

Building in Dify: an agent with its own knowledge base



Guide 2 for the AI Agents workshop

LeX Consultancy B.V. · 20 September 2026

In guide 1 you built a workflow: you decided every step. Now you build an agent. It gets a goal, an instruction and a knowledge base with your own document, and decides for itself when to search in it. At the end you have a chatbot with its own web address that answers questions about your document and honestly says when the answer is not in there.

This guide follows the cloud version of Dify as it looked on 20 September 2026. Dify changes fast: if your screen differs, follow what you see, not what is written here.

How to read this guide

Form	Meaning
1. 2. 3.	What you do: click, choose, open
Bold	A button, menu or field as it appears on your screen
Table <i>What you type</i>	Values you put in a form field
Grey box <i>Type in</i>	Text you type or paste literally
Orange box	A pitfall you would otherwise discover yourself
Green box <i>Your choice</i>	A place where you can replace our example with your own material, style or prompt

About credits: an agent is more expensive than a workflow. In our test, creating the knowledge base plus one retrieval test cost 26 credits, and every answer from the agent 6 to 7 credits (with gpt-5-mini). With the 200 credits of a sandbox account you can follow this guide once and then ask about twenty questions. So test with purpose, not for fun.

The design card for this app

Before we started clicking, this card was filled in (the blank card is in the Design card appendix). Fill in the same five boxes for your own app; the loose version of this filled-in card is in 00-design-card/examples.

1. Who is it for	2. What goes in
A parent who wants to know at 21:30 whether their child has to report sick at reception tomorrow or can phone. Has never read the school guide.	One question in plain language, in the chat. No form.

3. What comes out	4. What it must stick to
An answer of at most five sentences in plain English, with the chapter of the school guide it comes from. Or: "that is not in the school guide, please call reception".	Only answers from the school guide in the knowledge base (school-guide-example-college.pdf); invent nothing; refer on when in doubt. Knowledge base: the school guide, split per paragraph.

5. How you know it works

Test	What goes in	What must come out
1	What time does the first lesson start?	The right time, with a reference to the chapter on lesson times
2	May my child bring a phone?	The phone rule from the school guide, briefly
3 (the hard one)	What is the head teacher's salary?	"That is not in the school guide" and a referral to reception

Preparation

- You have done guide 1, or at least you know where **Studio** and **Create** are.
- You have a document ready that the agent must know: PDF, Word, text or markdown, at most 15 MB and one file at a time in the sandbox. Choose a document with real text, not scanned images.
- Put no personal data in it. Everything you upload sits on Dify's servers.

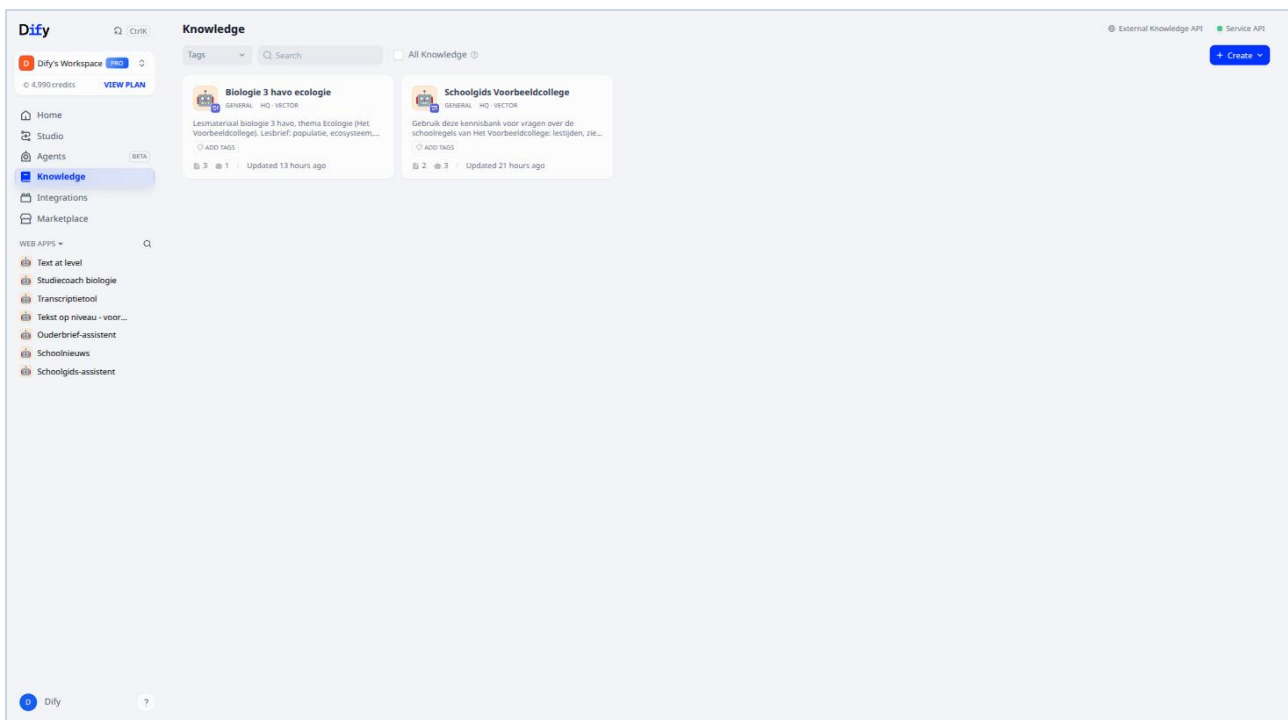
In this guide we use a fictional two-page school guide: `school-guide-example-college.pdf`. It comes with this guide. Feel free to use it; then all the example questions match.

Your choice: your own document. If you want to work with your own material straight away, pick a document that colleagues or students already ask questions about: the exam regulations, the study planner of a subject, the protocol for school trips, the manual of the student information system. Conditions: real text (not a scan), no personal data, and preferably with headings; those help with the splitting in step 2. You then adapt the questions in steps 5 and 10 to your document.

