

Hosting your Web Application on Azure

Azure Machine Learning provides a powerful platform for running and deploying large language models (LLMs). This guide will walk you through the steps to deploy your LLM and create an endpoint for your LLM in Azure Machine Learning.

Prerequisites

- An active Azure account - [Setting Up a free Azure Account - NLP Community - Confluence \(iabg.de\)](#)
- An API-accessible LLM, either from closed-source providers (GPT, Claude, Mistral Large etc.) OR a self-deployed LLM - [Deploying Large Language Model \(LLM\) on Azure](#)

Steps to Host your Web Application

- **Sign in to Azure Portal:**
 - Open your web browser and navigate to the Azure Portal: <https://portal.azure.com>.
 - Sign in with your Azure account credentials

- From the home page, select "Create a resource" and then "Web App":

Hi Paolo, see what more you can get from your Azure free account.

[View remaining credits](#) to try any service, or [browse free services](#) included with your account.

The screenshot shows the Azure portal home page. At the top, there are four cards: 'Take a free Azure fundamentals tutorial', 'Watch a demo and attend a live Q&A', 'Start a project with Quickstart Center', and 'Explore support resources'. Below these is the 'Azure services' section with a grid of icons. The 'Create a resource' icon (a blue plus sign) is highlighted with a red box. Other icons include 'Azure Machine Learning', 'Cost Management', 'Container registries', 'All resources', 'Azure Managed...', 'Kubernetes services', 'Quickstart Center', 'Virtual machines', and 'More services'. Below the 'Azure services' section is the 'Resources' section with tabs for 'Recent' and 'Favorite'. The 'Recent' tab is active, showing a table of resources:

Name	Type	Last Viewed
ml-test-us	Azure Machine Learning workspace	3 hours ago
ml-chatbot-test	App Service	7 days ago
Free trial	Subscription	2 weeks ago
TVL	Resource group	2 weeks ago

Below the 'Resources' section is the 'Navigate' section with icons for 'Subscriptions', 'Resource groups', 'All resources', and 'Dashboard'. At the bottom of the page, there is a 'Create a resource' section with a search bar and a list of popular Azure services and marketplace products. The 'Web App' service is highlighted with a red box in the 'Popular Azure services' section.

- Fill in the initial configuration for your Web App:

For this tutorial, we are deploying a Python web-based application directly from a code repository. Other options (e.g. Docker deployment, other programming languages) can also be specified here.

Under "Basics", select your free trial subscription and a resource group (if you have created one). Specify a name for your web app (which will also be used as the final URL), select "Code" as publish mode, your preferred Python version and Region (e.g. "Germany West Central").

[Basics](#) [Database](#) [Deployment](#) [Networking](#) [Monitoring](#) [Tags](#) [Review + create](#)

App Service Web Apps lets you quickly build, deploy, and scale enterprise-grade web, mobile, and API apps running on any platform. Meet rigorous performance, scalability, security and compliance requirements while using a fully managed platform to perform infrastructure maintenance. [Learn more](#)

Project Details

Select a subscription to manage deployed resources and costs. Use resource groups like folders to organize and manage all your resources.

Subscription * ⓘ

Free trial

Resource Group * ⓘ

TVL

[Create new](#)

Instance Details

Name *

my-test-iabg-webapp

.azurewebsites.net

Publish *

☒ Code ☐ Docker Container ☐ Static Web App

Runtime stack *

Python 3.12

Operating System *

☒ Linux ☐ Windows

Region *

Germany West Central

ⓘ Not finding your App Service Plan? Try a different region or select your App Service Environment.

Pricing plans

App Service plan pricing tier determines the location, features, cost and compute resources associated with your app. [Learn more](#)

Linux Plan (Germany West Central) * ⓘ

ASP-TVL-b586 (F1)

[Create new](#)

Pricing plan **Free F1** (Shared infrastructure)

Zone redundancy

An App Service plan can be deployed as a zone redundant service in the regions that support it. This is a deployment

[Review + create](#) [< Previous](#) [Next : Database >](#)

You can also customize other aspects of your applications in the other tabs. For example, if you want your app to be publicly accessible, ensure that "Enable public access" is turned on under "Networking":

[Basics](#) [Database](#) [Deployment](#) [Networking](#) [Monitoring](#) [Tags](#) [Review + create](#)

Web Apps can be provisioned with the inbound address being public to the internet or isolated to an Azure virtual network. Web Apps can also be provisioned with outbound traffic able to reach endpoints in a virtual network, be governed by network security groups or affected by virtual network routes. By default, your app is open to the internet and cannot reach into a virtual network. These aspects can also be changed after the app is provisioned. [Learn more](#)

Enable public access * ⓘ ☒ On ☐ Off

⚠ Network injection is only available in Basic, Standard, Premium, Premium V2, and Premium V3 Dedicated App Service plans.

Enable network injection ☐ On ☒ Off

In general, the default options for the advanced tabs are sufficient. When you are done, click on "Review + create" in the lower part of the screen.

- **Wait for the App to be deployed**

Delete
Cancel
Redeploy
Download
Refresh

Deployment is in progress

Deployment name: Microsoft.Web-WebApp-Portal-537fedc4-8f4e
Subscription: Free trial
Resource group: TVL

Start time: 3/12/2024, 11:47:46 AM
Correlation ID: d7fab962-1d35-4795-9aec-4fed20707bb8

Deployment details

Resource	Type	Status	Operation details
No results.			

