Embedding agents in applications

Last Updated: 2025-11-14

IBM watsonx Orchestrate makes it easy to embed intelligent agents directly into your web applications by using the embedded chat feature. It enables interactive, contextual conversations while maintaining enterprise-grade security and flexibility.

With embedded chat, you can:

- Integrate AI agents seamlessly into your web pages to create rich, interactive experiences.
- Customize the UI and behavior to match your brand and user needs.
- Handle advanced events securely, and work with both draft and live agents for testing and production.

Agents connect with users through channels with entry points like websites, Slack workspaces, or other messaging platforms. When you use the web chat channel, you embed your agent directly into a web page, enabling dynamic, interactive conversations that boost engagement and productivity.

General considerations

Before you embed an agent into your web application, it's important to prepare both the agent and the application environment to ensure a smooth integration. Consider the following essential concepts:

1. Draft and the Live mode

   - Use the Draft mode when you're testing or making changes. In Draft mode, the agent is not fully published, so you can safely experiment without affecting live users.
   - Deploying the agent moves it from Draft to Live for production use, thus making it accessible to your users.
   - When you embed an agent into a web page, the embedded script you use depends on the environment (Draft or Live).

2. Application setup

   Host your app on a server (local or cloud). For details on adding the embed snippet, see Procedure to embed agents in applications.

3. Security configurations

Security is *enabled by default in both Draft and Live environments but not configured*. Until configuration is complete, the embedded agent remains inactive. See Configuring security for embedded chat for more details.

> **Notes:**
>
> – **For IBM watsonx Orchestrate On-premises deployment:**
> – Version 5.2.1: Copying the embed script from the agent UI is not supported. Use the ADK docs to get the embed script. For more details, see Embedding Agent Web Chat.
> – Versions 5.2.2 and later: You can directly copy the embed scripts from the agent UI.

Procedure to embed agents in applications

You can customize your chat UI by embedding the agent. Start by embedding the agent in your draft environment before moving to production.

Steps to embed the agent:

1. Access the agent configuration:
   - Go to the agent configuration page.
   - Navigate to **Channels** > **Embedded agent**.
2. Select the environment:
   - Use **Draft** mode for testing and previewing the embedded experience.
   - The **Live** tab remains disabled until the agent is deployed.
3. Copy the embed script:
   - Click **Copy to clipboard** to copy the generated embed script.
4. Add the script to your web page:
   - Edit the HTML source of the page where you want the chat widget to appear.
   - Paste the embed script inside the <body> tag, as close as possible to the closing </body> tag for optimal page load performance.
   - Ensure the `root` element is present.

   Example HTML structure:

```
<body>
  <div id="root"></div>
  <script src="path-to-embed-script.js"></script>
</body>
```

5. Preview the chat widget

   - Open the web page in your browser.
   - Verify that the chat widget appears and opens when clicked.

6. Apply to multiple pages

   - Paste the same embed script into each page where you want the widget.
   - Do not include the script multiple times on a page to avoid conflicts.

Configuring security for embedded chat

To ensure secure communication between your application and IBM watsonx Orchestrate, the embedded chat feature uses public-key cryptography and JWT-based authentication. These mechanisms protect sensitive data and validate user identity during interactions.

Security setup involves three key steps:

– Generate RSA key pairs for IBM and your client application.
– Configure JWT authentication for secure user identity.
– Encrypt sensitive data by using IBM's public key.

Understanding default security settings

Security is **enabled by default**. However, it is **not configured** initially. Until you complete the configuration, the embedded agent remains inactive. Here's what you need to know:

– The embedded agent will not function until security setup is complete.
– You must configure both IBM and client key pairs to establish secure communication.
– Security can be disabled to allow anonymous access, but this is **not recommended for sensitive use cases**.

See Customizing your embedded chat for setting up security.

Customizing your embedded chat

After embedding your agent through the UI, you can extend functionality and customize the experience using the Agent Development Kit (ADK).

Security configuration

See Embedded chat security for full details and setup instructions.

Automated security setup

You can simplify the process of securing your embedded agent integration by using the security configuration tool. See Enabling security to automate key generation, JWT configuration, and encryption.

Using context variables

To use context variables in embedded chat, include them inside the JWT token. You can add context variables to a JWT token by using a JavaScript script. See Context variables for embedded chat for more details.

Custom styling and layout

The watsonx Orchestrate embed supports a flexible layout object that enables you to control how and where the chat UI appears on the page. See Customizing layout and Customizing styles for more details.

Enable feedback

By default, thumbs-up and thumbs-down feedback options are disabled. You need to enable them manually to collect user input on responses. See Enabling feedback for more details.

What to do next

After configuring your agent, the final step is to deploy it so it becomes available for users. Deployment moves the agent from Draft to Live, making it available for users in production.

See Deploying your agent for more details.