

SRE Agent – Autonomous Incident Remediation

Learning Objective

You will gain hands-on knowledge and practical skills to design, build, and deploy intelligent AI agents tailored for Site Reliability Engineering (SRE) workflows in the insurance domain using Watsonx Orchestrate ADK and Watsonx Orchestrate. Specifically, you will be able to:

- Develop custom AI agents using Watsonx Orchestrate ADK that can autonomously monitor system health, analyze logs and metrics, and trigger predefined remediation playbooks for core insurance services such as policy management, claims processing, and customer portals.
- Orchestrate end-to-end incident management workflows using Watsonx Orchestrate, integrating multiple tools, APIs, and observability platforms to streamline root cause analysis and reduce resolution time.
- Apply these skills to real-world SRE challenges, improving system reliability, minimizing downtime, and enhancing customer satisfaction through proactive and intelligent incident response in the insurance domain.

The Problem

MetLife Insurance is striving to improve service reliability and enhance customer satisfaction by modernizing the operations behind its core insurance services—including policy management, claims processing, payments, and the customer portal.

However, the Site Reliability Engineering (SRE) team currently faces significant challenges in maintaining seamless service delivery and minimizing downtime during system incidents.

The incident management process is predominantly manual, requiring SREs to monitor disparate system dashboards, interpret performance metrics, analyze logs, and identify root causes before applying remediation. When services degrade or fail, it often leads to delays in detection and resolution, affecting both internal workflows and end-user experience.

Moreover, the lack of intelligent automation in incident response extends resolution times, introduces operational inefficiencies, and increases the risk of prolonged service disruptions. This reactive approach limits the team's ability to proactively prevent issues or scale operations effectively.

To summarize, some of the key challenges faced by SREs at MetLife Insurance include:

- Manual and fragmented processes slow down the detection, diagnosis, and resolution of service incidents.
- Reliance on human expertise to interpret logs and metrics delays response time during critical outages.

- Lack of integrated tools hinders rapid identification of root causes and effective execution of fixes.
- Extended service downtime negatively impacts customer experience and trust in digital insurance services.

An intelligent, autonomous SRE AI Agent equipped with advanced observability, automated root cause analysis, and pre-defined remediation playbooks could dramatically improve system resilience—enabling faster, more reliable, and scalable operations while reducing the burden on human engineers.

Objective

MetLife Insurance plans to implement an AI-powered SRE Agent Assistant to support their Site Reliability Engineering (SRE) team in maintaining service reliability and minimizing system downtime. The goal is to create an AI-powered agentic solution that assists SREs in executing the following tasks:

- **Check service health**
Check the health status of a service in the system.
Queries the service registry to retrieve current health metrics and determine if a service is functioning accordingly. This helps verify service availability, troubleshoot incidents, validate deployments, and assess the health of upstream/downstream dependencies.
- **Restart services**
Perform a service restart to recover from degraded or unhealthy states.
Initiates a restart of the specified service—commonly used as a remediation action during incident response when a service is unhealthy.
- **View incident history**
Retrieve historical incident data for pattern analysis and reporting.
Fetches past incidents to identify recurring issues, support root cause analysis, and improve incident response over time.
- **Find unhealthy services**
List services that are degraded or unhealthy.
Enables quick identification of problematic services to prioritize investigation and resolution.
- **List all services**
Return the current health status of all services.
Provides a unified view of overall system health, aiding in monitoring, incident management, and proactive maintenance.

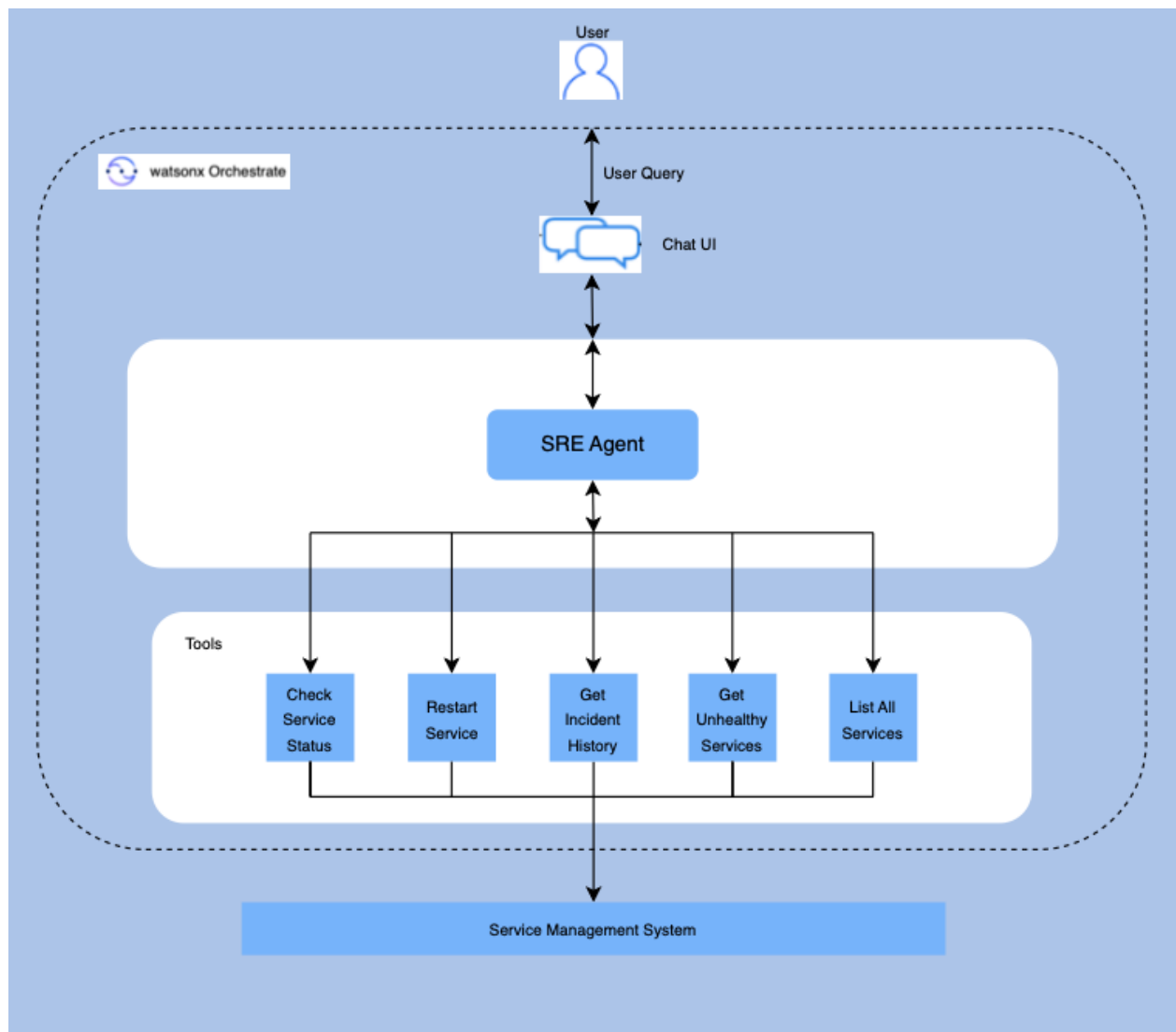
By automating these tasks, MetLife aims to reduce operational overhead, accelerate incident response, and improve system resilience—ultimately enhancing customer experience and trust in their digital insurance services at scale.

