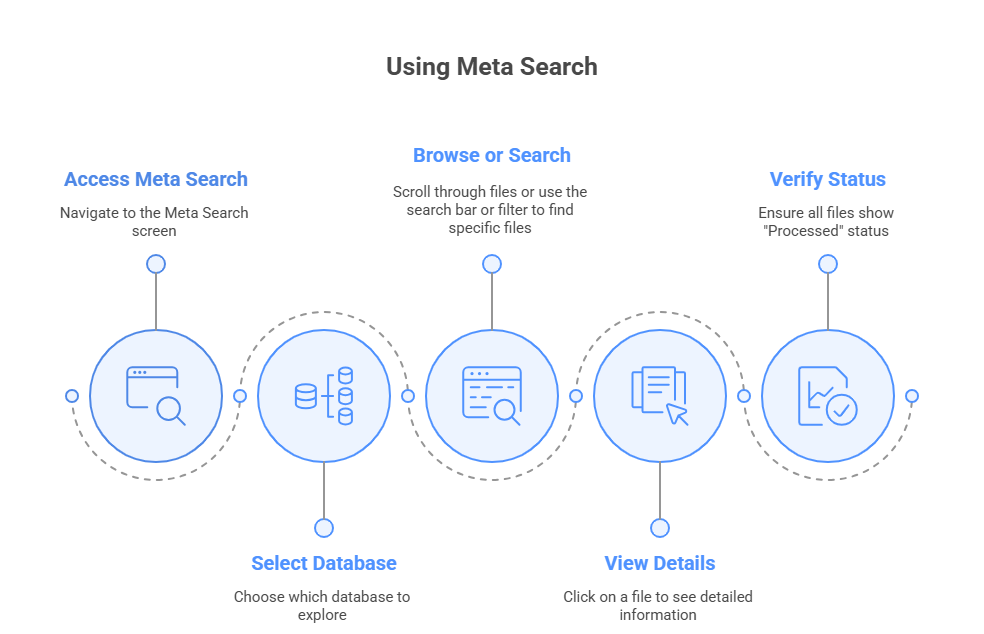
1. **Validates** file format and size
2. **Extracts** text content and metadata
3. **Chunks** large documents into searchable segments
4. **Generates** vector embeddings using your selected model
5. **Indexes** content for fast retrieval
6. **Stores** both original content and vectors

## **7.4 Upload Best Practices**

|  |  |  |
| --- | --- | --- |
| Practice | Why It Matters | Example |
| Organize files before upload | Easier management and search | Group related documents together |
| Use descriptive filenames | Better search results | "Q4-2024-Sales-Report.pdf" vs "document1.pdf" |
| Check file quality | Better extraction results | Ensure PDFs are text-based, not just images |

# **Meta Search**

**Meta Search** is your data exploration tool that allows you to inspect, verify, and manage the files you've uploaded to any database. Think of it as your "file manager" with search capabilities.



