

Building a Knowledge Buddy in Pega: Creating Data Sources, Collections, and Ingesting Document Data

In today's digital landscape, intelligent virtual assistants—what we often call "Knowledge Buddies"—are revolutionizing the way enterprises deliver support, enable employees, and manage customer interactions. Pega Knowledge offers a powerful platform to structure and serve contextual knowledge through intelligent search and AI-driven suggestions.

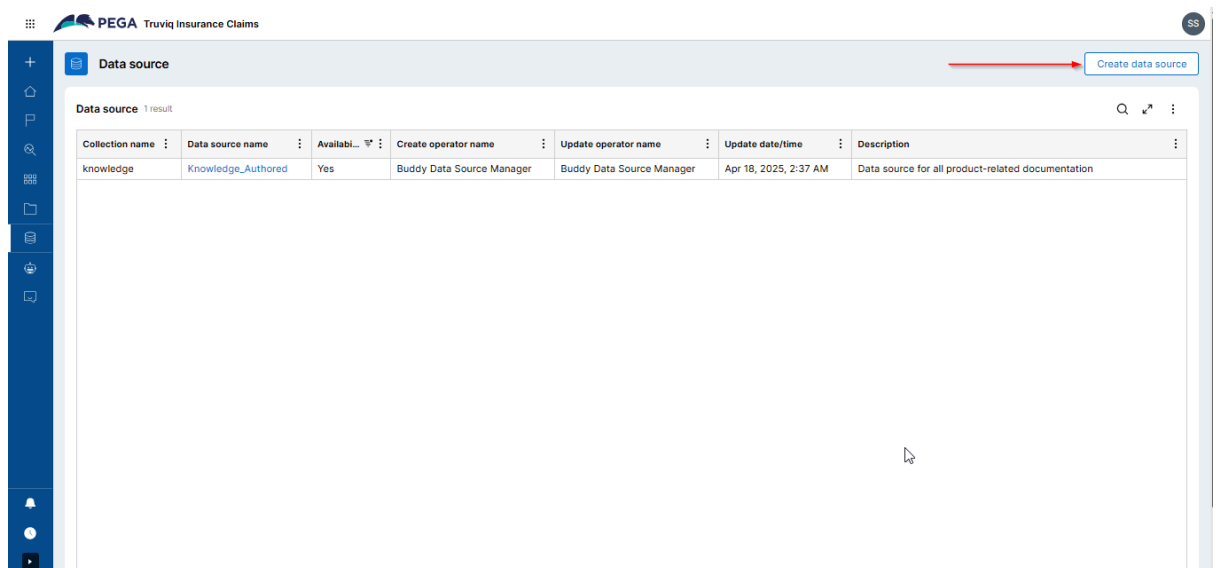
In this guide, we walk through the step-by-step process of preparing content by creating data sources and collections and ingesting external documents into the Knowledge Management (KM) system—forming the knowledge base backbone for your Knowledge Buddy.

Step 1: Define Your Data Source

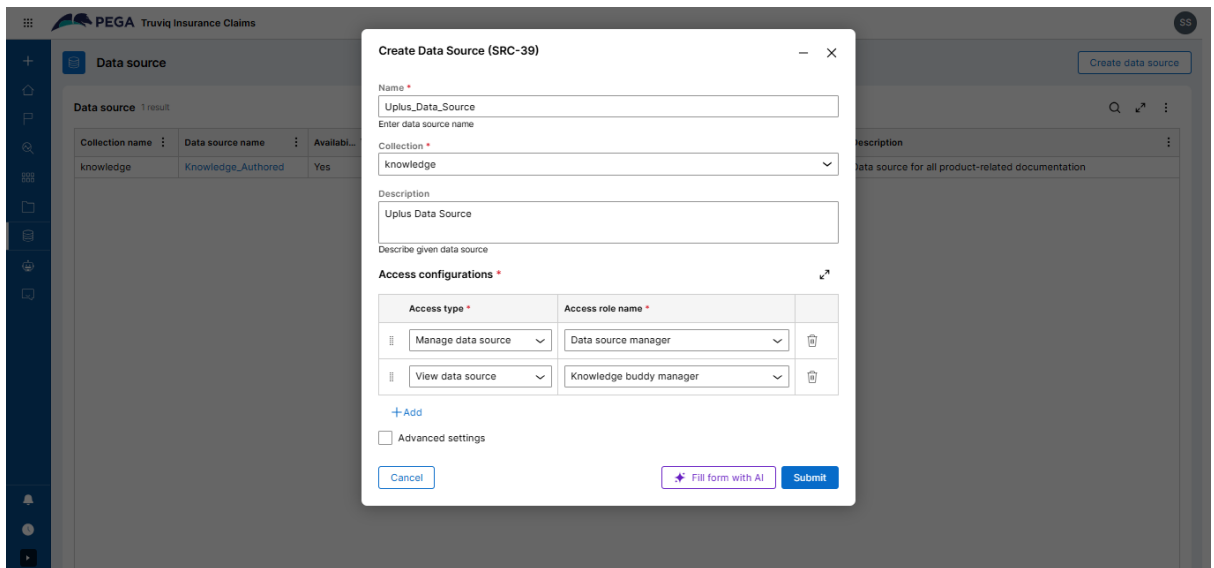
Before content ingestion, you must first create a Data Source in Pega Knowledge. This represents the origin of external documents—whether PDF manuals, Word documents, or web content.