Business Value

Accelerating Incident Resolution & Reinvesting Human Capital

- An SRE AI agent that autonomously troubleshoots and remediates issues can dramatically reduce Mean Time to Detect (MTTD), Mean Time to Acknowledge (MTTA), and Mean Time to Resolve (MTTR).
- By leveraging intelligent automation and contextual reasoning, the agent identifies root causes faster than manual processes, minimizing downtime and improving service reliability.
- Reducing MTTR translates directly into cost savings and improved customer experience. Every minute of reduced outage time prevents revenue loss and protects brand reputation.
- By offloading repetitive troubleshooting tasks to an autonomous agent, SRE teams reclaim valuable time.
- Instead of firefighting, engineers can focus on strategic initiatives such as capacity planning, performance optimization, and building reliability features into the architecture.

Architecture



Pre-requisites

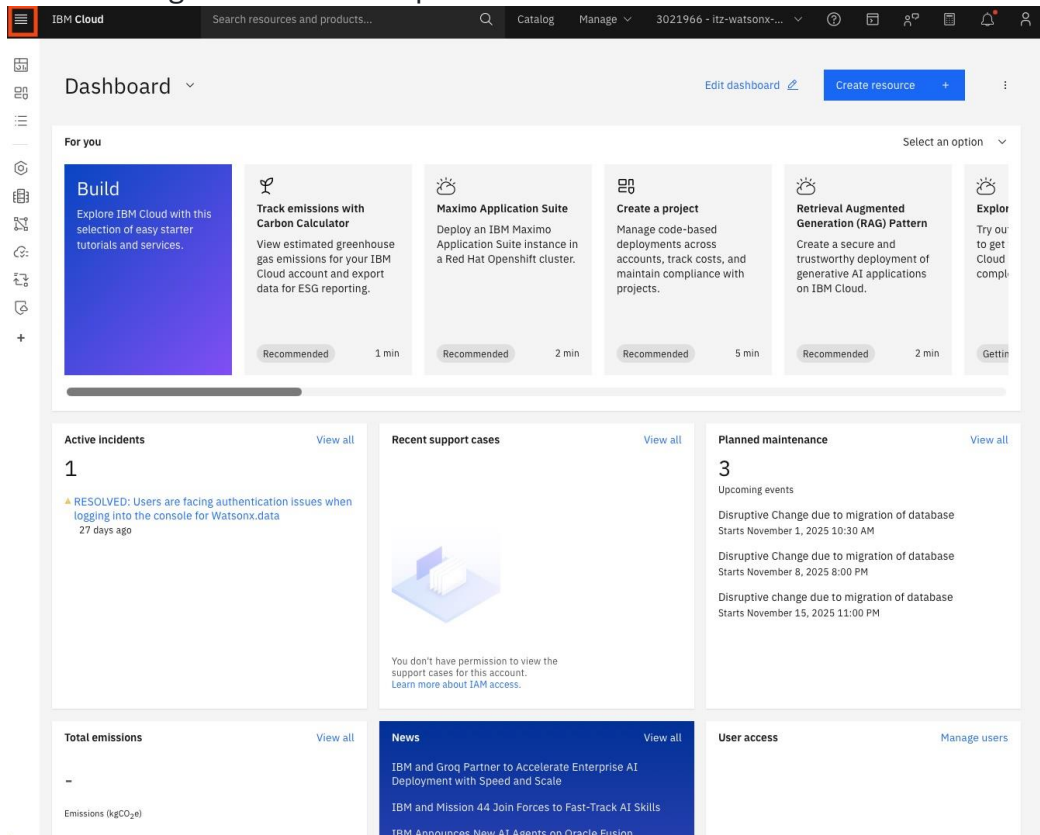
- Check with your instructor to make sure all systems are up and running before you continue.
- Validate that you have access to the right techzone environment for this lab.
- Validate that you have access to a credentials file that you instructor will share with you before starting the labs.