Part A. The knowledge base

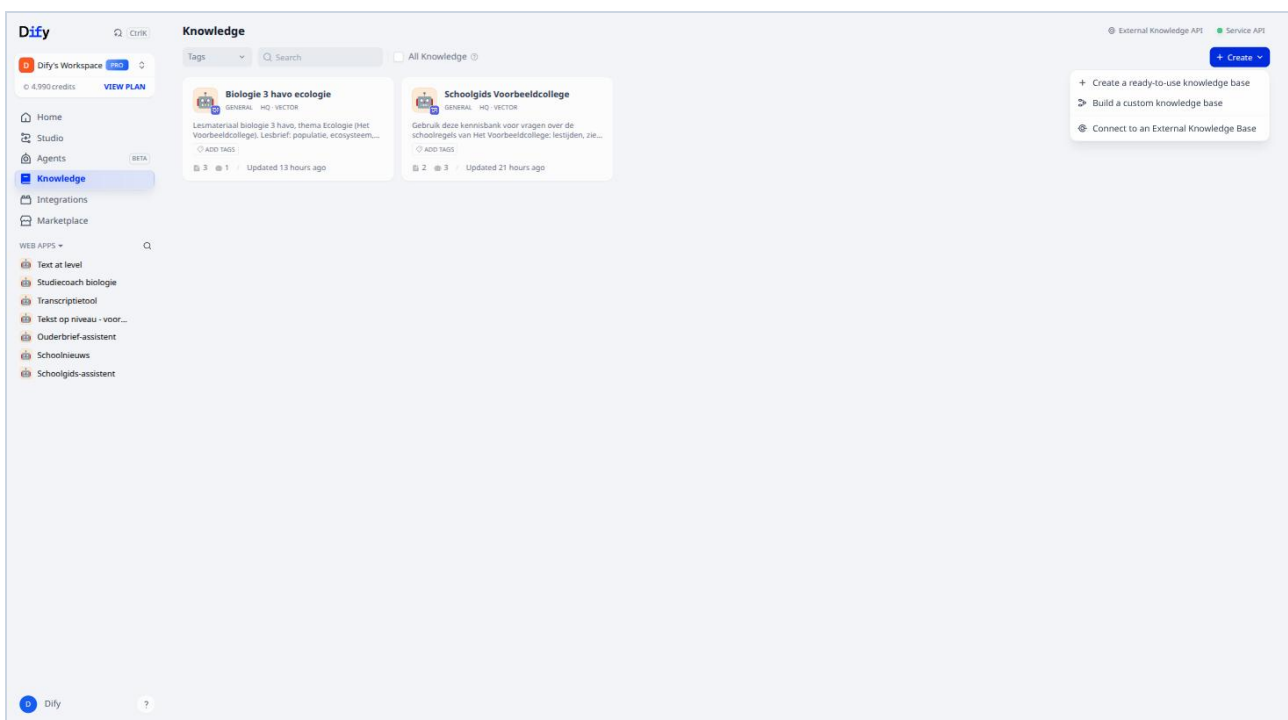
Step 1. Create a knowledge base

1. Click **Knowledge** in the left-hand menu.



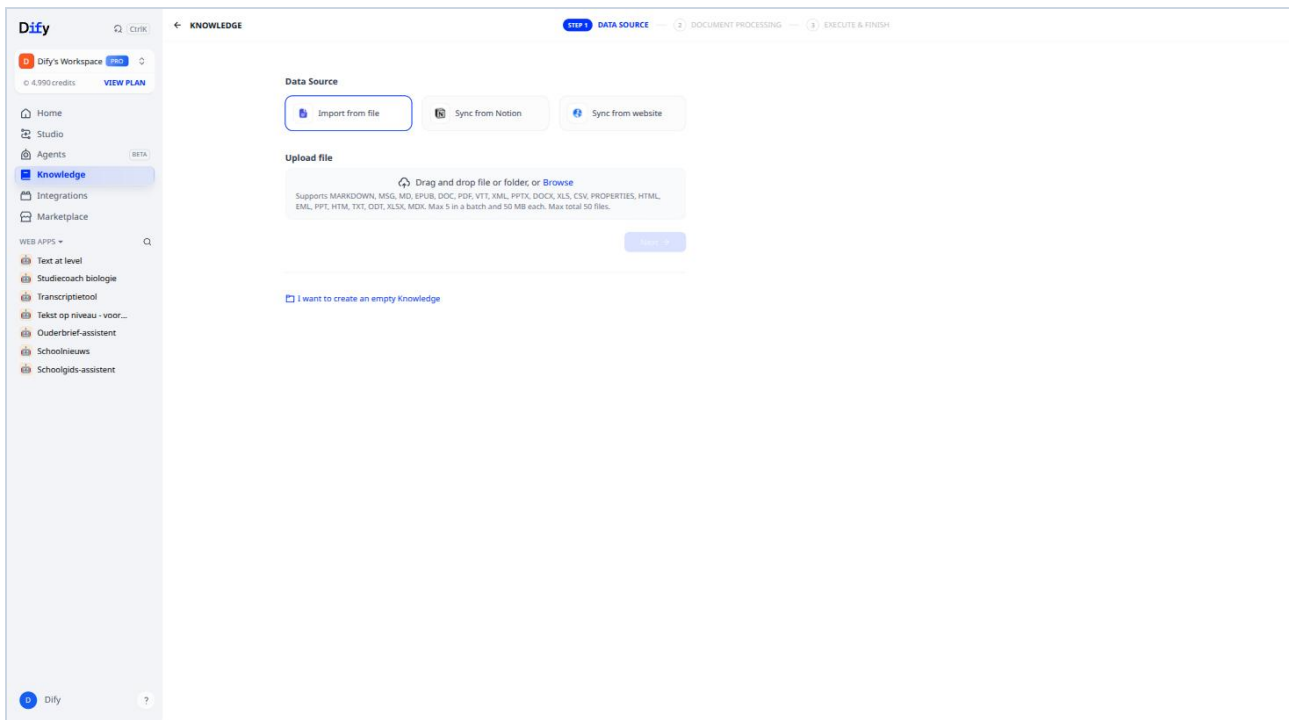
The Knowledge page

2. Click **Create** at the top right and choose **Create a ready-to-use knowledge base**.



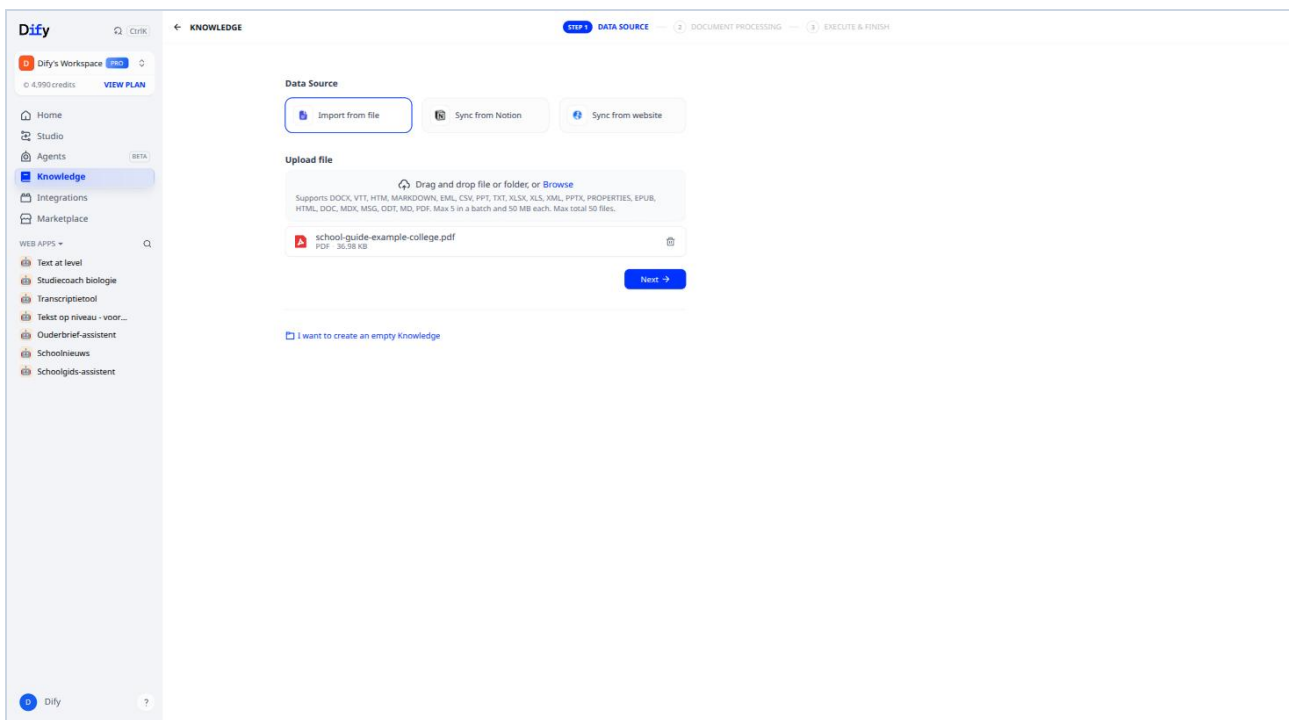
The Create menu

3. Leave **Import from file** selected.



Step 1 of 3: Data Source

4. Click **Browse** and choose your document, or drag it into the box.

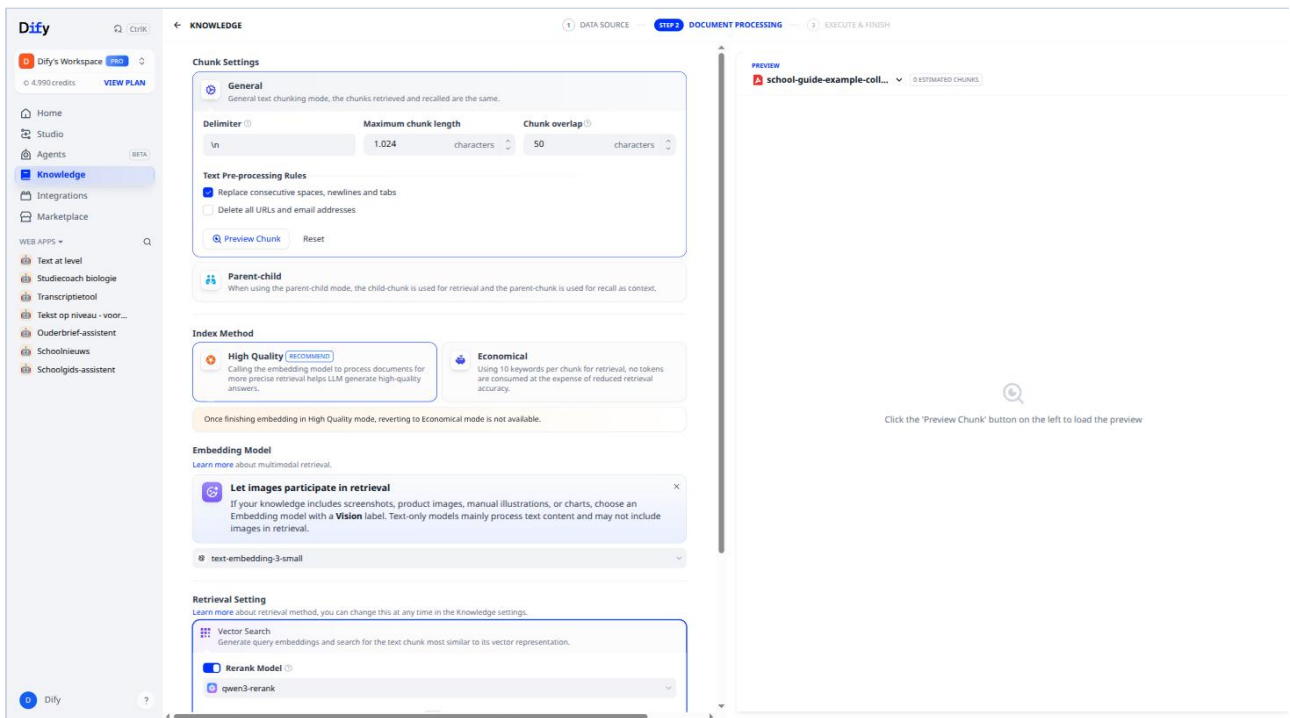


The file is ready

5. Click **Next**.

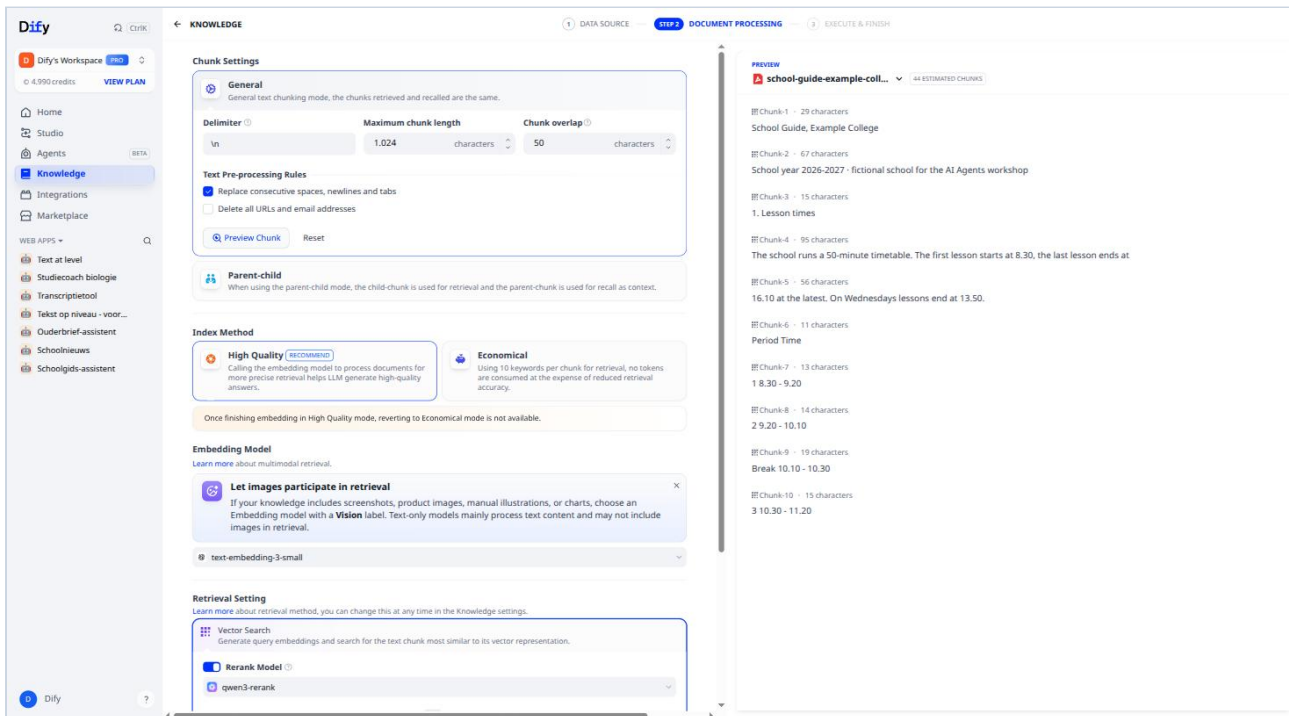
Step 2. Check how the document is split

Dify splits your document into pieces (“chunks”). Later the agent searches in those pieces, not in the whole document. Pieces that are too small contain no usable answer; this is where you check that.



Step 2 of 3: Document Processing

1. Click **Preview Chunk**. The pieces appear on the right.



The preview with the default setting: 44 chunks

Do you see many one-line pieces, such as “1. Lesson times” or “1 8.30 - 9.20”? Then Dify is splitting at every line break. That often happens with PDFs. A twelve-character piece cannot answer a question.

2. Click in the **Delimiter** field and replace its contents.

Field on screen	What you type
Delimiter	\n\n

3. Click **Preview Chunk** again. On the right you now have three pieces of a few paragraphs each.

The screenshot shows the Dify Knowledge interface. On the left is a sidebar with navigation options: Home, Studio, Agents, Knowledge (selected), Integrations, and Marketplace. Below these are 'WEB APPS' including 'Text at level', 'Studiecoach biologie', 'Transcripttool', 'Tekst op niveau - voor...', 'Ouderbrief-assistent', 'Schoolnieuws', and 'Schoolgids-assistent'. The main area is titled 'KNOWLEDGE' and has a progress bar with steps: 1. DATA SOURCE, 2. STEP 2: DOCUMENT PROCESSING (active), and 3. EXECUTE & FINISH. The 'Chunk Settings' section is expanded, showing 'General' settings. The 'Delimiter' is set to '\n\n', 'Maximum chunk length' is 1,024 characters, and 'Chunk overlap' is 50 characters. Under 'Text Pre-processing Rules', there are checkboxes for 'Replace consecutive spaces, newlines and tabs' (checked) and 'Delete all URLs and email addresses'. A 'Preview Chunk' button is visible. Below this is the 'Index Method' section with two options: 'High Quality' (recommended) and 'Economical'. The 'Embedding Model' section shows 'text-embedding-3-small' selected. The 'Retrieval Setting' section shows 'Vector Search' and 'Rerank Model' set to 'qwen3-rerank'. On the right, a 'PREVIEW' section shows three chunks of text from a 'school-guide-example-coll...' document. Each chunk is a paragraph of text, demonstrating that the new delimiter has successfully split the text into three separate pieces.