Give feedback
Tell us about your experience with deployment

Your deployment is complete

Deployment name: Microsoft.Web-WebApp-Portal-537fedc4-8f4e
Subscription: Free trial
Resource group: TVL

Start time: 3/12/2024, 11:47:46 AM
Correlation ID: d7fab962-1d35-4795-9aec-4fed20707bb8

Deployment details

Next steps

Manage deployments for your app. Recommended
Protect your app with authentication. Recommended

Go to resource

Give feedback
Tell us about your experience with deployment

when deployment is complete, click on "Go to resource". This page gives you an overview on everything related to your current WebApp deployment (usage, logs, monitoring, accessibility, etc..)

Connect a repository to your Web App

Under "Deployment Center" in the left pane, you can specify the source of your app code. Fill in the Source provider (in our example "GitHub"), repository name, branch and other details. This code repository will be copied in the app container. You can access it later on via SSH (see [Debugging](#)).

my-test-iabg-webapp | Deployment Center

Web App

Search
Save
Discard
Browse
Manage publish profile
Sync
Leave Feedback

Overview
Activity log
Access control (IAM)
Tags
Diagnose and solve problems
Microsoft Defender for Cloud
Events (preview)

Deployment
Deployment slots
Deployment Center

Settings
Configuration
Authentication
Application Insights
Identity
Backups
Custom domains
Certificates
Networking
Scale up (App Service plan)
Scale out (App Service plan)
Service Connector
Locks
App Service plan
Quotas
Change App Service plan

Deploy and build code from your preferred source and build provider. [Learn more](#)

Source *
GitHub

Building with GitHub Actions. [Change provider.](#)

GitHub

App Service will place a GitHub Actions workflow in your chosen repository to build and deploy your app whenever there is a commit on the chosen branch. If you can't find an organization or repository, you may need to enable additional permissions on GitHub. You must have write access to your chosen GitHub repository to deploy with GitHub Actions. [Learn more](#)

Signed in as
paolo-notaro [Change Account](#)

Organization *
tv1-ai-microclass-exercises

Repository *
radio-chatbot-demo

Branch *
main

Workflow Option *
Add a workflow: Add a new workflow file 'main_my-test-iabg-webapp.yml' in the selected repository and branch.
Use available workflow: Use one of the workflow files available in the selected repository and branch.

Build

Runtime stack
Python

Version
Python 3.12

Authentication settings

Select how you want your GitHub Action workflow to authenticate to Azure. If you choose user-assigned identity, the identity selected will be federated with GitHub as an authorized client and given write permissions on the app. [Learn more](#)

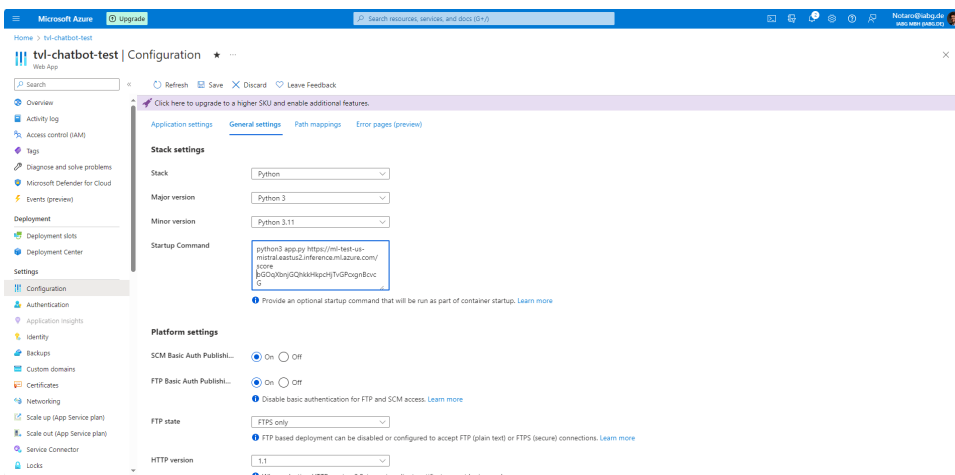
Authentication type *
User-assigned identity
Basic authentication

Subscription *
Free trial

Set up an endpoint for your Web App

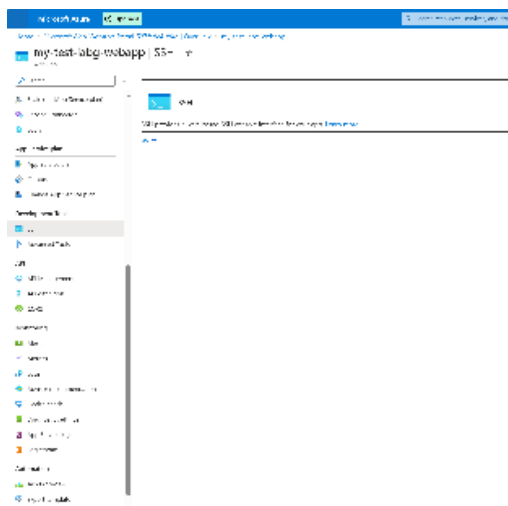
Under "Settings" in the left pane, select the "General Settings" tab. Then, specify your endpoint options.

In the example below, we specified a Python start-up script with CLI arguments (LLM API endpoint URL, LLM API key).



Debugging your Web App

You can use the SSH connection option from the left pane to connect inside your app container. There you have all your application code and computing resources. You can also set up a debugger via SSH tunnel (e.g. in VSCode) or try running tests directly.





Documentation: <http://aka.ms/webapp-linux>

Python 3.11.7

Note: Any data outside '/home' is not persisted

(antenv) root@c08314d92787:/tmp/8dc426cfc44c37b# cd src/

(antenv) root@c08314d92787:/tmp/8dc426cfc44c37b/src# █