Accessing Your Watsonx Studio and adding the Python Notebook

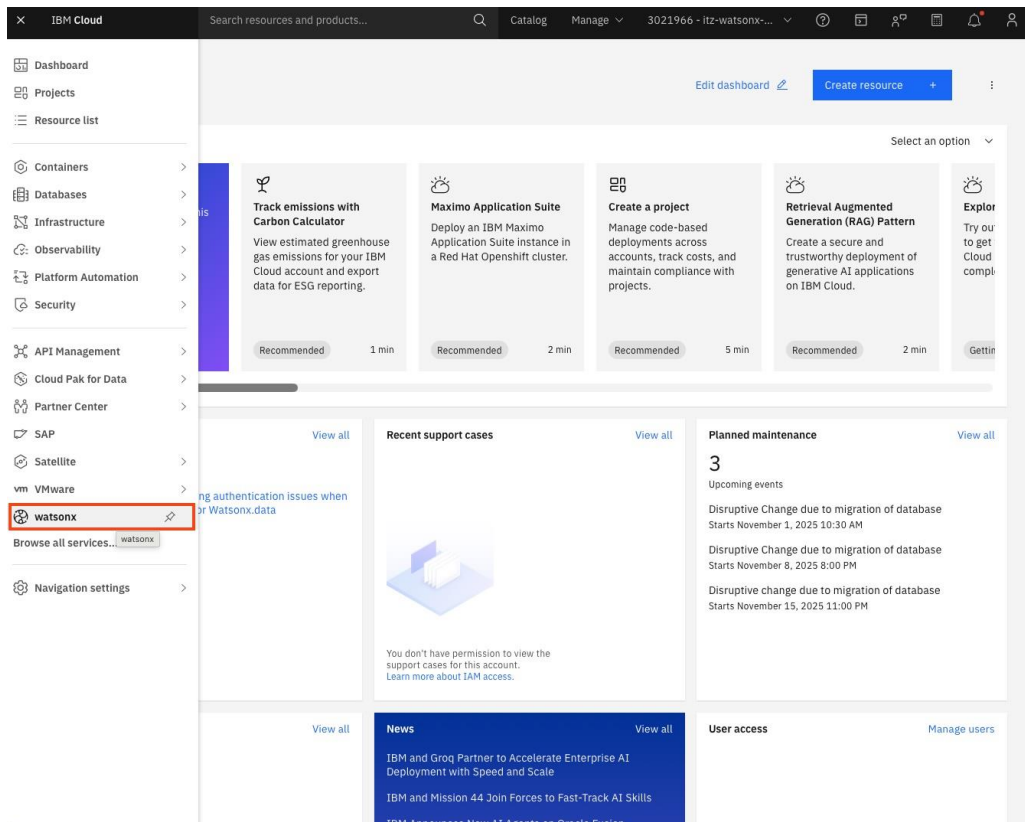
- A TechZone environment has been reserved for you, and you will receive the access details via email.
- To access your Watsonx Orchestrate and Studio instance, **please check with your instructor** for the access link and the necessary steps.

Step by step instructions:

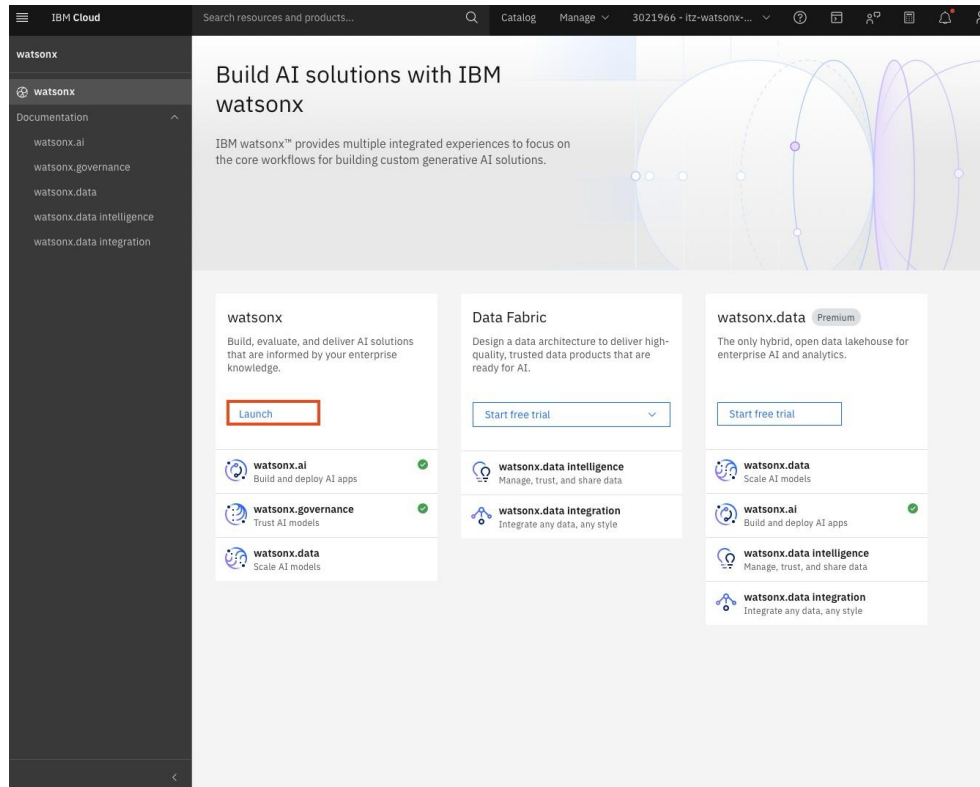
1. When you launch IBM Cloud, you'll be directed to this page. Click on the hamburger menu in the top left corner:



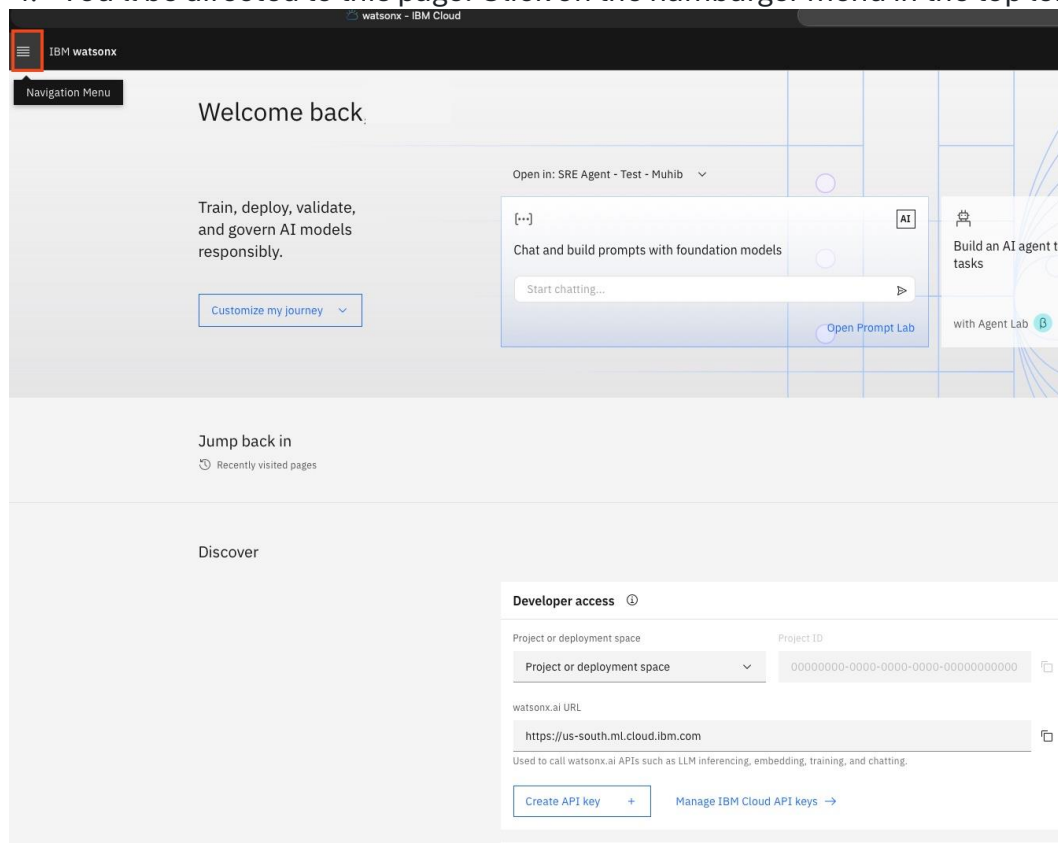
2. Click on **watsonx**:



3. Click on Launch:



4. You'll be directed to this page. Click on the hamburger menu in the top left corner:



5. Click on **View all projects**:

The screenshot shows the IBM watsonx console interface. On the left, a dark sidebar contains a navigation menu with the following items: Home, Data, Connectivity, Projects, View all projects (highlighted with a red box), Jobs, AI governance, AI use cases, Inventories, Governed agentic catalog, Guardrail manager, Deployment spaces, Resource hub, Administration, and Support. The main content area displays a dashboard with several cards: 'Open in: SRE Agent - Test - Muhib', 'Chat and build prompts with foundation models' (with a 'Start chatting...' input and 'Open Prompt Lab' button), 'Build an AI agent to automate tasks' (with 'with Agent Lab' and a beta icon), and 'Tune a foundation model with labeled data' (with 'with Tuning Studio'). Below these cards, there are sections for 'Developer access' (showing Project ID and watsonx.ai URL) and 'Developer hub' (with a 'New watsonx Developer Hub' announcement).

6. Click New Project +

The screenshot shows the 'Projects' page in the IBM watsonx console. At the top, there's a search bar with the text 'Find a project'. To the right of the search bar, a blue button labeled 'New project +' is highlighted with a red box. Below the search bar is a table with columns: Name, Date created, Your role, Collaborators, and Tags. The table is currently empty. At the bottom of the page, there is an illustration of a person standing next to a large blue cube with a white plus sign, with the text 'Start creating projects' and 'Click New project.' next to it.

7. Create a new project and name it

IBM watsonx

New project | IBM watsonx

3021952 - itz-watsonx-eve... Dallas MK

Create a project

Start with a new, blank project or select from where to import an existing project.

+ New

- Local file
- Sample

Define details

Name

SME Agent Notebook

Description (optional)

What's the purpose of this project?

Tags (optional)

Add tags

Add tags to make projects easier to find. To add tags, separate them with commas and press Enter.

Storage

cos-itx-wxo-68dec3b9e2b86cbdfa15df

Project includes integration with [Cloud Object Storage](#) for storing project assets.

Advanced settings

Cancel Create

8. Click **Add data to work with:**

IBM watsonx

SME Agent Notebook — Projects | IBM watsonx

3021952 - itz-watsonx-eve... Dallas MK

Projects / SME Agent Notebook

Overview Assets Jobs Manage

Start working

Recommended

- Add users as collaborators
- Add data to work with**
- Chat and build prompts with foundation models
- Tune a foundation model with labeled data

View all Collapse

Jump back in

By all

Assets that you create with tools show here. See all assets, including data assets, on the Assets page.

View all

Resource usage

For this month in this project

2.4 cuh

0 Tokens

0 Hosting hours

0 Pages

AI governance

No associations

AI use case you associate with this project show here.

Your documentation

New!

Get started with your documentation

You can create and manage documents about work that you do in this project.