How to Create a Data Source:

1. Navigate to **Knowledge > Administration > Data Sources**
2. Click on **Add Data Source**



3. Fill out the following fields:
 - **Name:** Give your data source a meaningful name (e.g., *Product Manuals Repository*)
 - **Source Type:** Choose the appropriate type (e.g., *Document Upload*, *Web Crawler*, or *API Feed*)
 - **Description:** Provide a clear description of the context



4. Click **Submit** to save.

Once your data source is created, it becomes selectable during the content ingestion process.

Step 2: Set Up Collections

Collections in Pega Knowledge allow you to categorize content into meaningful groups—for example, by product line, department, or content type. This structure is crucial for filtering and delivering relevant articles to your end users.

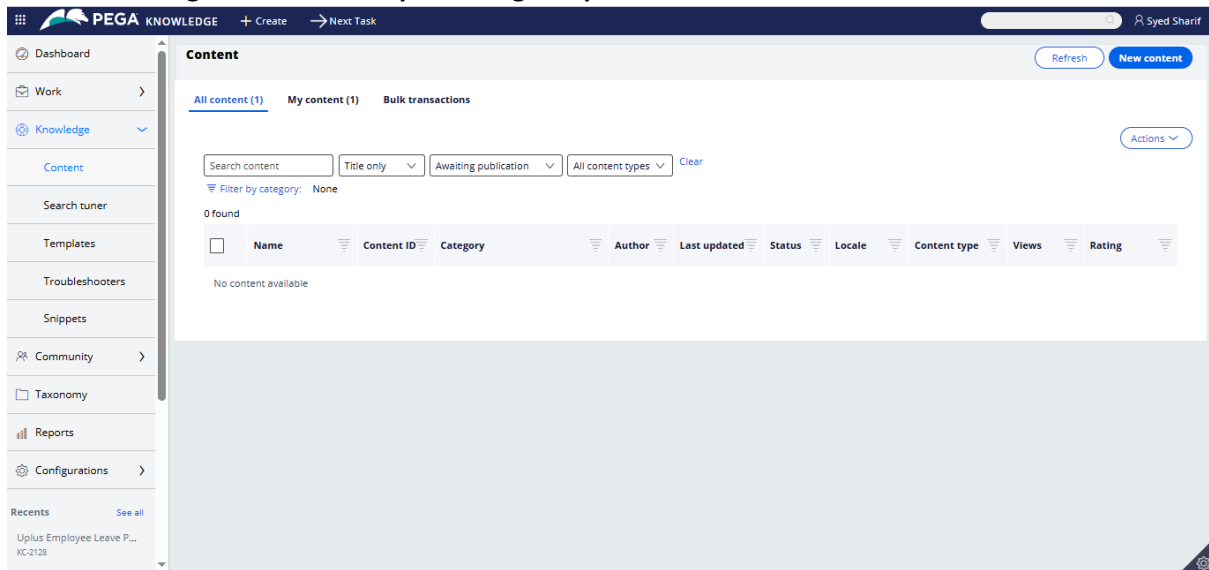
By default, Pega provides a default collection of **Knowledge** and all the Data Sources will be associated with it.