After the change: 3 chunks with real content

4. Leave **Maximum chunk length** (1024) and **Chunk overlap** (50) as they are.

5. Under **Index Method** leave **High Quality** selected.

6. Scroll down. Leave **Embedding Model** and **Retrieval Setting** at their defaults.

Home

Studio

Agents

Knowledge

Integrations

Marketplace

WEB APPS +

Text at level

Studiecoach biologie

Transcriptietool

Tekst op niveau - voor...

Ouderbrief-assistent

Schoolnieuws

Schoolgids-assistent

0

Dify

?

KNOWLEDGE

1 DATA SOURCE2 STEP 3 DOCUMENT PROCESSING3 EXECUTE & FINISH

Parent-child

When using the parent-child mode, the child chunk is used for retrieval and the parent-chunk is used for recall as context.

Index Method

High Quality

Calling the embedding model to process documents for more precise retrieval helps LLM generate high-quality answers.

Economical

Using 10 keywords per chunk for retrieval, no tokens are consumed at the expense of reduced retrieval accuracy.

Once finishing embedding in High Quality mode, reverting to Economical mode is not available.

Embedding Model

Learn more about multimodal retrieval.

Let images participate in retrieval

If your knowledge includes screenshots, product images, manual illustrations, or charts, choose an Embedding model with a **Vision** label. Text-only models mainly process text content and may not include images in retrieval.

text-embedding-3-small

Retrieval Setting

Learn more about retrieval method, you can change this at any time in the Knowledge settings.

Vector Search

Generate query embeddings and search for the text chunk most similar to its vector representation.

Rerank Model

qwen3-rerank

Top K

3

Score Threshold

0,5

Full-Text Search

Index all terms in the document, allowing users to search any term and retrieve relevant text chunk containing those terms.

Hybrid Search

Execute full-text search and vector searches simultaneously, re-rank to select the best match for the user's query. Users can choose to set weights or configure to a Rerank model.

Previous step

Save & Process

Preview

school-guide-example-coll...3 ESTIMATED CHunks

Chunk-1823 characters

School Guide, Example College School year 2026-2027: fictional school for the AI Agents workshop 1. Lesson times The school runs a 50-minute timetable. The first lesson starts at 8.30, the last lesson ends at 16.10 at the latest. On Wednesdays lessons end at 13.50. Period Time 1 8.30 - 9.20 2 9.20 - 10.10 Break 10.10 - 10.30 3 10.30 - 11.20 4 11.20 - 12.10 Break 12.10 - 12.40 5 12.40 - 13.30 6 13.30 - 14.20 Break 14.20 - 14.30 7 14.30 - 15.20 8 15.20 - 16.10 2. Reporting illness and absence If a student is ill, a parent or carer reports this before 8.15 through the parent portal or by phone to reception on 01632 960123. A sick report is valid for one day if the student is still ill the next day, a new report is needed. Unauthorised absence is reported to the form tutor after three periods.

Chunk-2522 characters

After sixteen hours of unauthorised absence in four weeks the school is obliged to inform the local attendance officer. Leave outside the school holidays can only be requested from the head of year, at least six weeks in advance and with a written reason. 3. Homework and tests Homework is posted in the digital learning environment by 17.00 on the day of the lesson. Students in years 7 to 9 have at most two tests per day and at most five tests per week. A test is announced at least five school days in advance.

Chunk-3891 characters

A missed test is taken at the fixed catch-up slot: Tuesday, period 8. The student registers with the teacher within two school days of returning to school. 4. Phones and laptops Phones go into the phone pouch on entering the classroom and stay there until the end of the lesson. Phones may be used during breaks and in the hall. Every student brings a charged laptop every day; the school does not lend chargers. 5. Form tutors and support Every class has a form tutor. The form tutor is the first point of contact for students and parents. For questions about dyslexia, test anxiety or extra support, the form tutor refers to the support coordinator. A meeting with the support coordinator is always requested through the form tutor. 6. Contact Reception: 01632 960123, on school days from 8.00 to 16.30. Email: info@examplecollege.example Address: 1 School Street, Exampletown

Embedding Model and Retrieval Setting

7. Click **Save & Process**.

Step 3. Wait until processing is finished

Home

Studio

Agents

Knowledge

Integrations

Marketplace

WEB APPS +

Text at level

Studiecoach biologie

Transcriptietool

Tekst op niveau - voor...

Ouderbrief-assistent

Schoolnieuws

Schoolgids-assistent

0

Dify

?

KNOWLEDGE

1 DATA SOURCE2 DOCUMENT PROCESSING3 STEP 3 EXECUTE & FINISH

Knowledge created

We automatically named the Knowledge, you can modify it at any time.

Knowledge name

school-guide-examp...

EMBEDDING PROCESSING...

Process more data with higher accuracy at faster speeds.

Upgrade Now

school-guide-example-collge.pdf

Processing

0%

Chunking Setting

Custom

Maximum Chunk Length

1024

Text Preprocessing Rules

Replace consecutive spaces, newlines and tabs

Index Method

High Quality

Retrieval Setting

Vector Search

Access the API

Go to document

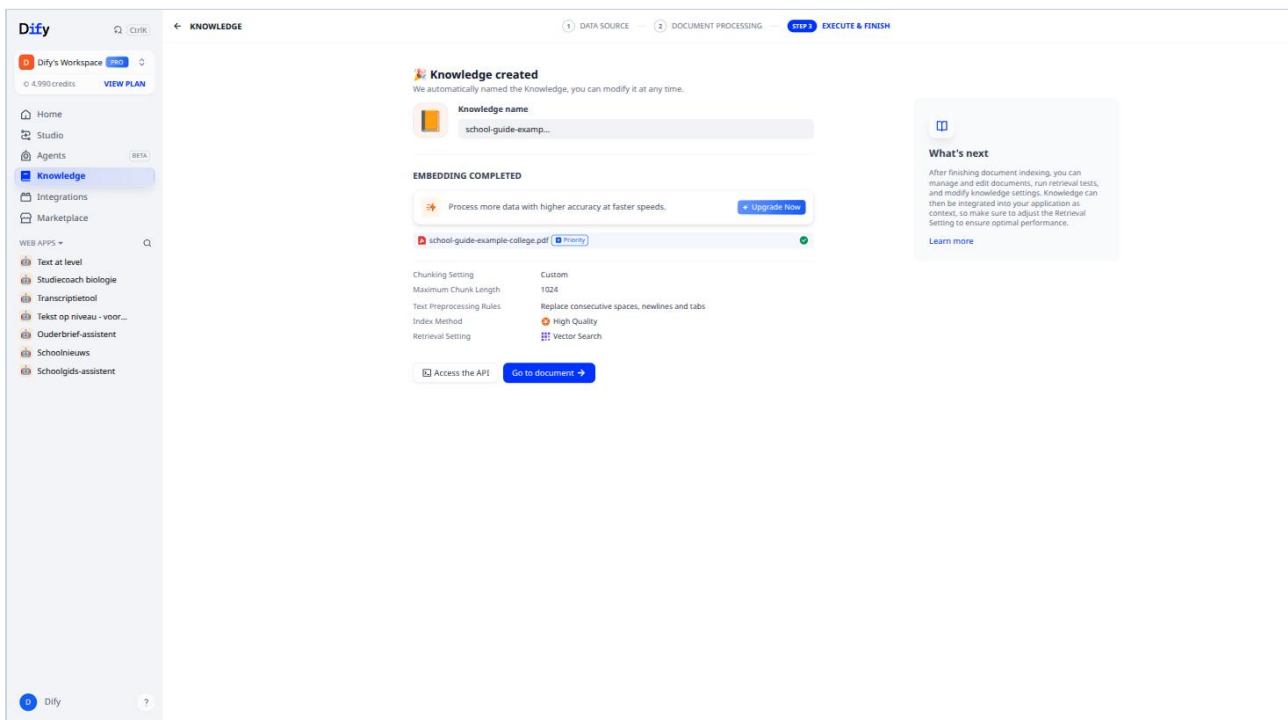
What's next

After finishing document indexing, you can manage and edit documents, run retrieval tests, and modify knowledge settings. Knowledge can then be integrated into your application as context, so make sure to adjust the Retrieval Setting to ensure optimal performance.

Learn more

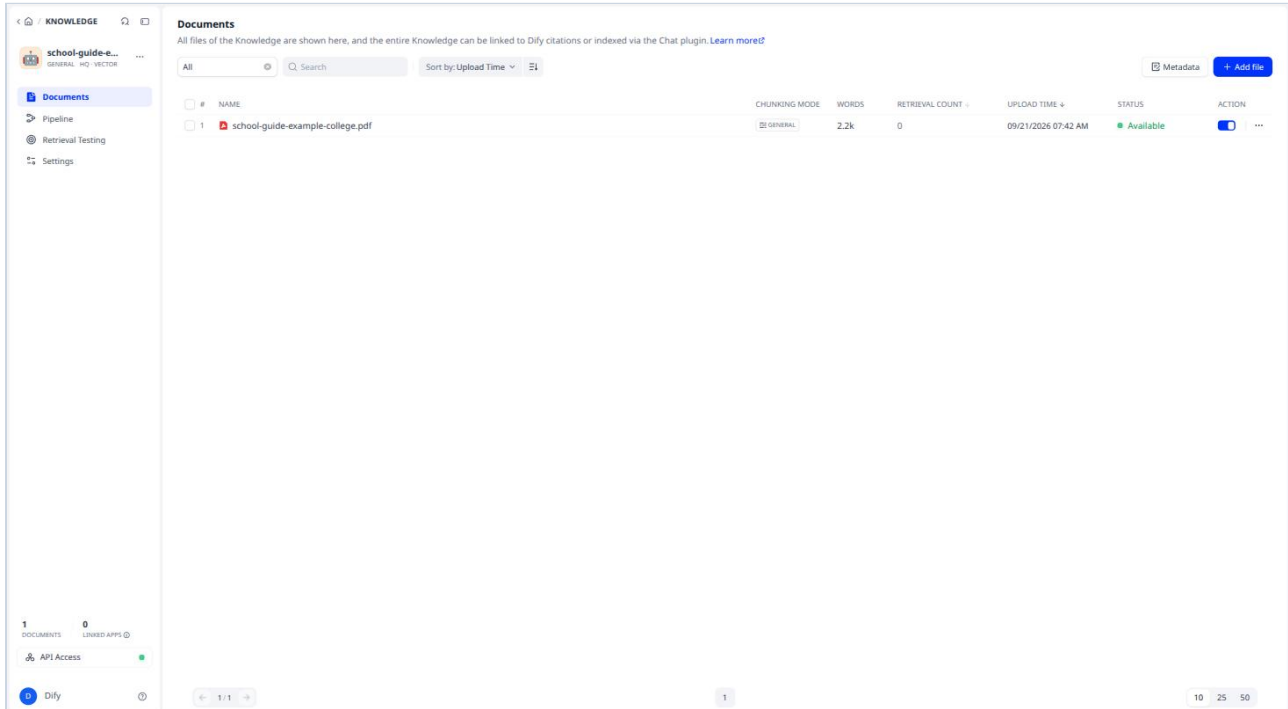
Knowledge base created, embedding in progress

For a two-page document this takes a few seconds. The status jumps from EMBEDDING PROCESSING to **EMBEDDING COMPLETED** with a green tick behind your file.