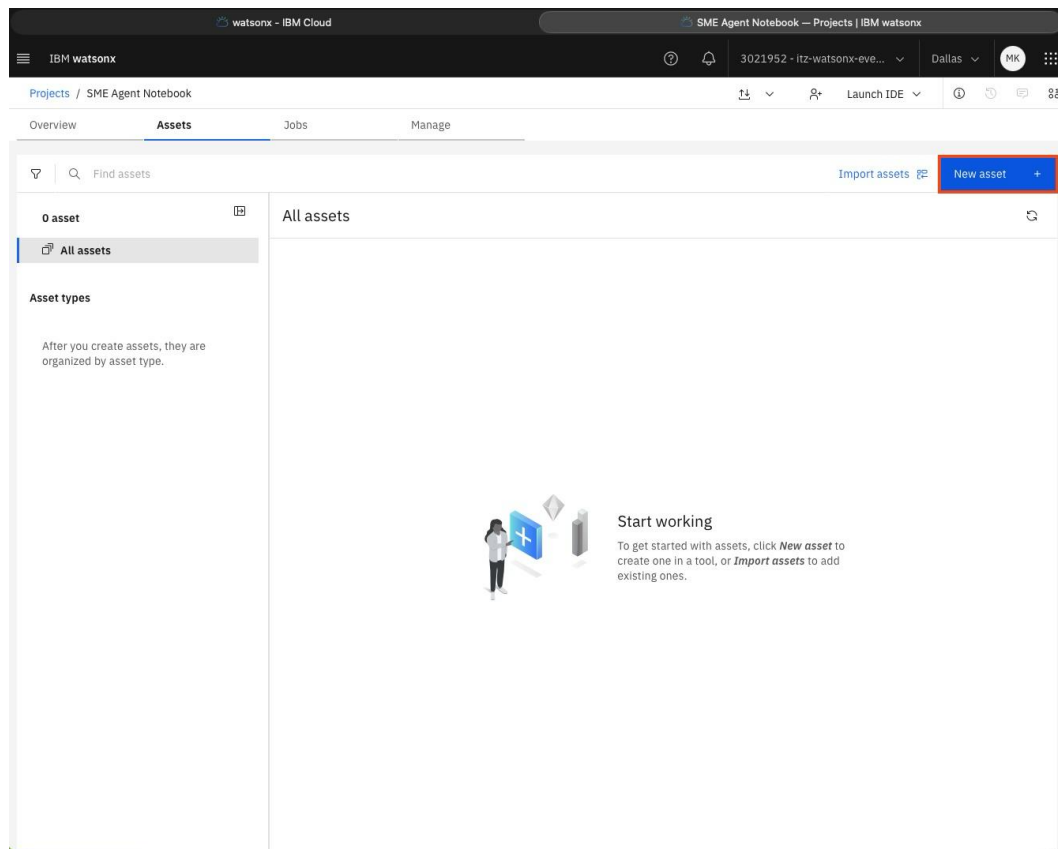
Open Documentation editor

Project history

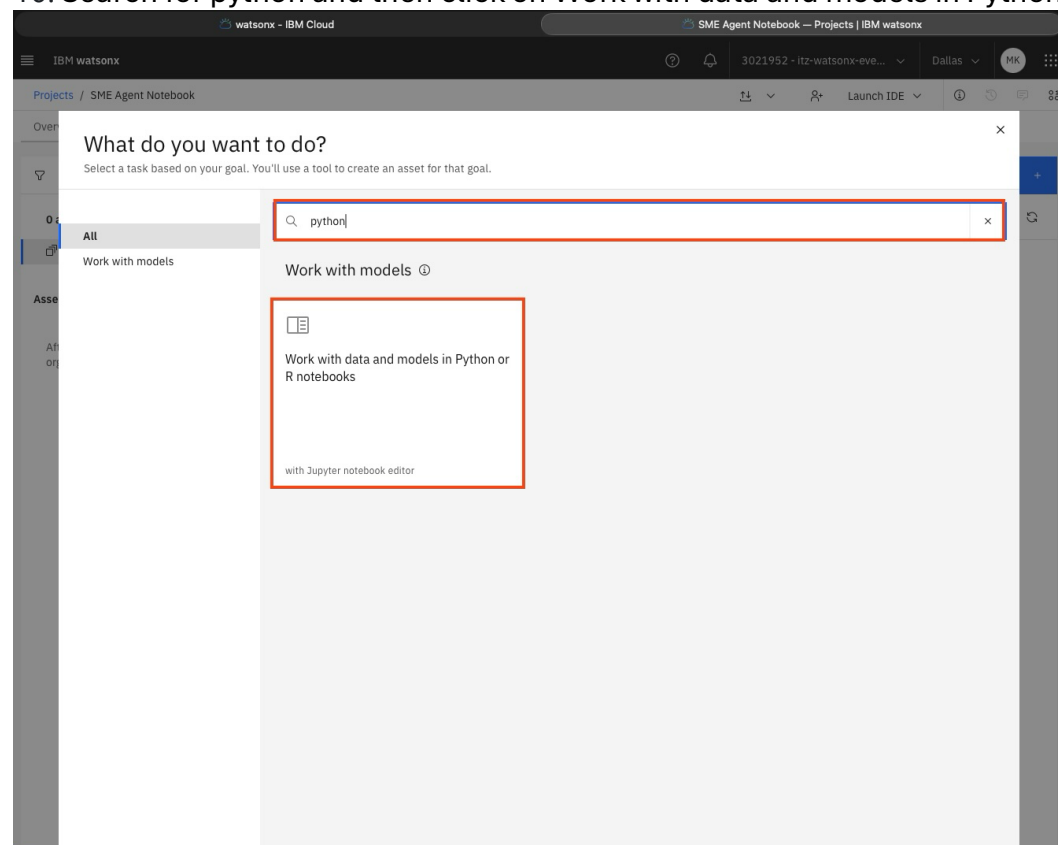
No notifications

You will see your most recent notifications here.

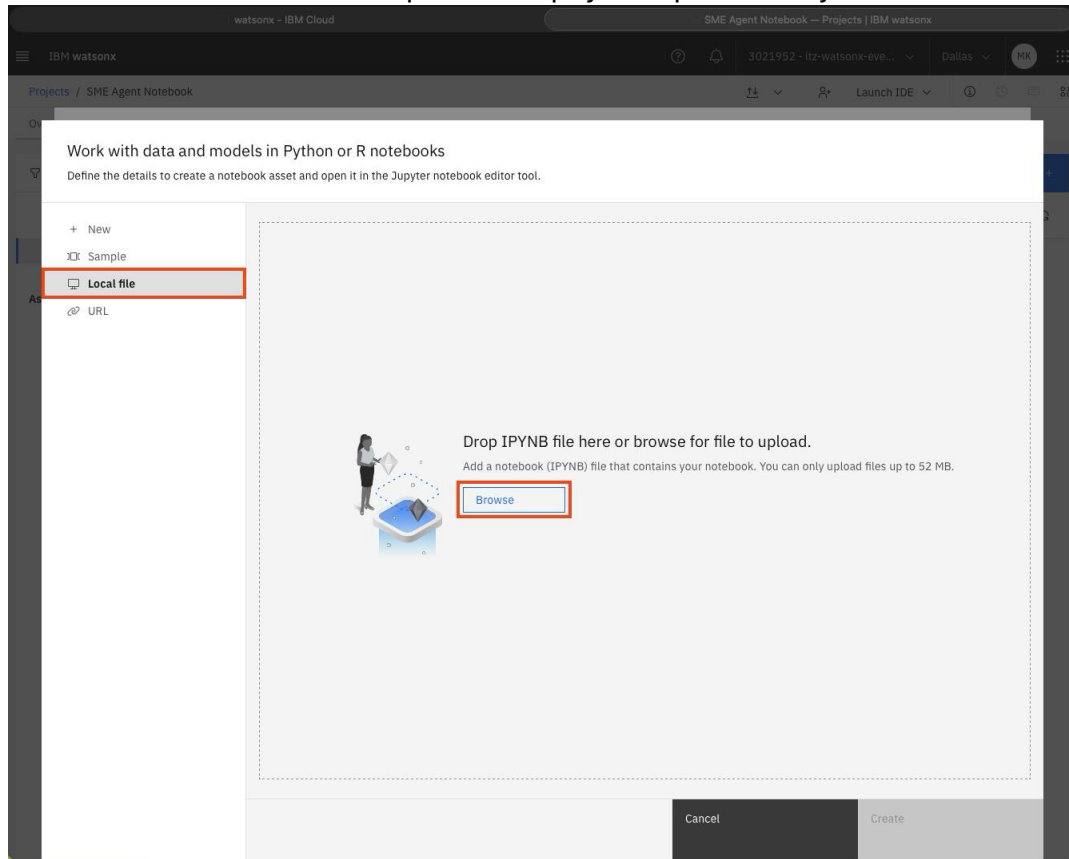
9. Click on New Asset +:



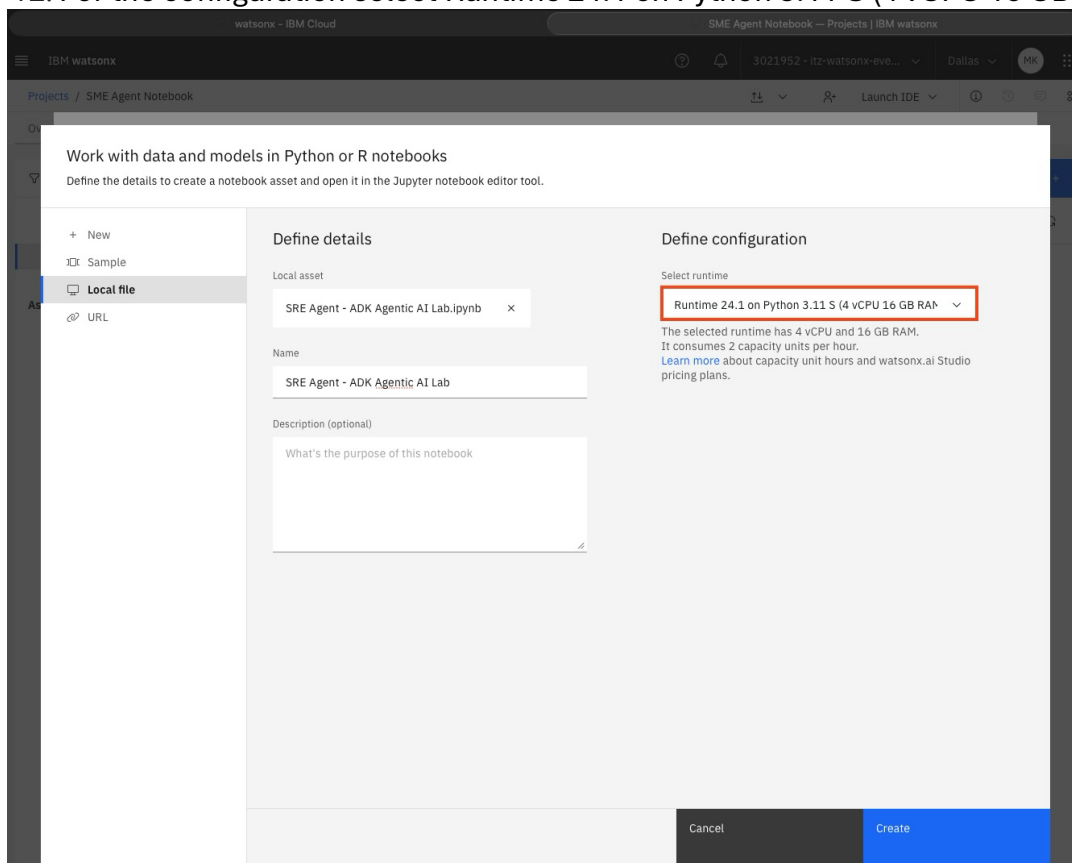
10. Search for python and then click on Work with data and models in Python or R notebooks:



11. Now click Local file and upload the ipnyb file provided by the instructor



12. For the configuration select Runtime 24.1 on Python 3.11 S (4 vCPU 16 GB RAM)



13. You should end up with this page with the notebook loaded and follow the instructions within the notebook.

IBM watsonx

Projects / SME Agent Notebook / SRE Agent - ADK Agentic AI Lab

File Edit View Run Kernel Help

Not Trusted Memory:154 / 16384 MB Python 3.11

Getting Started with the ADK

Required Software

```
[1]: !python --version
```

Python 3.11.13

1.) Installing the ADK

```
[2]: !pip install --upgrade ibm-watsonx-orchestrate
```

Collecting ibm-watsonx-orchestrate
 Downloading ibm_watsonx_orchestrate-1.13.0-py3-none-any.whl.metadata (1.4 kB)
Collecting certifi==2024.8.30 (from ibm-watsonx-orchestrate)
 Downloading certifi-2024.8.30-py3-none-any.whl.metadata (2.5 kB)
Collecting click<8.2.0,>=8.0.0 (from ibm-watsonx-orchestrate)
 Using cached click-8.1.8-py3-none-any.whl.metadata (2.3 kB)
Collecting docstring-parser<1.0,>=0.16 (from ibm-watsonx-orchestrate)
 Using cached docstring_parser-0.17.0-py3-none-any.whl.metadata (3.5 kB)
Collecting httpx<1.0.0,>=0.28.1 (from ibm-watsonx-orchestrate)
 Using cached httpx-0.28.1-py3-none-any.whl.metadata (7.1 kB)
Collecting ibm-cloud-sdk-core>=3.24.2 (from ibm-watsonx-orchestrate)
 Using cached ibm_cloud_sdk_core-3.24.2-py3-none-any.whl.metadata (8.7 kB)
Collecting ibm-watsonx-orchestrate-evaluation-framework==1.1.5 (from ibm-watsonx-orchestrate)
 Downloading ibm_watsonx_orchestrate_evaluation_framework-1.1.5-py3-none-any.whl.metadata (1.7 kB)
Collecting jsonref==1.1.0 (from ibm-watsonx-orchestrate)
 Using cached jsonref-1.1.0-py3-none-any.whl.metadata (2.7 kB)
Collecting langchain-core==0.3.63 (from ibm-watsonx-orchestrate)
 Using cached langchain_core-0.3.63-py3-none-any.whl.metadata (5.8 kB)

2.) Configure Your Env in ADK

To connect to IBM Cloud you need both the **service instance URL** and an **API key** of your watsonx orchestrate instance

