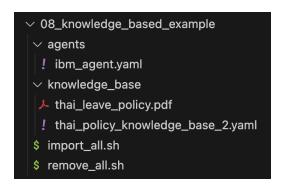
# Lab 8 - Knowledge based RAG agent

In this lab we will dive deeper in to building agent with Watsonx Orchestrate ADK. For this lap please **make sure that you have Python version 3.11 installed**. If not, please follow the <u>link</u> to install python 3.11 with brew.

1. Clone the following git repo

git clone https://github.com/edsml-kl121/watsonx-orchestrate-adk-tutorial.git

2. cd into the 08\_knowledge\_based\_example folder. Here is the structure of the folder



3. create virtual environment using the following command. Make sure to use python 3.11 to create the environment and source to the .venv

python3.11 -m venv .venv source .venv/bin/activate

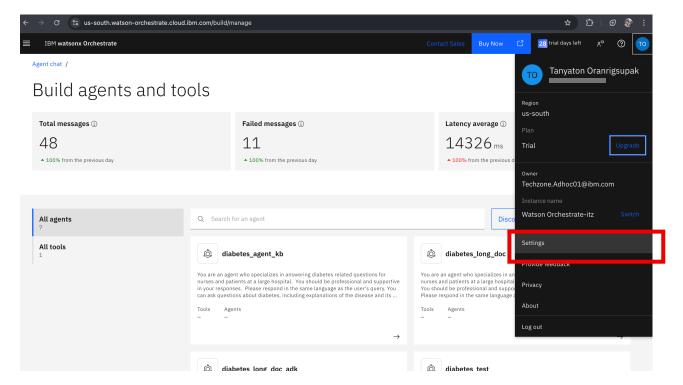
4. install the requirement library

pip install ibm-watsonx-orchestrate

To check if the library is installed properly, you could try command **orchestrate**. The following text should pop up, indicates that Watsonx orchestrate ADK is installed

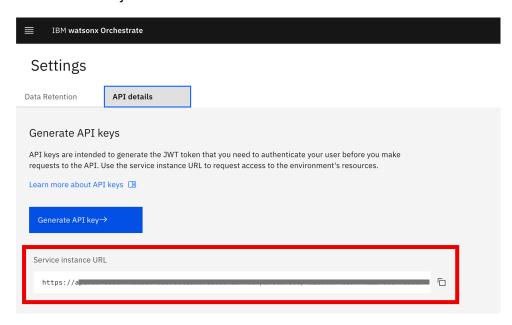
```
(.venv) → 08_knowledge_based_example git:(main) × orchestrate
 Usage: orchestrate [OPTIONS] COMMAND [ARGS]...
  Options
  --version
                                Show the installed version of the ADK and Developer Edition Tags
                                Enable debug mode
  --debug
  --install-completion
                                Install completion for the current shell.
  --show-completion
                                Show completion for the current shell, to copy it or customize the installation.
  --help
                                Show this message and exit.
  - Commands -
                    Add, remove, or select the activate env other commands will interact with (either your local server or a production
  env
                    instance)
                    Interact with the agents in your active env
  agents
  tools
                    Interact with the tools in your active env
  toolkits
                    Interact with the toolkits in your active env
  knowledge-bases
                   Upload knowledge your agents can search through to your active env
  connections
                    Interact with the agents in your active env
  server
                    Manipulate your local Orchestrate Developer Edition server [requires entitlement]
                    Launch the chat ui for your local Developer Edition server [requires entitlement]
  models
                    List the available large language models (llms) that can be used in your agent definitions
  channels
                    Configure channels where your agent can exist on (such as embedded webchat)
  evaluations
                    Evaluate the performance of your agents in your active env
  copilot
                    Access AI powered assistance to help refine your agents
  settings
                    Configure the settings for your active env
```

- 5. To move on to the next step, we will need to obtain your API KEY and URL. You can find your API KEY and URL on the Watsonx Orchestrate webpage
  - 5.1 Go to the **profile** on the top right corner > **setting**