## **8.1 Meta Search Features**

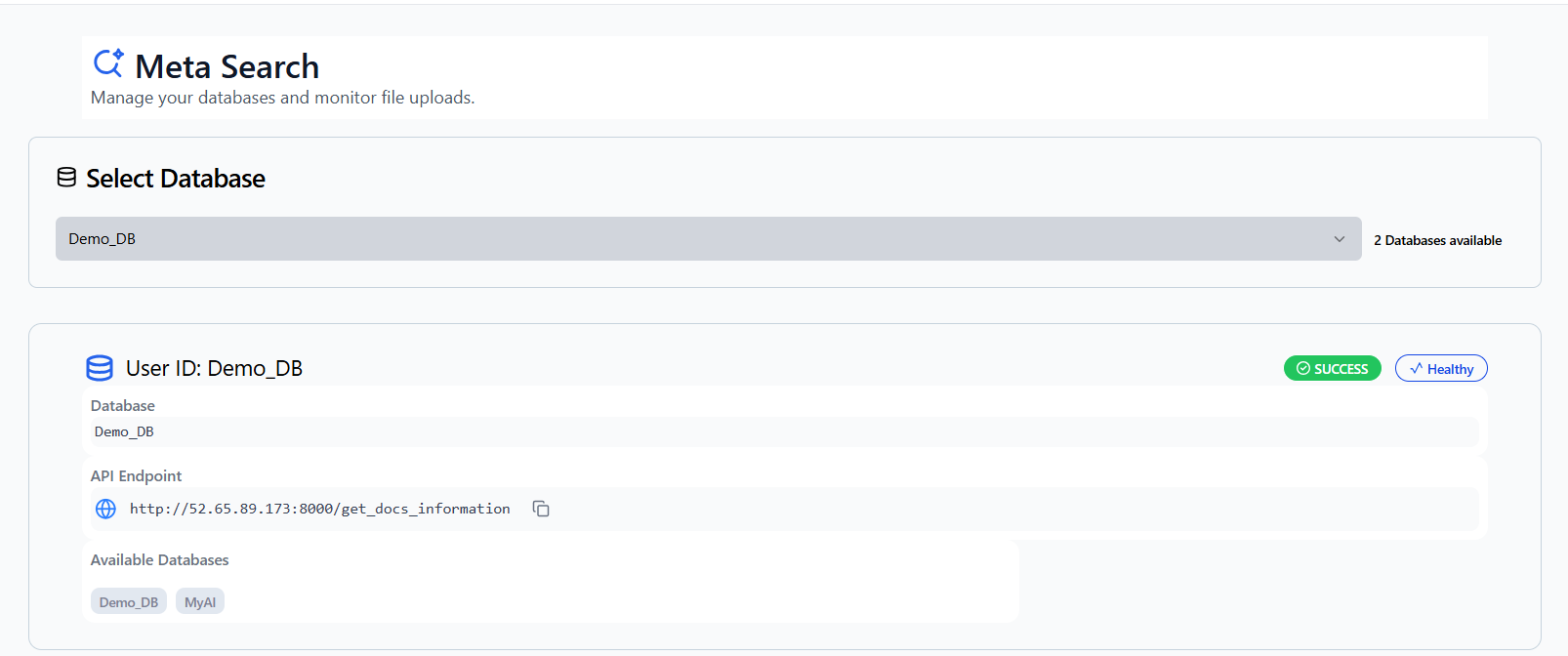
|  |  |  |
| --- | --- | --- |
| Feature | Description | Use Case |
| File Listing | View all files in a database | Verify successful uploads |
| Search by Name | Find files by filename | Locate specific documents quickly |
| Filter by Type | Show only certain file formats | Focus on PDFs, CSVs, etc. |
| View Metadata | See upload date, size, status | Troubleshoot processing issues |