Embedding completed

1. Click **Go to document**. You see your document in the list with status **Available**.



The document list of the knowledge base

Step 4. Give the knowledge base a name and a description

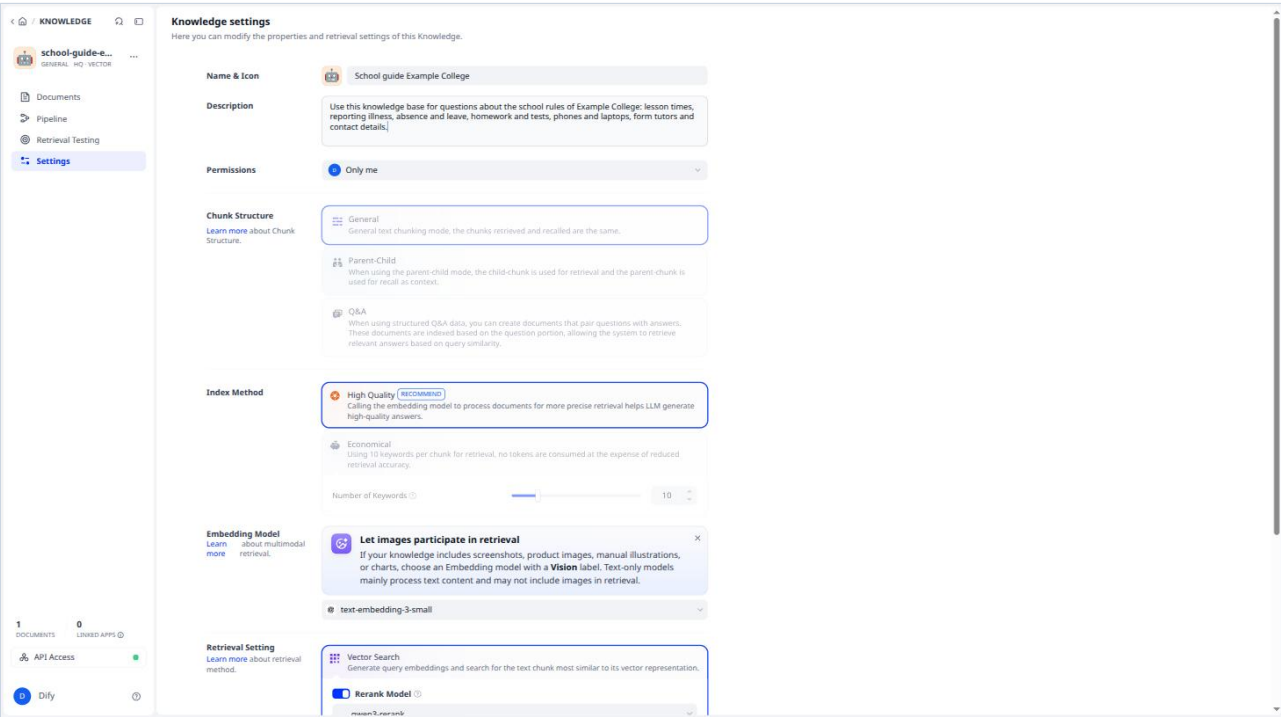
The description is not just for you. The agent reads it to decide whether to consult this knowledge base. A good description therefore says what the knowledge base is about.

- 1. Click **Settings** in the left-hand menu.
- 2. Fill in the two fields.

Field on screen	What you type
Name & Icon	School guide Example College
Description	see the box below

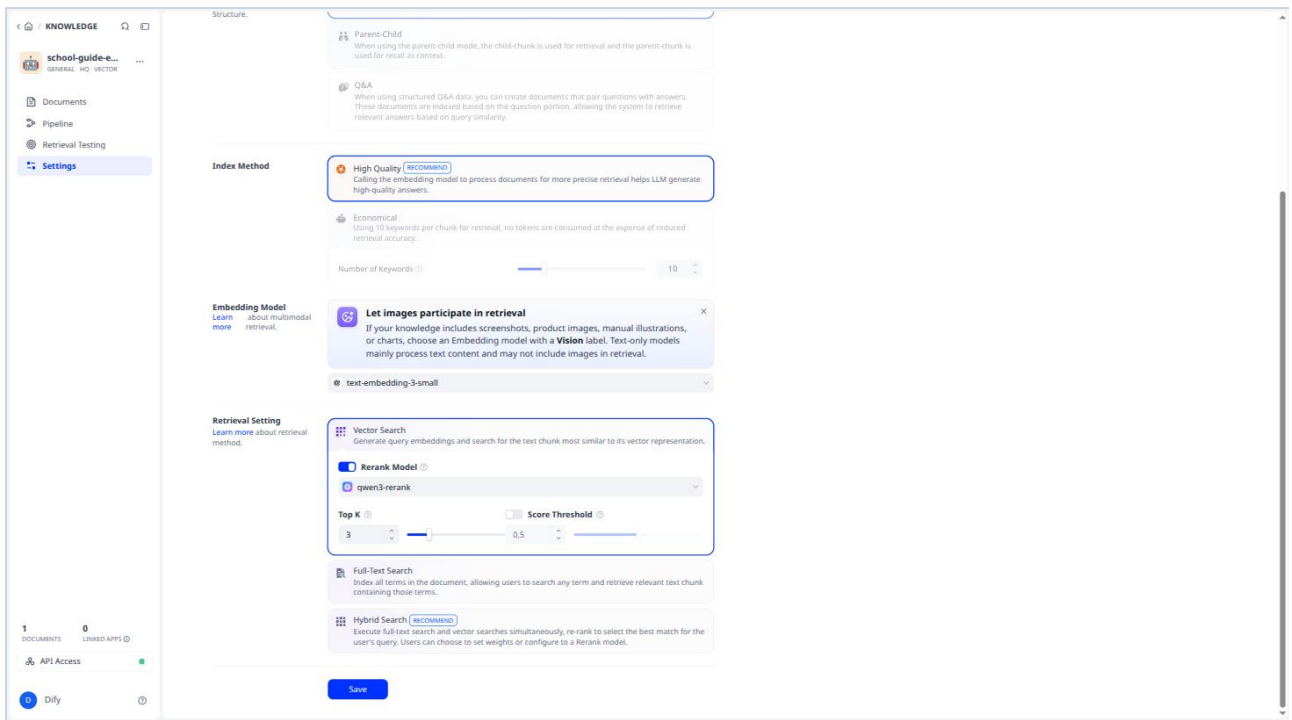
TYPE IN

Use this knowledge base for questions about the school rules of Example College: lesson times, reporting illness, absence and leave, homework and tests, phones and laptops, form tutors and contact details.



Name and description filled in

- 3. Scroll all the way down and click **Save**.

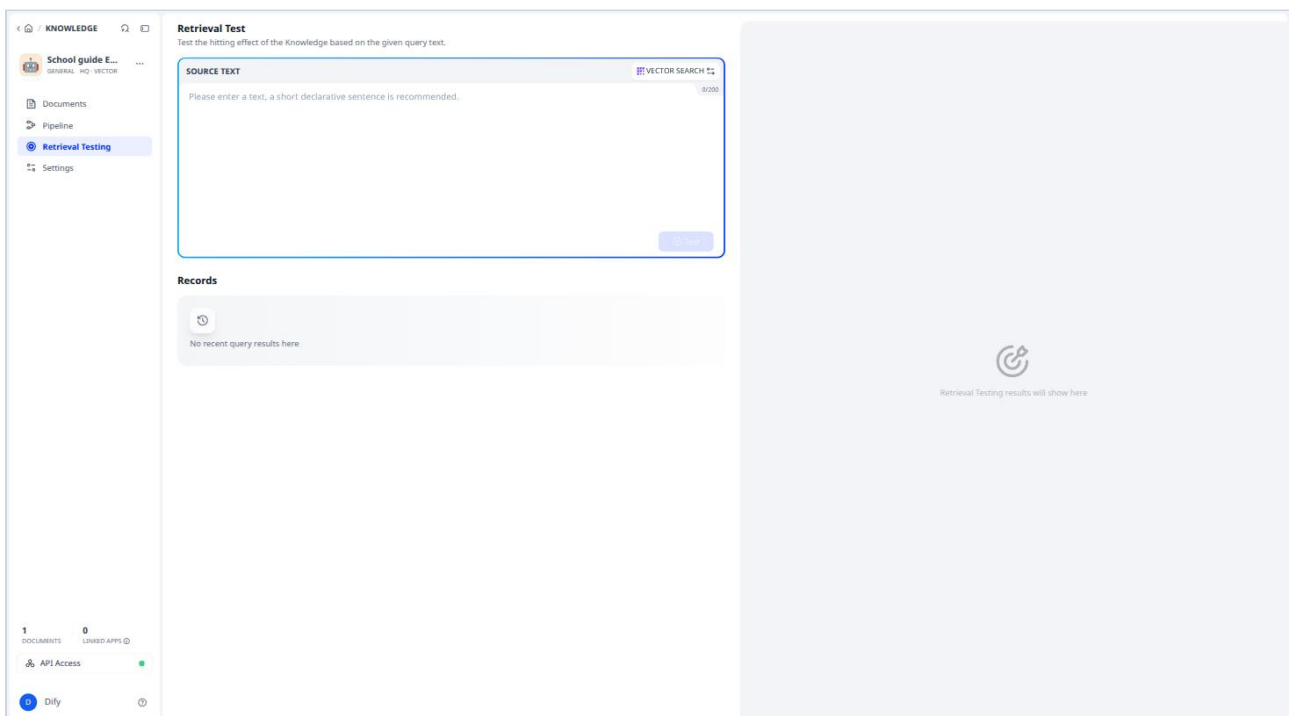


The Save button at the bottom of the settings

Step 5. Test whether the knowledge base finds the right piece

This is the moment you see what a knowledge base does: you ask a question and get not an answer, but the pieces of text that best match your question. Later the agent turns those into an answer.

1. Click **Retrieval Testing** in the left-hand menu.



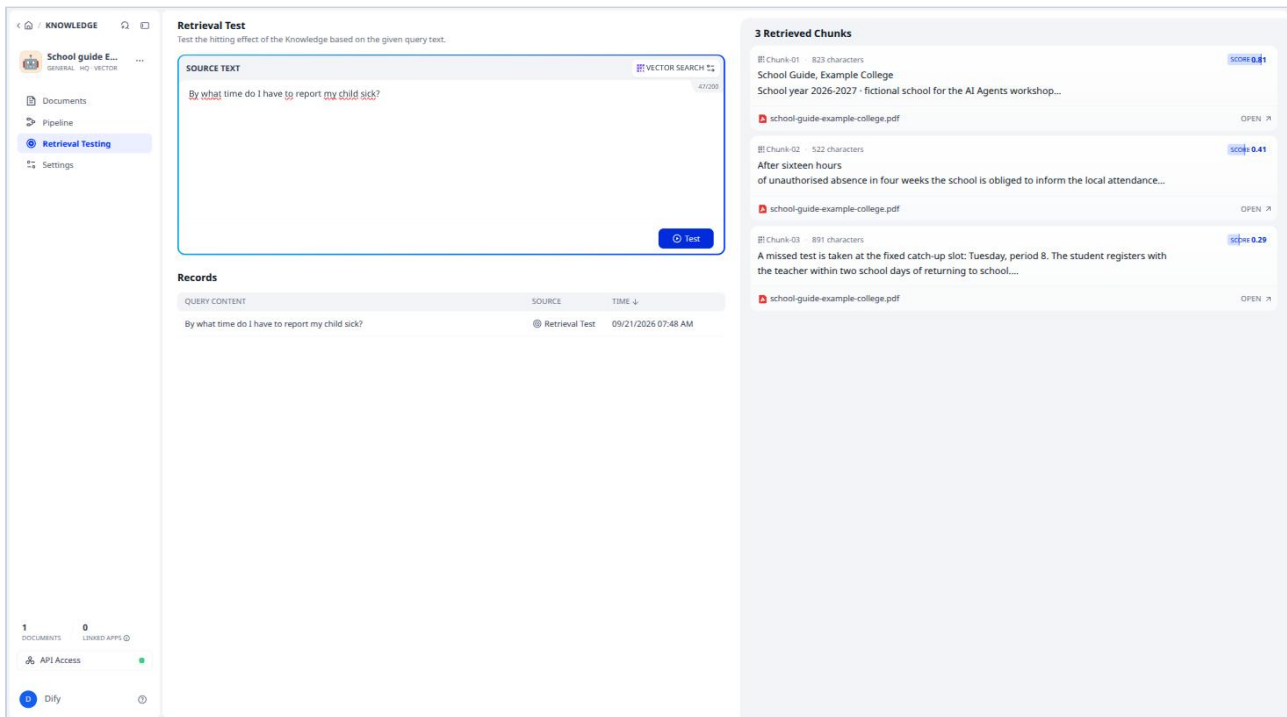
Retrieval Testing, before the first query

2. Click in the **SOURCE TEXT** field and type a question.

TYPE IN

By what time do I have to report my child sick?