Step 3: Ingest Document Data into Knowledge Articles

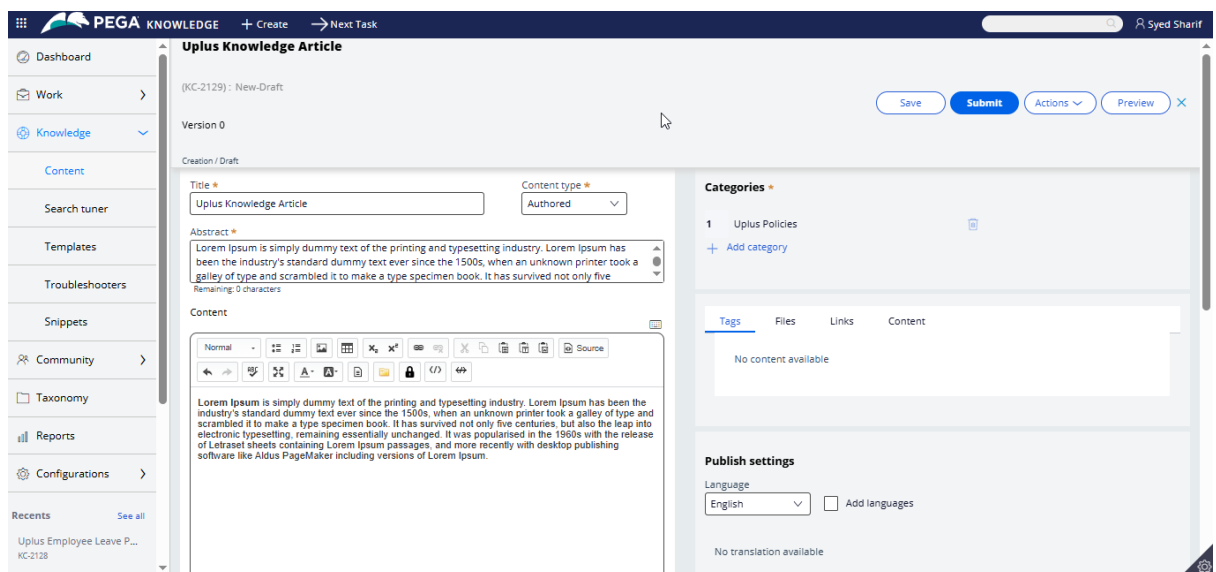
Pega Knowledge supports ingesting content from documents using the built-in AI-based ingestion pipeline. The platform will parse, extract, and transform document content into structured knowledge articles.

Ingestion Steps:

Go to Knowledge > Content > Import Using Template or Click on New Content



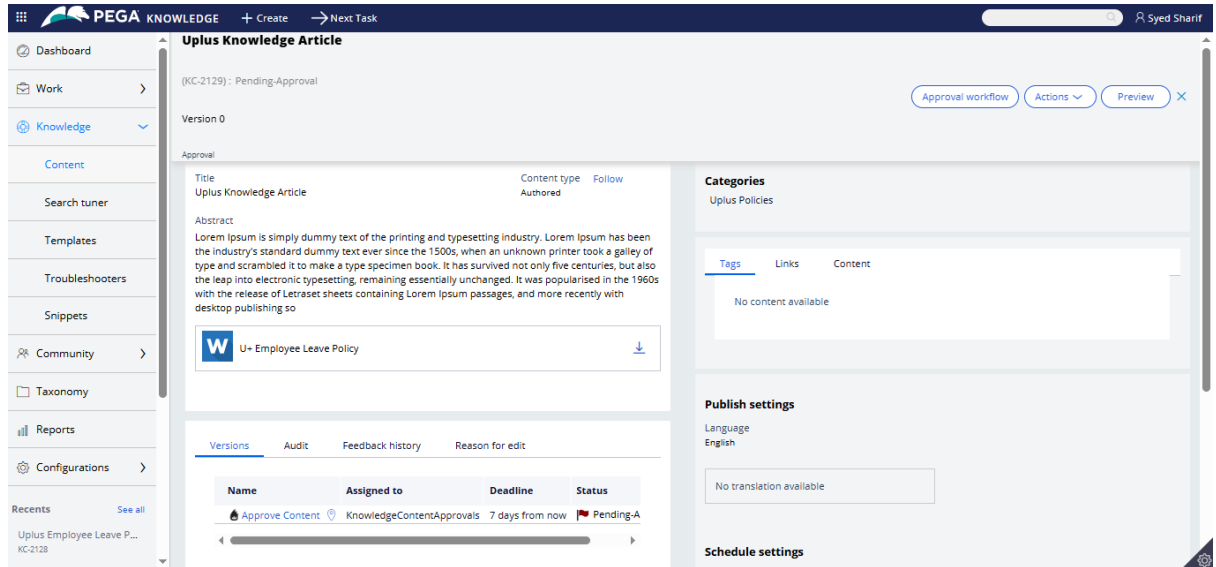
1. Give Title, Content type, Abstract, and Content, and Select the Category for text ingestion.