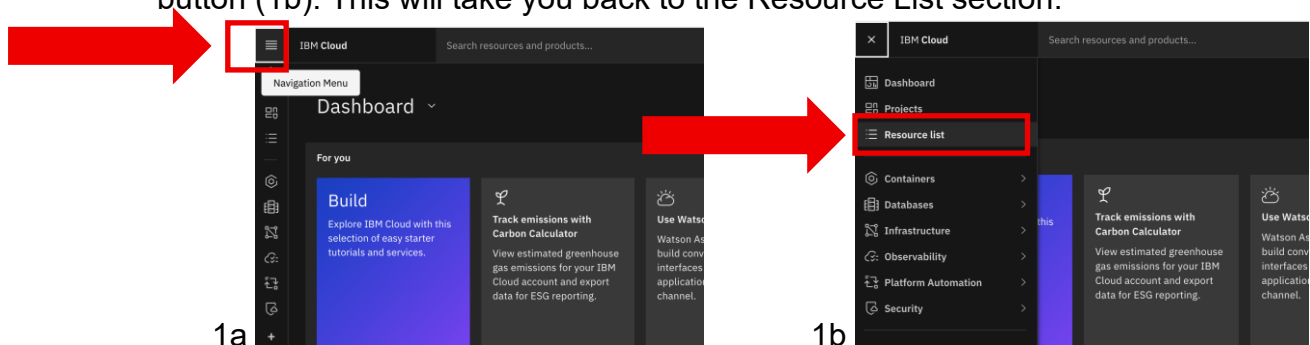
- 1.) Login to watsonx Orchestrate Instance. [URL](#)
- 2.) Click your user icon on the top right and click **Settings**

IBM watsonx Orchestrate

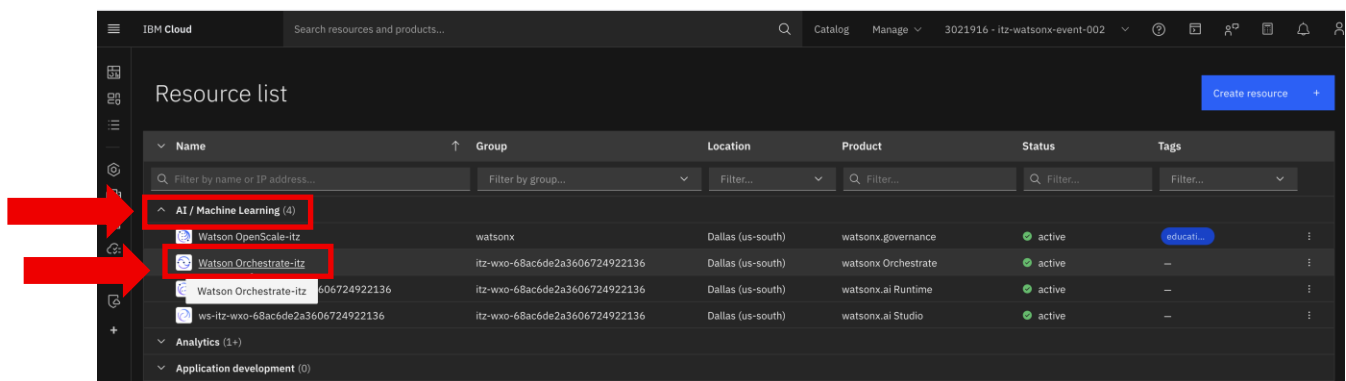
Accessing Your Watsonx Orchestrate Instance :

During the lab you will use the watsonx Orchestrate ADK to import agents, tools, knowledge base to the Orchestrate environment. Once you have imported these assets you can access the Orchestrate instance using the instructions below

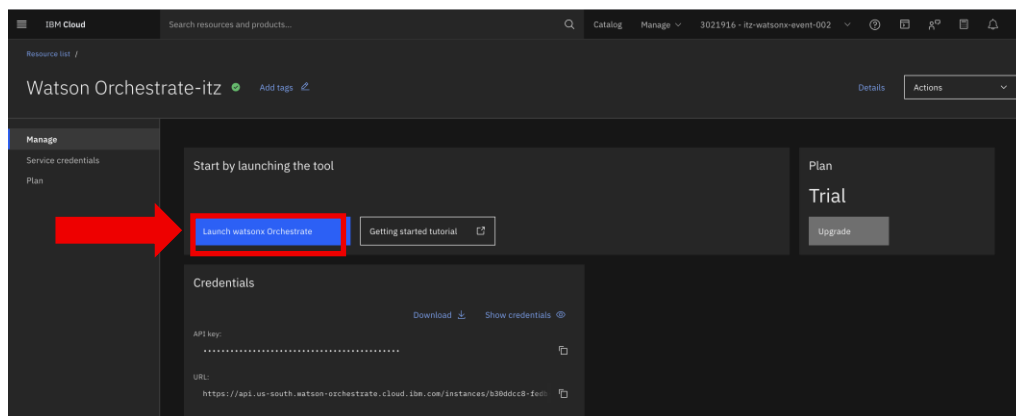
- 1.) Navigate to the [IBM Cloud](#) Home Page. To view IBM Resources select the “Navigation Menu” button on the top left (1a), and then select the “Resource List” button (1b). This will take you back to the Resource List section.