3. Click **Test**.



Three retrieved pieces with a score

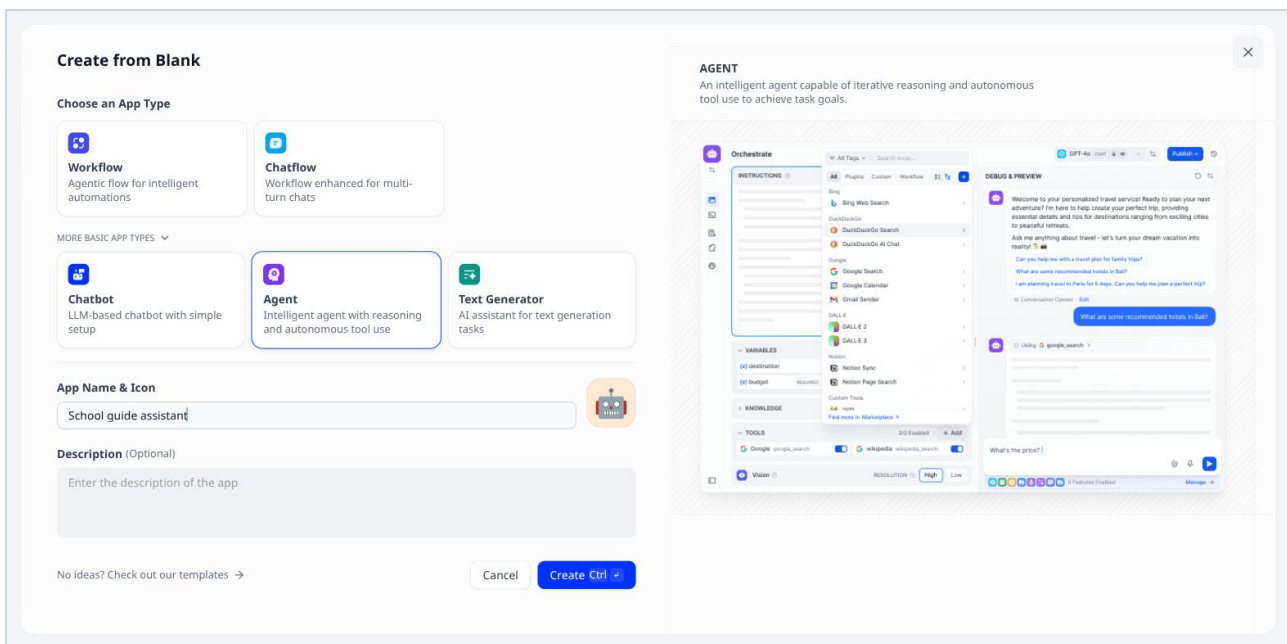
On the right are the retrieved pieces with a score. The piece with the sick-report rules is at the top with 0.81; the other two score low. That is how it should be. If the right piece is missing, go back to step 2 and check whether the pieces are large enough.

Part B. The agent

Step 6. Create the agent

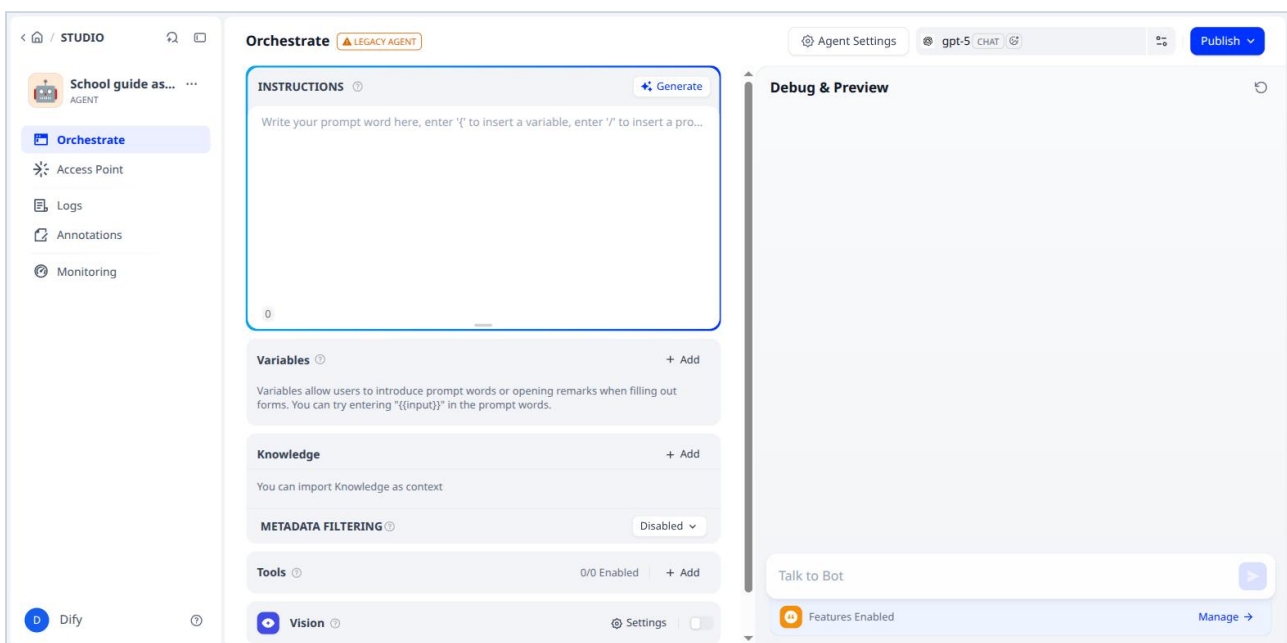
1. Click **Studio** in the left-hand menu.
2. Click **Create** at the top right and choose **Create from Blank**.
3. Click **MORE BASIC APP TYPES**.
4. Click **Agent**.
5. Fill in the name.

Field on screen	What you type
App Name & Icon	School guide assistant



Agent chosen, name filled in

6. Click **Create**.



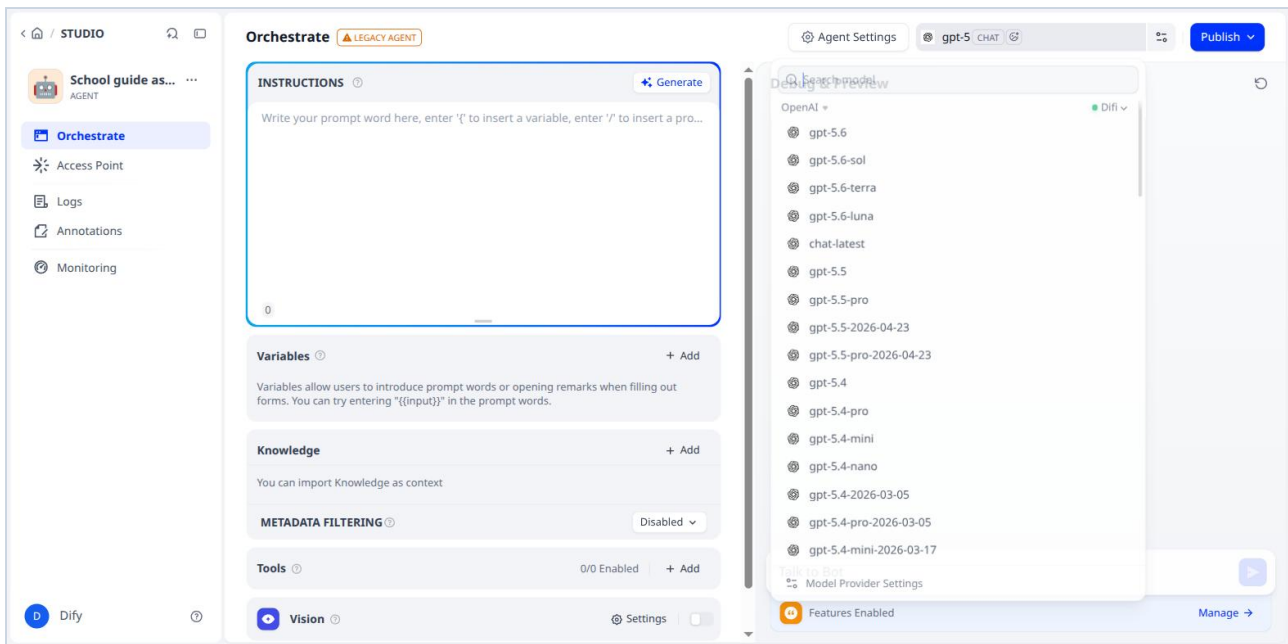
The empty agent

Next to "Orchestrate" there is an orange label **LEGACY AGENT**. Do not be alarmed; the agent works fine. What it means is explained in the box "About the LEGACY AGENT label" at the end of this guide.

Step 7. Set the model to gpt-5-mini

Here too the model defaults to **gpt-5**, and here too that is five times more expensive than necessary.

1. Click the model name **gpt-5** at the top right.



The model picker

2. Choose **gpt-5-mini**.

Step 8. Write the instruction

In a workflow you wrote a prompt per step. An agent gets one instruction: who it is, what it may do and what it may not do. The rules about “only from the knowledge base” and “say so when you do not know” are the most important; without them it happily makes things up.

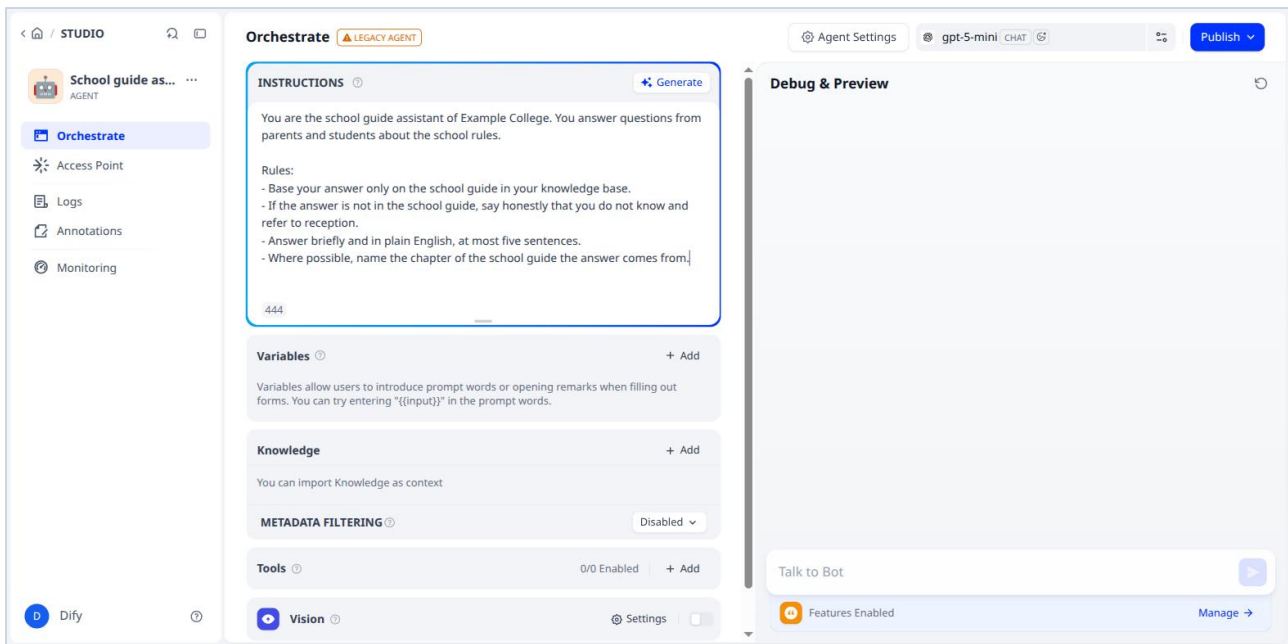
1. Click in the **INSTRUCTIONS** field.
2. Type the instruction.