Give Title, Content type, Abstract, and Content, and Select the Category. Select the Article type as an attachment article only, upload the document, and submit.

2. Upload your document(s):
 - Supported types: .docx, .pdf, .txt, .html

3. Select Approval workflow and Approve the Knowledge Content.



Step 4: Review and Publish Articles

Once the documents are processed:

1. Navigate to **Knowledge > Content > Articles**
2. Open articles in **Draft** or **Pending Review**
3. Review the extracted content for accuracy and completeness
4. Edit metadata, tags, and categories as needed
5. Set visibility roles and publication dates
6. Click **Publish**

Published articles are now part of the searchable knowledge base, ready to be served by the Knowledge Buddy or suggested in case workflows.

Step 5: Create Knowledge Buddy

Once the Knowledge Articles are Published:

1. Navigate to **Knowledge Portal > Knowledge Buddy > Create Knowledge Buddy**
2. Give the **Name and Description** of the Buddy

3. Configure the Access

UPlus_Buddy BUDDY-18

Urgency 10 Work Status NEW

Created Syed Sharif 1 minute ago

Updated Syed Sharif now

Details

Execution history

Pulse

History

Configure security Assigned to Syed Sharif • In BUDDY-18 • Urgency 10

Access

Type *	Role name *
Manage knowledge buddy	Knowledge buddy manager
Use knowledge buddy	KnowledgeBuddy:Public

+ Add

Cancel

Fill form with AI Save for later Submit

Details

Name UPlus_Buddy

Description UplusBuddy

Instructions

I will give you some CONTEXT in the user prompt which will be enclosed in #####.

You MUST use this data only to answer questions.

You are a customer service representative and I would like you to answer some questions based on the CONTEXT I provide. I have some rules you need to follow. Please follow all the rules.

1. Use only the CONTEXT you are provided with.
2. In your answer do not mention you are a customer service representative.
3. Do not refer to anyone or anything that is not part of the CONTEXT provided.
4. If you don't know the answer, just say you don't know and you have flagged the question for internal review.
5. Never make up answers which is not in the CONTEXT provided.
6. You can only use NOUNS from the CONTEXT provided in your response.
7. Never repeat the customers question in your response if you do not have an answer.

4. In Configure Prompt Select Edit

UPlus_Buddy BUDDY-18

Urgency 10 Work Status NEW

Created Syed Sharif 3 minutes ago

Updated Syed Sharif now

Details

Execution history

Pulse

History

Configure prompt

Information

CONTEXT: #####

(SEARCHRESULTS)

#####

This is the end of CONTEXT. Only text above can be used to answer the question.

QUESTION: (QUESTION)

Contextual data definitions 1 result

Context name	Collection	Must return	Max chunk	Data sources	Response attributes	Minimum similarity
SEARCHRESULTS	—	Yes	5,000	—	—	80.00

+ Add Contextual data definition

GenAI model

Select...

Output format

Custom

☐ Apply text replacements to user request

Cancel

Fill form with AI Save for later Submit

Details

Name UPlus_Buddy

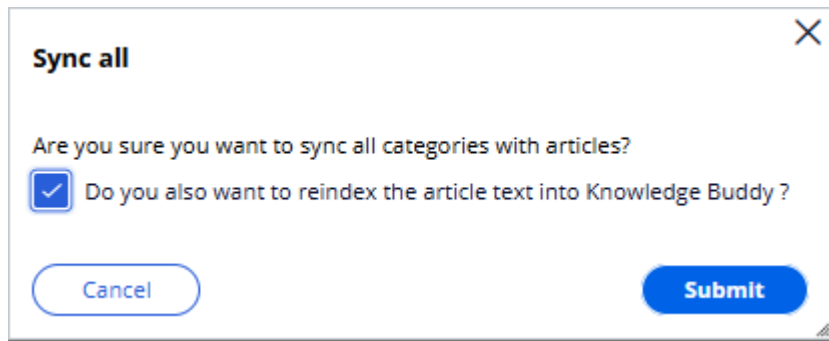
Description UplusBuddy

5. Select Collection and Data Source
6. By default, the Data Source will be associated with the Knowledge Collection
7. After submission, we can use the Buddy Playground to test the prompts and responses.