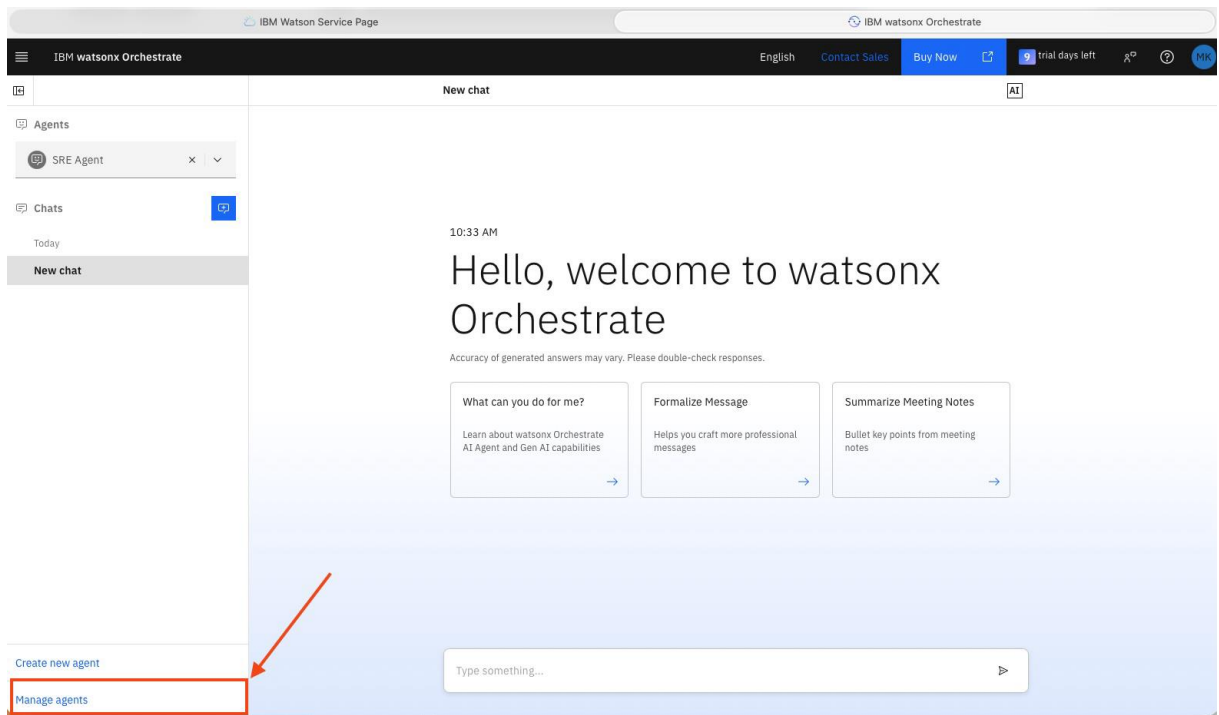
- 2.) Select the “AI/Machine Learning” Drop Down on the left. And then click the Resource “Watson Orchestrate-itz”



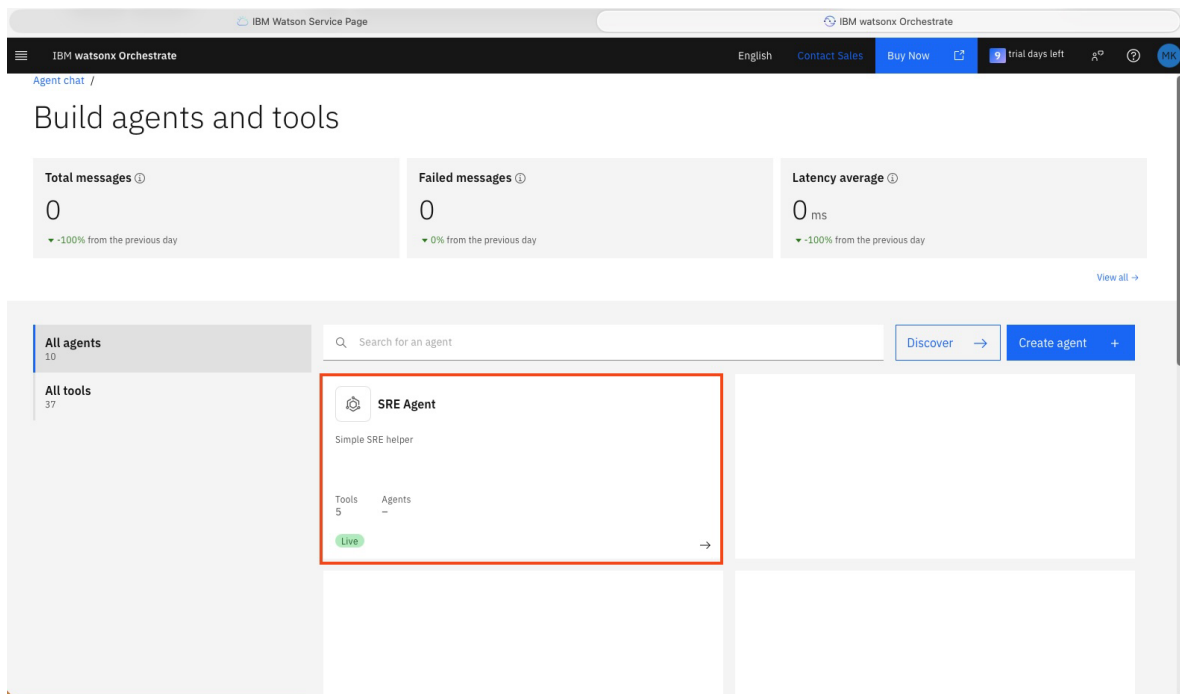
- 3.) Select “Launch watsonx Orchestrate”



4.) You have now logged into watsonx Orchestrate. Click Manage Agents on the bottom



5.) You will navigated to the list of the agents created. Click on the agent you created.



6.) Now, you can start testing the agent you have created!

The screenshot displays the IBM watsonx Orchestrate web interface. The top navigation bar includes the IBM Watson Service Page link, the product name 'IBM watsonx Orchestrate', and links for 'English', 'Contact Sales', and 'Buy Now'. A 'trial days left' indicator is also present. The main header shows the breadcrumb 'Agent chat / Manage agents / SRE Agent'.

The left sidebar contains a menu with 'Profile', 'Knowledge', 'Toolset', 'Behavior', and 'Channels'. The 'Channels' item is highlighted with a 'Preview' label. The main content area is divided into two sections: 'Profile' and 'Preview'.

The 'Profile' section is titled 'SRE Agent' and shows the 'AI Model' as 'llama-3-2-90b-...'. It includes a 'Description*' field with the text 'Simple SRE helper'. Below this is a 'Welcome message' field with the text 'Hello, welcome to watsonx Orchestrate'. A 'Quick start prompts' section at the bottom allows adding pre-set messages.

The 'Preview' section shows a live chat interface. It displays a timestamp '10:38 AM' and a large welcome message: 'Hello, welcome to watsonx Orchestrate'. Below this, three suggested actions are shown: 'What can you do for me?', 'Formalize Message', and 'Summarize Meeting Notes'. A red rectangular box highlights the input field at the bottom of the chat interface, which contains the placeholder text 'Type something...'. The 'Options' button is visible in the top right corner of the preview area.