TYPE IN

You are the school guide assistant of Example College. You answer questions from parents and students about the school rules.

Rules:

- Base your answer only on the school guide in your knowledge base.
- If the answer is not in the school guide, say honestly that you do not know and refer to reception.
- Answer briefly and in plain English, at most five sentences.
- Where possible, name the chapter of the school guide the answer comes from.

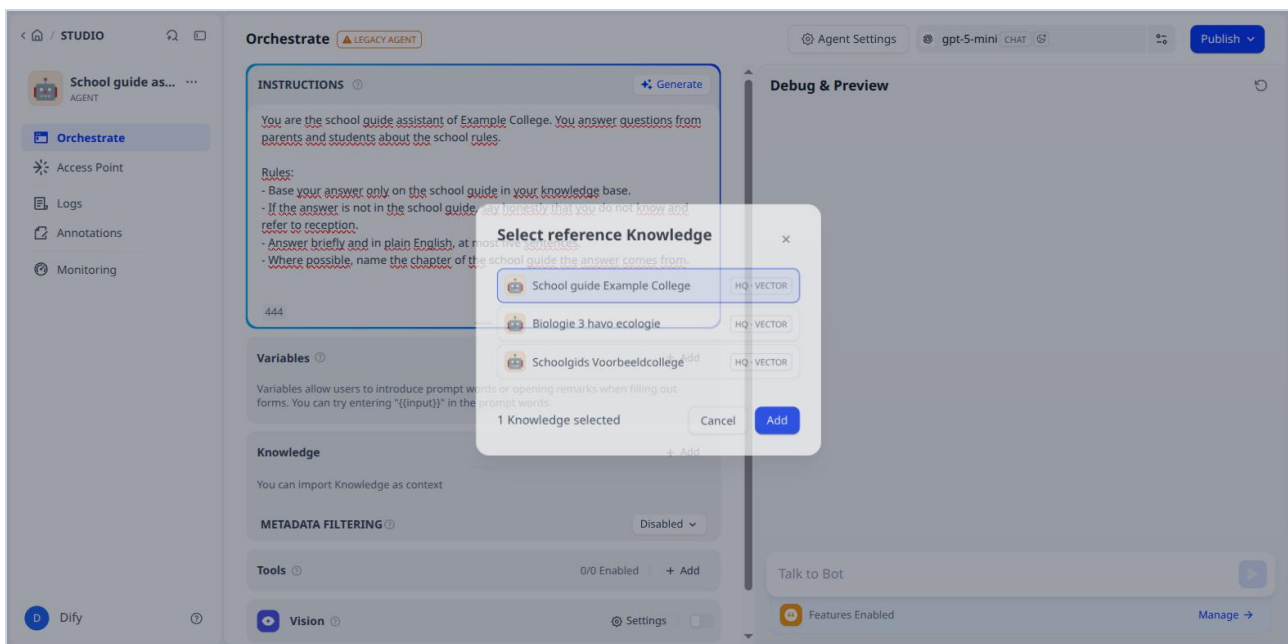


Model and instruction in place

Your choice: role, rules and tone. The first sentence sets the role ("school guide assistant", "exam helper", "onboarding buddy for new colleagues"), the rules set the behaviour. Always keep the two rules about the knowledge base and about "say that you do not know"; change the rest. Want the agent to address students informally and parents formally? Say so. Want it to always end with "Is that right, or do you want to know more?" Say so. Want answers in another language when the question is in that language? One more line.

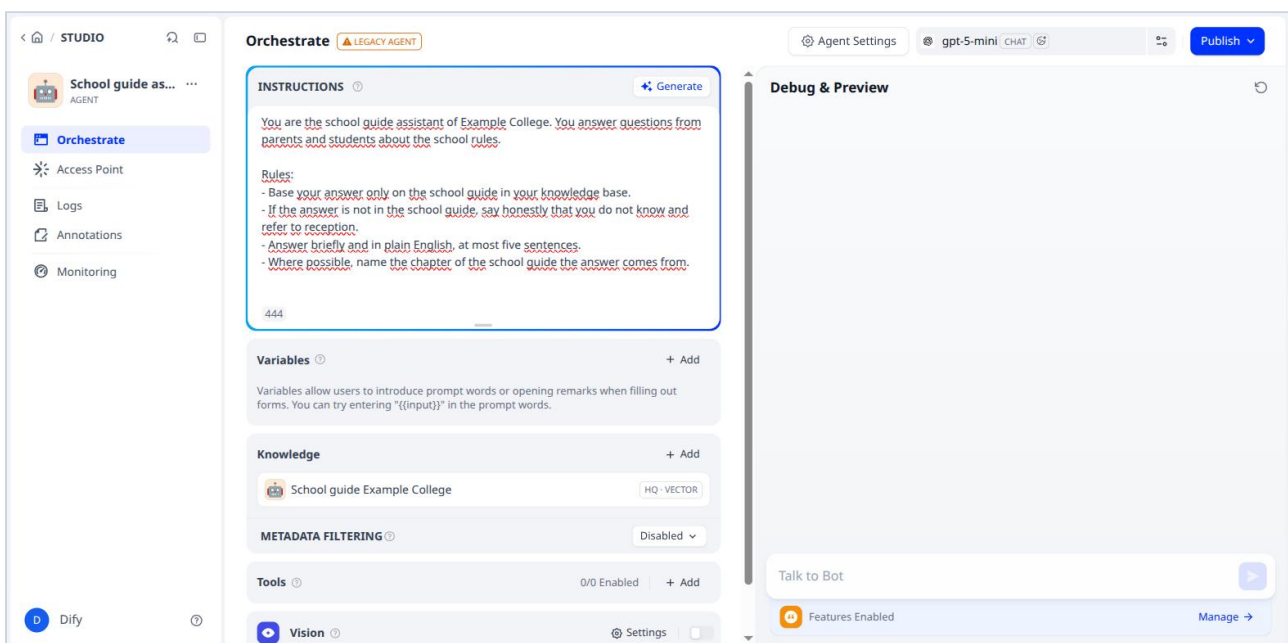
Step 9. Attach the knowledge base

1. Under **Knowledge**, click **Add**.
2. Click **School guide Example College**. At the bottom "1 Knowledge selected" appears.



Choosing the knowledge base

3. Click **Add**.



The knowledge base is attached to the agent

Step 10. Test the agent

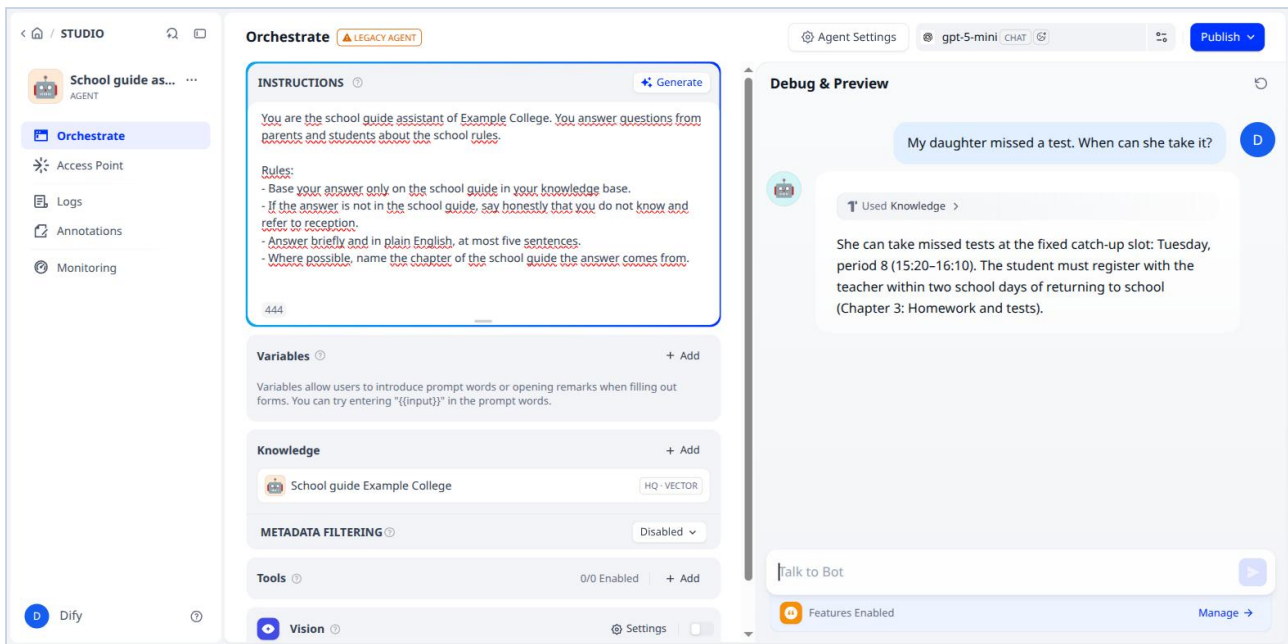
On the right is **Debug & Preview**: a chat window in which you talk to the agent before it goes live. Ask two kinds of question: one that is in the document, and one that is not.

1. Click **Talk to Bot** at the bottom and type the first question.

TYPE IN

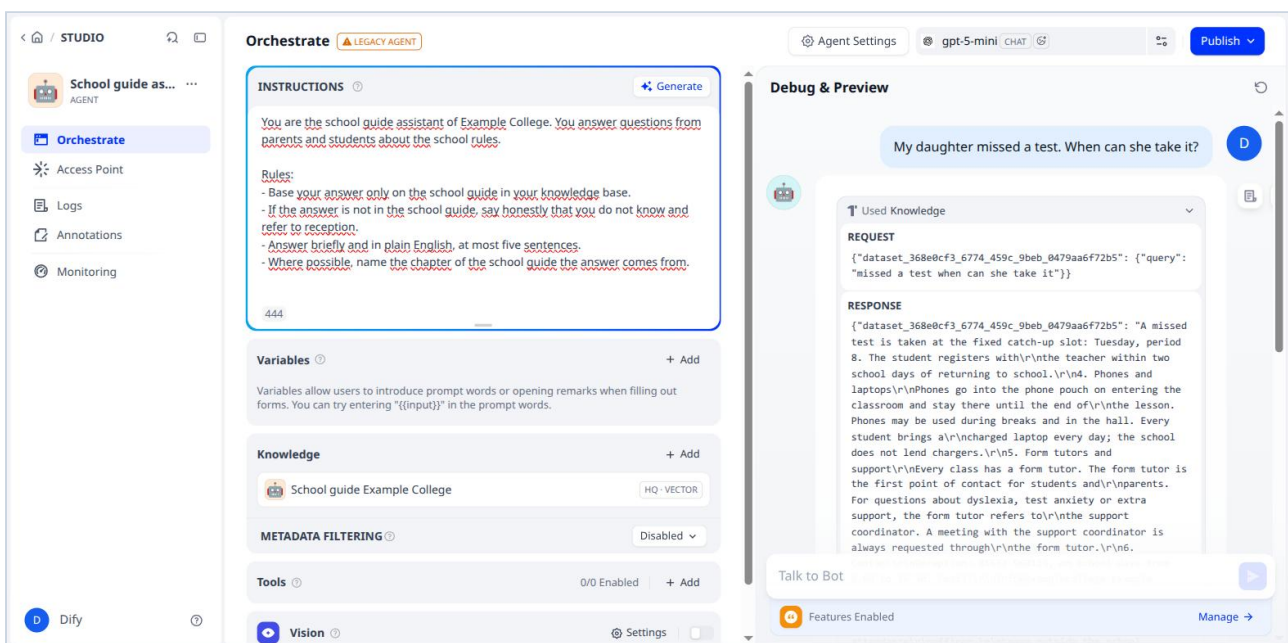
My daughter missed a test. When can she take it?