Step6: Syncing Knowledge Articles into Knowledge Buddy

Once the Articles are Published:

1. Navigate to **Knowledge > Taxonomy > Article Synchronization > Sync all**
2. Select **Do you also want to reindex the article text into Knowledge Buddy** and Submit.



Exploring Pega Knowledge Buddy: Leveraging APIs for Smart Knowledge Management

In today's fast-paced digital landscape, organizations are looking for smarter ways to manage and interact with knowledge. **Pega Knowledge Buddy** is an innovative solution that brings intelligence and interactivity to enterprise knowledge systems. What makes it even more powerful is its **API-first architecture**, allowing seamless integration and automation.

We'll explore how Pega Knowledge Buddy supports APIs to enable functionalities such as:

1. Creating collections
2. Defining data sources
3. Ingesting documents
4. Asking questions
5. Listing available buddies

API Capabilities

Here's a breakdown of the core API functionalities supported by Knowledge Buddy

1. Create a Knowledge Collection

Collections are logical groupings of knowledge documents. With the API, you can programmatically create a collection by providing metadata and configuration details.

Example API Call:

POST /knowledgebuddy/v1/collections

Payload:

```
{  
  "name": "Loan Servicing FAQ",  
  "description": "FAQs and documents related to loan servicing operations."}
```

2. Add or Configure a Data Source

Knowledge Buddy supports various data sources like files, SharePoint, websites, and enterprise systems.

Example API Call:

```
POST /knowledgebuddy/v1/datasources
```

Payload:

```
{  
  "type": "web",  
  "url": "https://example.com/loan-faq"}
```

3. Ingest Documents

You can upload and ingest documents (PDFs, Word files, etc.) into a collection to make them searchable and usable in conversations.

Example API Call:

```
POST /knowledgebuddy/v1/collections/{collectionId}/documents
```

Payload:

```
{  
  "file": "<binary-pdf-or-docx-content>",  
  "title": "Loan Terms and Conditions"}
```

4. Ask Questions

You can send a user question and receive an intelligent, contextual answer based on the ingested knowledge and LLM integration.

Example API Call:

```
POST /knowledgebuddy/v1/query
```

Payload:

```
{  
  "question": "What are the penalties for late loan payments?",  
  "collectionId": "loan-servicing-faq"}
```

5. List All Knowledge Buddies

Organizations might have multiple Knowledge Buddies for different departments or functions (HR, IT, Finance, etc.). You can fetch a list of available buddies via API.

Example API Call:

```
GET /knowledgebuddy/v1/buddies
```

Use Cases Enabled by These APIs

1. **Automated Knowledge Curation:** Automatically ingest updated content from external systems.
2. **Dynamic FAQ Generation:** Use form submissions or tickets to generate and query FAQs.
3. **Personalized Knowledge Assistants:** Create department-specific assistants with tailored collections.
4. **AI-Driven Support Bots:** Seamlessly integrate Knowledge Buddy with chatbots or voice bots.

Best Practices

1. **Chunk large documents** into logical sections before ingestion to improve accuracy.
2. Use **consistent formatting** (e.g., headings, bullet points) in source documents to aid parsing.
3. Periodically **audit your content** for relevance and freshness.
4. Train support staff on using the Knowledge Buddy effectively.

Final Thoughts

Establishing a well-organized, searchable knowledge base is the foundation of a successful Knowledge Buddy. By leveraging Pega's document ingestion and taxonomy features, you can rapidly scale your knowledge assets and empower both customers and employees with the answers they need—right when they need them. It's an intelligent, API-driven assistant designed to evolve with your organization's needs. With its powerful API support, you can integrate knowledge capabilities into any part of your digital ecosystem—whether that's a chatbot, service desk, or employee portal.