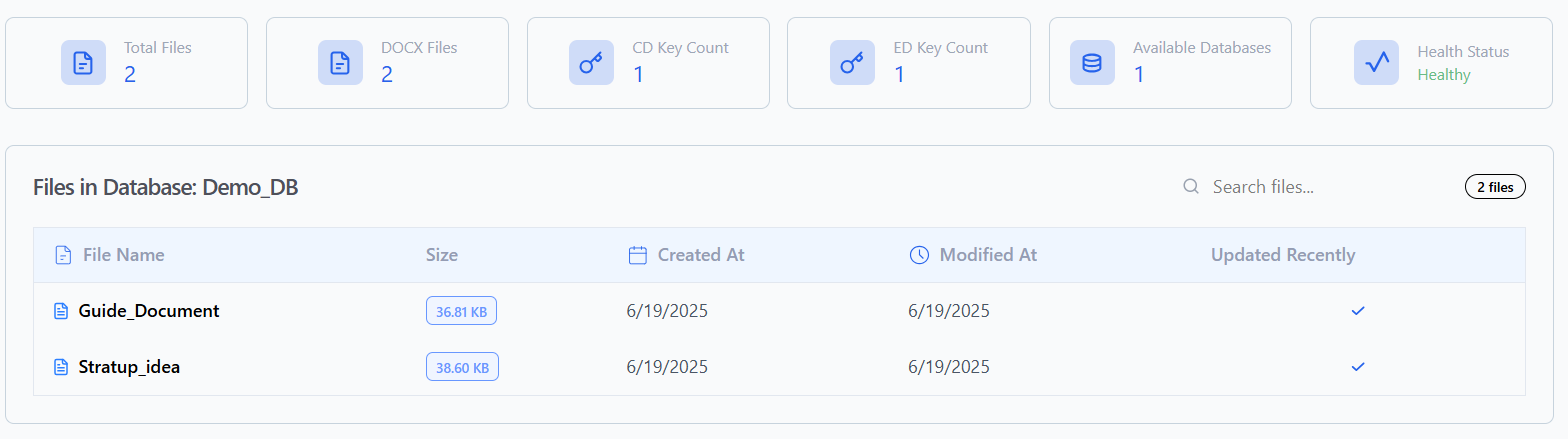
## **8.2 How to Use Meta Search**

1. **Access Meta Search**: Navigate to the Meta Search screen
2. **Select Database**: Choose which database to explore
3. **Browse or Search**:

* **Browse**: Scroll through all files
* **Search**: Use the search bar to find specific files
* **Filter**: Apply filters by file type or upload date

1. **View Details**: Click on any file to see detailed information
2. **Verify Status**: Ensure all files show "Processed" status





## **8.3 Troubleshooting Common Issues**

|  |  |  |
| --- | --- | --- |
| Issue | Possible Cause | Solution |
| File not appearing | Still processing | Wait a few minutes and refresh |
| Processing failed | Unsupported format or corruption | Re-upload with supported format |
| Empty content | File was image-only PDF | Convert to text and re-upload |
| Partial processing | Large file timeout | Split large files into smaller chunks |

# **Dashboard**

The Dashboard is your central command canter, providing a real-time, at-a-glance overview of your entire WaveflowDB account. It is designed to give you immediate insight into your usage and resources without needing to navigate to different sections.

From this single screen, you can instantly see:

* **Account-Wide Summary:**
  + **Total Databases:** The total number of isolated data containers you have created across all your clusters.
  + **Total Files:** A cumulative count of all files uploaded across every one of your databases.
  + **Total Storage Used:** The total disk space consumed by all your files and indexed data, giving you a clear picture of your overall footprint.
* **Detailed Database Breakdown:**
  + A clear list showing the **names of each individual database**.
  + The **size of each database**, so you can identify which projects are consuming the most storage.
  + The **number of files within each specific database**, helping you track content volume per project.

# **Manage Access**

**Manage Access** is your team collaboration hub where you can invite team members, assign roles, and control who can access different parts of your WaveflowDB workspace.

**10.1 Access Control Features**

|  |  |  |
| --- | --- | --- |
| Feature | Description | Benefit |
| User Invitations | Send email invites to team members | Easy onboarding |
| Role-Based Access | Assign different permission levels | Secure collaboration |

**10.2 User Roles and Permissions**

|  |  |  |
| --- | --- | --- |
| Role | Permissions | Best For |
| Owner/Admin | • Full system access  • Manage all users  • Create/delete clusters  • Billing management | Team leads, IT administrators |
| Member | • Create/modify databases  • Upload files  • Create AI assistants  • Invite users | Content managers, developers |
| Viewer | • View databases  • Use AI assistants  • Download files  • Meta search | End users, customers, researchers |

**10.3 How to Invite Team Members**

1. **Navigate to Manage Access**: Go to the team management section
2. **Click Invite Member**: Start the invitation process
3. **Enter Details**:

* **Email Address**: Provide the team member's email
* **Access Level**: Choose which databases they can access

1. **Send Invitation**: The user will receive an email invitation
2. **Track Status**: Monitor invitation acceptance and user activity