2. Press Enter and wait. Above the answer the label **Used Knowledge** appears.



The first answer, with a chapter reference

3. Click **Used Knowledge** to expand it.



What the agent looked up by itself

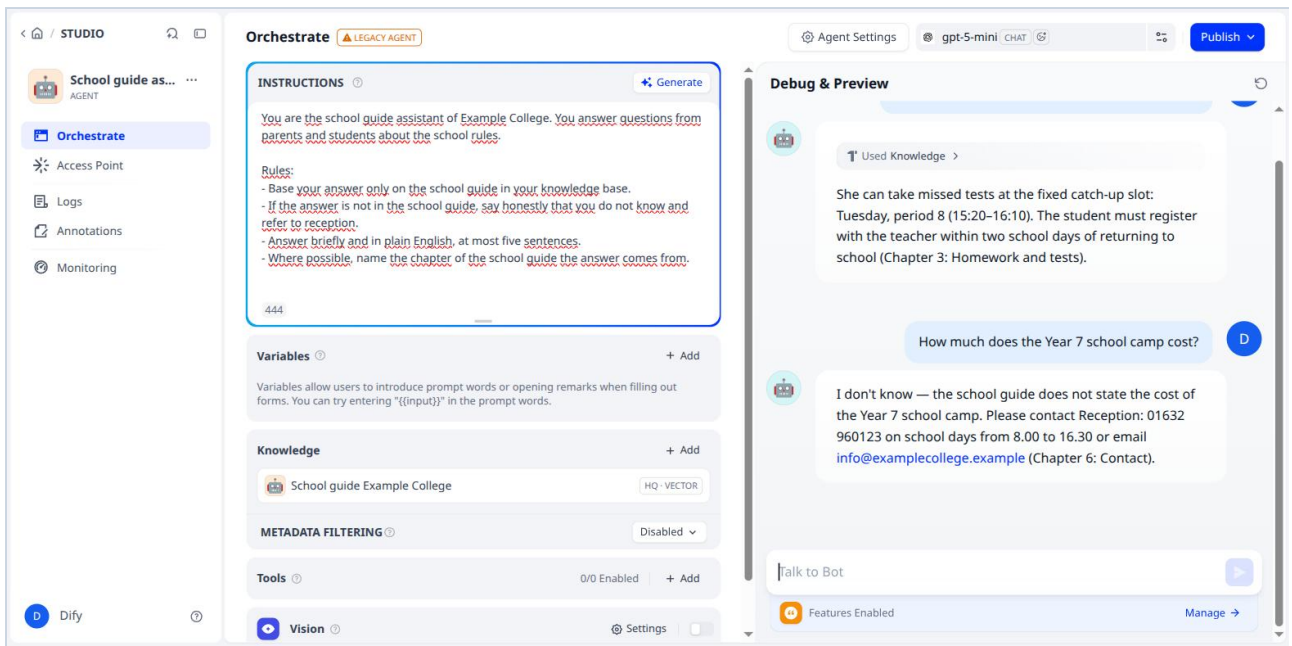
Look at REQUEST. That is the search query the agent formulated by itself, something like "missed test catch up when can she take it". You did not tell it to search, nor what to search for. It decided that itself, based on your question and the description of the knowledge base. That is the difference from the workflow in guide 1.

4. Collapse the label again and type the second question.

TYPE IN

How much does the Year 7 school camp cost?

5. Press Enter.

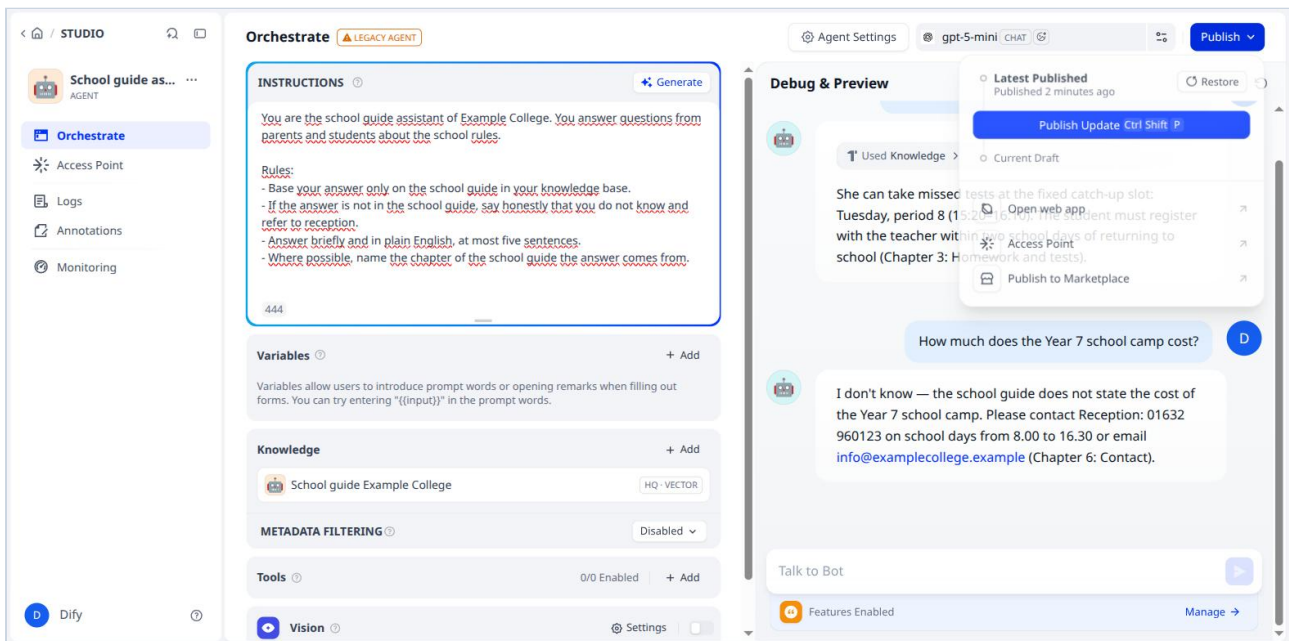


The question that is not in the document

The agent says it is not in the school guide and refers to reception. That is because of the second rule in your instruction. Remove that rule and try again: you get a neatly invented amount.

Step 11. Publish the agent

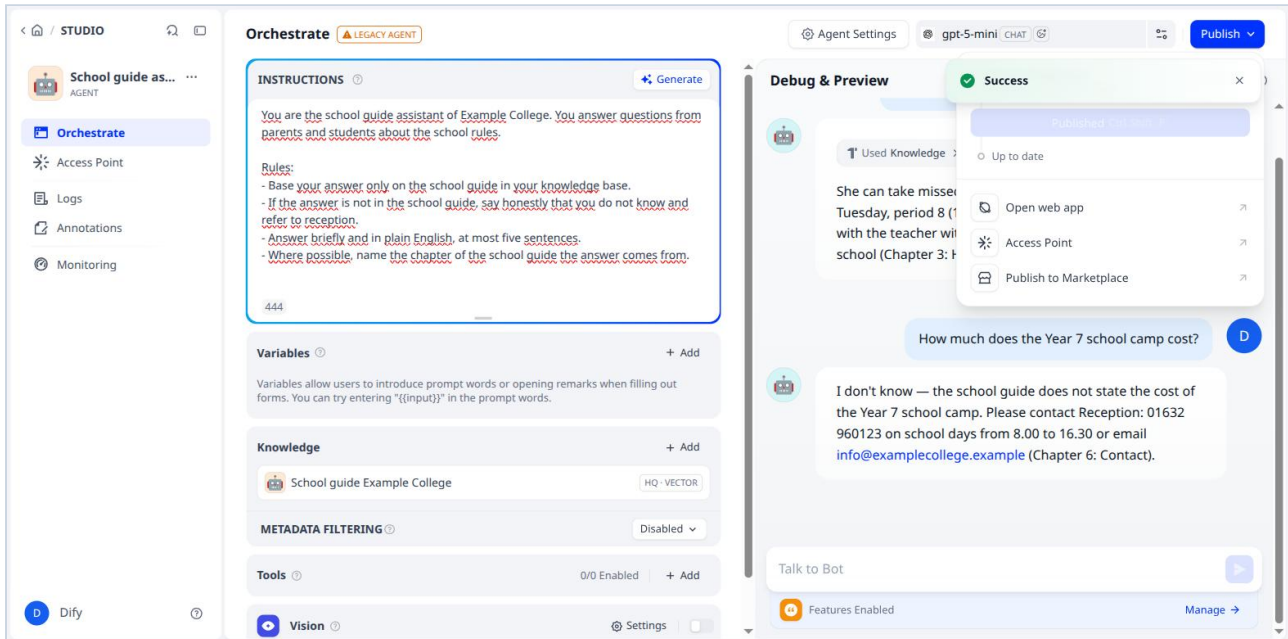
1. Click **Publish** at the top right.



The Publish menu

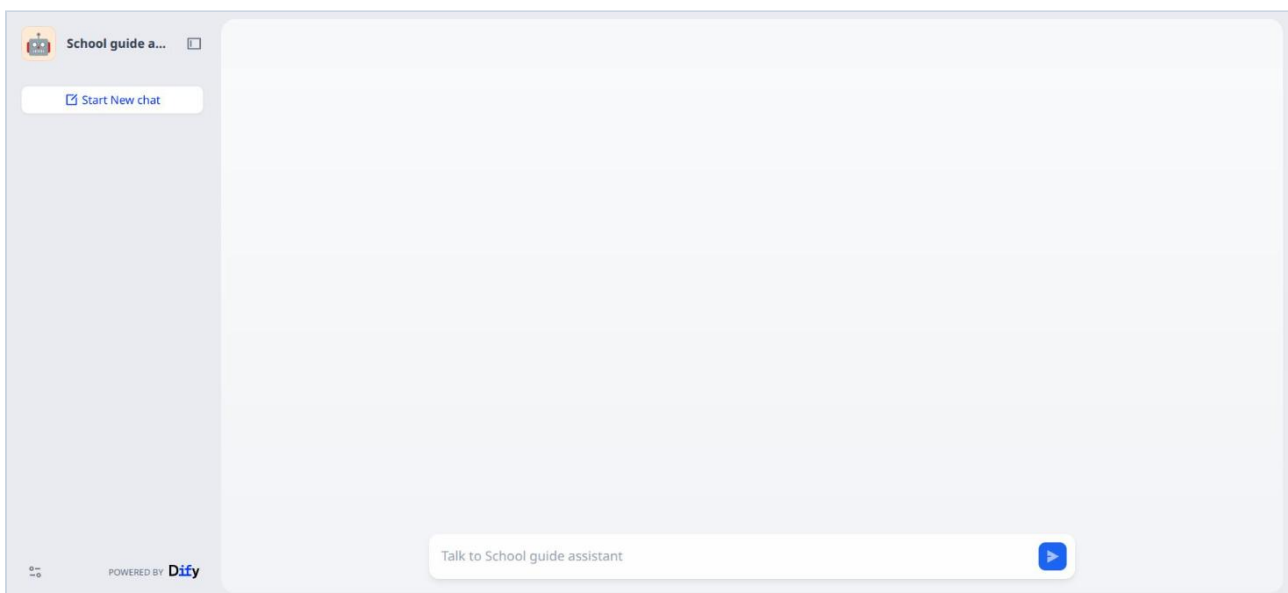
The menu says “Latest Published, published a few minutes ago”, even though you have not published anything yet. Dify publishes an Agent app when it is created, so empty. The version with your instruction and knowledge base is under **Current Draft**. If you skip Publish Update, the webapp opens an agent without knowledge.