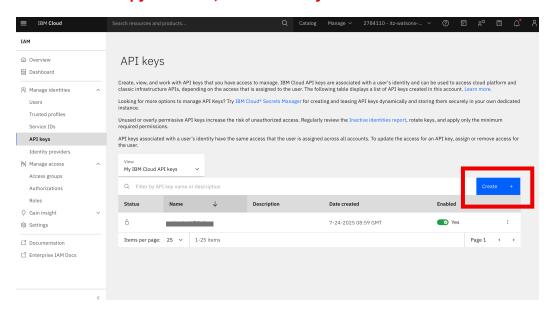
#### 5.2 Go to API details

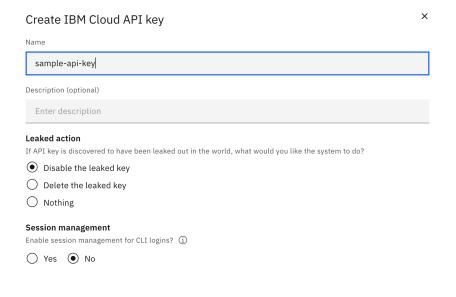
# 5.3 Note down your "Service instance URL"



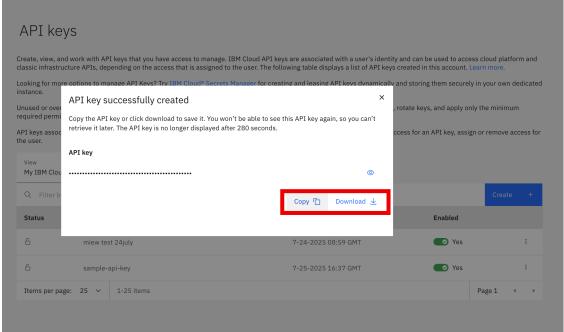
- 5.4 Click on "Generate API keys". This will redirect you to a new webpage
- 5.6 Click on Create > add the Name for you key > Click Create

# Make sure to copy and save, or download your API KEY and URL









Now you should already have URL and API KEY saved in your local device:

- 6. Add new environment to watsonx Orchestrate
  - 6.1 Run the following command on your terminal at the **08\_knowledge\_based\_example** folder

```
orchestrate env add -n .venv -u <YOUR_URL> orchestrate env activate
```

put in your URL you have saved from the previous step in the first command. After running the second command, you will be asked to put in your API\_KEY. Please do so using the API KEY you have saved.

#### Example:

```
(.venv) → 08_knowledge_based_example git:(main) × orchestrate env add -n .venv -u https://api.dl.watson-orchestrate
[INFO] - Environment 'new-venv' has been created
(.venv) → 08_knowledge_based_example git:(main) × orchestrate env activate .venv

Please enter WXO API key:
[WARNING] - Using MCSP Auth Type. If this is incorrect please use the '--type' flag to explicitly choose one of 'ibm_iam', 'mcsp_v1', 'mcsp_v2' or 'cpd'
[INFO] - Environment '.venv' is now active
```

6.2 Add and activated the environment, you can check the list of orchestrate environment with the following command. The .venv environment should have the activate mark on the right

orchestrate env list

### Example:

```
(.venv) → 08_knowledge_based_example git:(main) × orchestrate env list

.venv    https://api.dl.watson-orchestrate.ibm.com/instances/20250515-0617-4459-503c-8c686a2c306d (active)
local    http://localhost:4321
miew-env    https://api.eu-central-1.dl.watson-orchestrate.ibm.com/instances/20250616-0240-4655-107d-25a33ce99b20
new-venv    https://api.us-south.watson-orchestrate.cloud.ibm.com/instances/4a1f0aa7-fa04-41af-86f4-8b8fcbaa91b8
```

- 7. We are now going to **create a RAG agent** using the following code in the folder
  - 7.1 understanding file structure

```
08_knowledge_based_example
agents
ibm_agent.yaml
knowledge_base
thai_leave_policy.pdf
thai_policy_knowledge_base_2.yaml
import_all.sh
remove_all.sh
```