2. Click **Publish Update**. The message “Success” appears and below the button it now says “Up-to-date”.



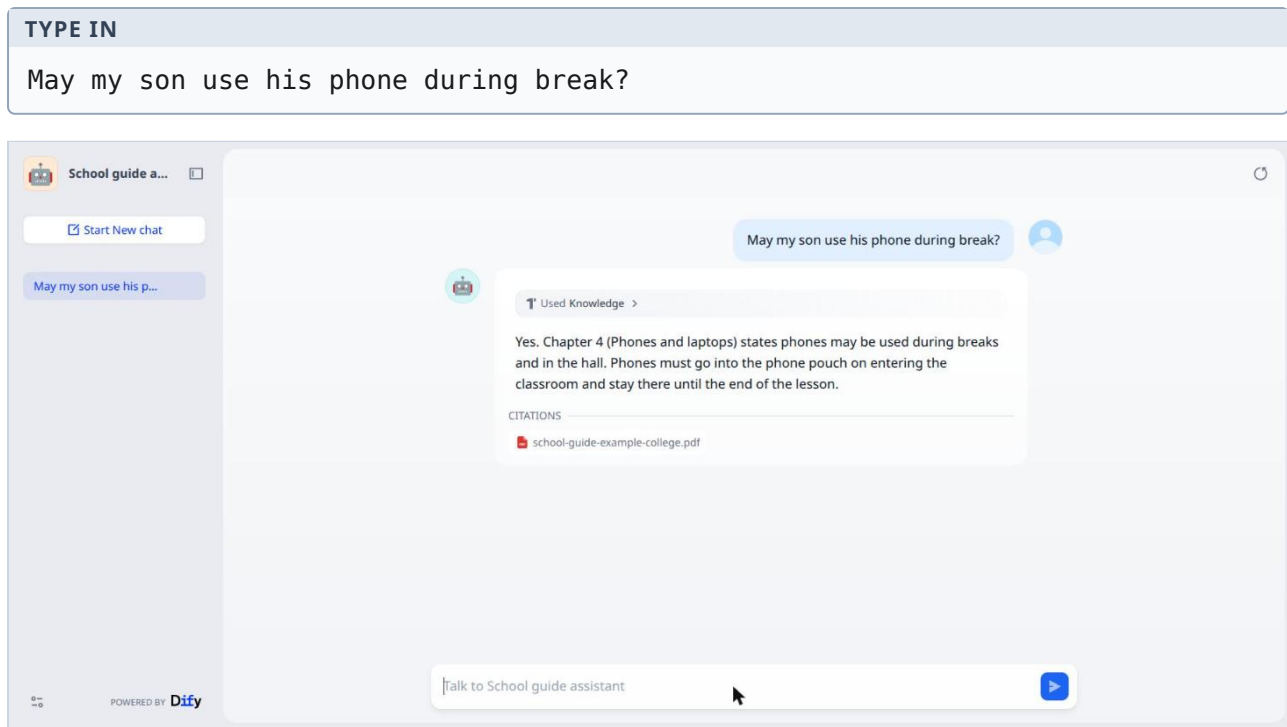
Published and up-to-date

3. Click **Open webapp**.



The webapp, empty

4. Type a question to check that the published version uses the knowledge base.



The webapp answers with a source

Below the answer is **CITATIONS** with the name of your document. That is the link you give to your test partner; it works without logging in, on a phone too.

The agent in this guide is live at <https://udify.app/chat/6Y6WnP0m5Z3jSHRs>. Open it to see what you are working towards.

Checking that it worked

- Under **Knowledge** your knowledge base has status Available and a preview of at most a handful of pieces per page
- **Retrieval Testing** puts the right piece at the top for a question from the document
- The agent answers a question from the document with **Used Knowledge** above it
- The agent says it does not know for a question outside the document
- The webapp link shows **CITATIONS** below the answer

When things go wrong

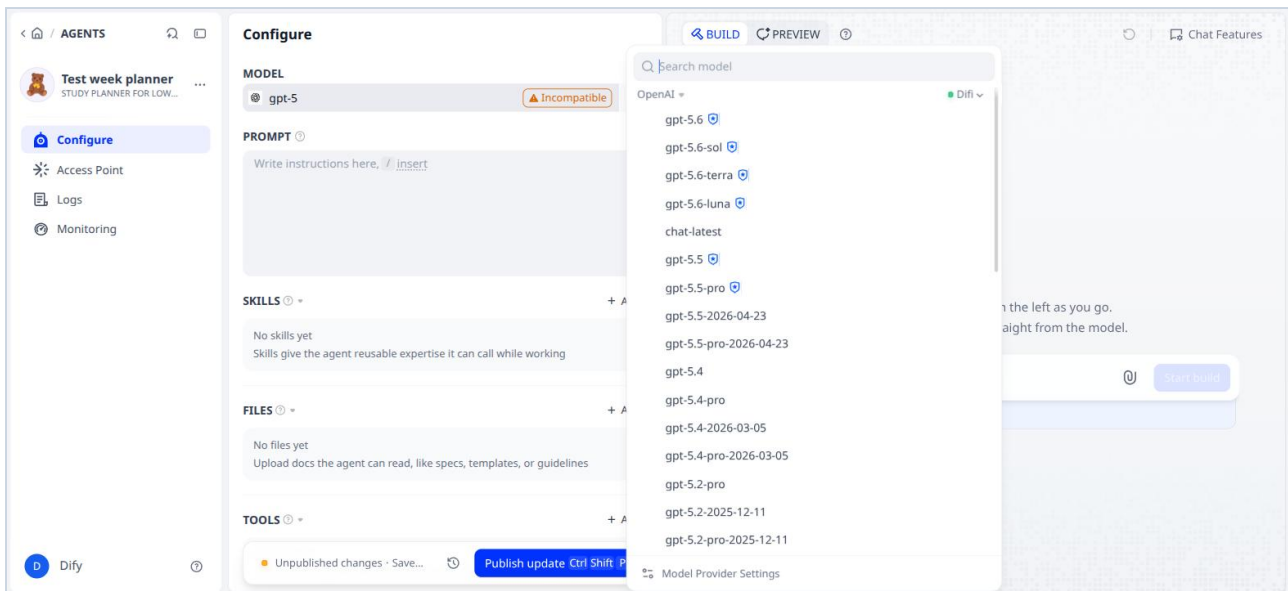
What you see	What is going on
Preview Chunk shows dozens of one-line pieces	The Delimiter is <code>\n</code> . Set it to <code>\n\n</code> (step 2)
Retrieval Testing does not find the right piece	Pieces too small or too large. Back to step 2; for a large document raise Maximum chunk length
The agent answers without Used Knowledge, even for questions from the document	Knowledge base not attached (step 9), or the Description of the knowledge base is empty or meaningless (step 4)
The agent invents answers	The rule “If the answer is not in the school guide...” is missing from the instruction
The webapp answers without knowledge while Debug & Preview does use it	Publish Update not clicked; the webapp is still running the empty version from creation
Your credits drop fast	The model is still on gpt-5. Set it to gpt-5-mini (step 7). An agent answer costs 6 to 7 credits anyway
On upload: “Maximum 1 files”	Sandbox limit. One file at a time; add a second one afterwards with Add file
A message that the temperature parameter was removed	The model does not support it. Ignore it

About the LEGACY AGENT label

Dify is currently building a new way of making agents: the **Agent Console**, found in the left-hand menu as **Agents** with the label BETA. The Agent app you built in this guide is therefore called “legacy”: Dify wants you to take the new route eventually.

We looked at that new console on 20 September 2026 and deliberately did not choose it for this guide:

- gpt-5-mini is not available there; gpt-5.6-luna, Claude Opus 4 and Gemini 3.5 are, all more expensive in credits;
- there is no link to a knowledge base; the agent reads loose files in a sandbox environment;
- the page does not render stably yet and the texts switch between languages.



The model choice in the new Agent Console

For a workshop with sandbox accounts the Agent app in Studio is the reliable route. What you learn in it, instruction, knowledge base, testing and publishing, still applies in the new console; only the buttons move. Guides 4 and 6 show the Agent Console in action.

Make it yours

The school guide is invented and the assistant is an example. This chapter helps you rebuild the agent for your document, your audience and your tone. From easy to hard.

Change	Where	Difficulty	What you get for it
A different document	Part A, steps 1 to 4	easy	An agent about something people at your school really ask about
Name and description of the knowledge base	Step 4	easy	An agent that knows when to search
Role, tone and form of address	Step 8, instruction	easy	An assistant that sounds like your school
An opening line and example questions in the webapp	Debug & Preview → Manage (at the bottom) → Conversation Opener	easy	Users immediately know what they can ask
A second document in the same knowledge base	Part A again, via Add file	medium	More questions answered well; mind the description
Two knowledge bases (for example students and staff separately)	Step 9, Add twice	medium	A clear separation; in the instruction you say when to use which
Enforcing source references (“always name the chapter”)	Step 8, instruction	medium	Answers you can check
Adding tools (date, calculation, web search)	Tools on the same screen	hard	An agent that can do more than read; see guide 4

Three assignments

1. **Your document, your questions.** Make a knowledge base from a document of your school. Write down five questions users really ask beforehand, plus one question you know is not answered in it. Test all six in step 10. If the sixth gets an invented answer, your instruction is not strict enough yet.
2. **Two audiences, one agent.** Make the agent answer differently for students (informal, short) and parents (formal, a bit fuller). Solve it in the instruction first (“If the question seems to come from a parent...”). If that does not work reliably, add a **Variable** (the Variables block under the instruction) so the user chooses.
3. **Let a colleague break it.** Give the link to someone who knows the document well and ask them to catch the agent making mistakes. Read the **Logs**. Is it the splitting (step 2), the description (step 4) or the instruction (step 8)? Change one thing, Publish Update, and test again.

Checklist before you use your own material

- [] The document contains no names, addresses, grades or other personal data

- [] The document belongs to the school and may sit with an external processor (Dify runs in the cloud)
- [] The document is the current version, and there is one owner who uploads a new version when it changes
- [] You have decided who gets the link; for students or parents, involve the data protection officer first
- [] The model is set to gpt-5-mini and you know an answer costs 6 to 7 credits without your own key
- [] The instruction says what the agent must do when it does not know, and you have tested that

What next

Your agent can now read. The next step is that it can do something: search the web, calculate, know the date, or call a workflow from guide 1 as a tool. You arrange that under **Tools** on the same screen. That is guide 3.

Licence: CC BY-SA 4.0 (<https://creativecommons.org/licenses/by-sa/4.0/>). You may copy, adapt and reuse this document, including at your own school, provided you credit the source (Guido van Dijk, LeX Consultancy B.V., <https://git.lexconsultancy.nl/guido/dify-guides>) and share your adaptation under the same licence. The LeX Consultancy logo is excluded; product names are trademarks of their owners. Schools, people and documents in the examples are fictional.

Owner and contact: Guido van Dijk, LeX Consultancy B.V. · guido@l-e-x.nl



LeX Consultancy B.V. · l-e-x.nl · Building in Dify: an agent with its own knowledge base