#### ! ibm\_agent.ymal

ymal file used to specify <u>agent</u> name, description, *llm model used*, style, and *knowledge base* source

## ! thai\_policy\_knowledge\_base\_2.ymal

ymal file used to specify <u>agent's knowledge</u> name, description, source document, and embedding model

# ♦ Thai\_leave\_policy.pdf

PDF document contain information for agent knowledge

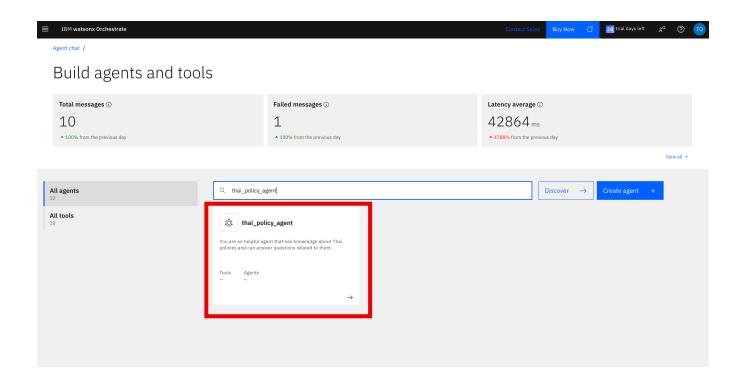
You can discover more by going into each file

To create this agent on watsonx orchestrate, there are a list of command you need to execute. That list of commands is provided in the file **import\_all.sh** 

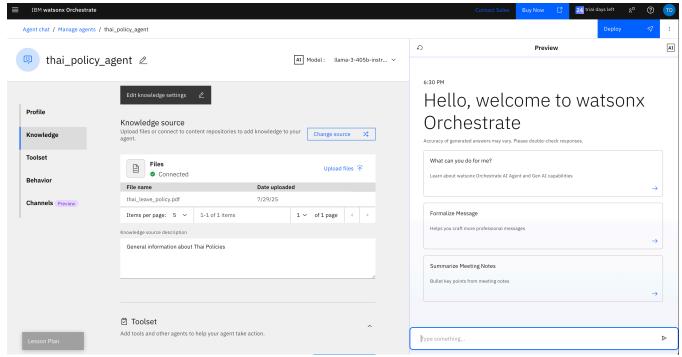
7.2 Run the file using the following command

bash import\_all.sh

8. After running the command successfully, you should be able to see the agent on the watsonx Orchestrate UI. Click on the agent



As you can see, the information specified in the ymal file (eg. Name, description, knowledge) is now in the agent detail



9.	The agent can now	be tested with the	following prompt

- พนักงานสามารถลาชั่วคราวได้กี่วันต่อปี
- การลาคลอดได้รับค่าจ้างกี่สัปดาห์
- พนักงานชายมีสิทธิ์ลาพ่อได้กี่วัน
- สามารถนำการลาไปปีถัดไปได้สูงสุดกี่วัน
- การลาโคยไม่ได้รับค่าจ้างสามารถลาได้สูงสุดกี่เดือน



thai\_policy\_agent 06:39 PM | Show Reasoning 

wนักงานหญิงมีสิทธิได้รับการลาคลอดพร้อมกับค่าจ้างเต็มจำนวน
สำหรับระยะเวลา 26 สัปดาห์ติดต่อกัน (ไม่รวมวันหยุดนักขัตฤกษ์)

thai\_leave\_policy.pdf

HRMS ได้รับ ได้รับล่วงหน้า 15
วัน หลังจากยี ่นคำร้องแล้ว การแจ้ง
เตือนจะถูกส่งไปยังผู ่อการลาได้รับ
การอนุมัติ การแจ้งเตือนจะถูกส่งไป
ยังพนักงานและแผนกทรัพยากร...

View source

You 06:39 PM