

# Explora la visión por computadora

En este ejercicio, utilizarás un entorno de chat interactivo para interactuar con una solución de IA generativa capaz de analizar e interpretar imágenes.

Este ejercicio debería tardar aproximadamente **15** minutos en completarse.

## Utilizar visión artificial para analizar imágenes

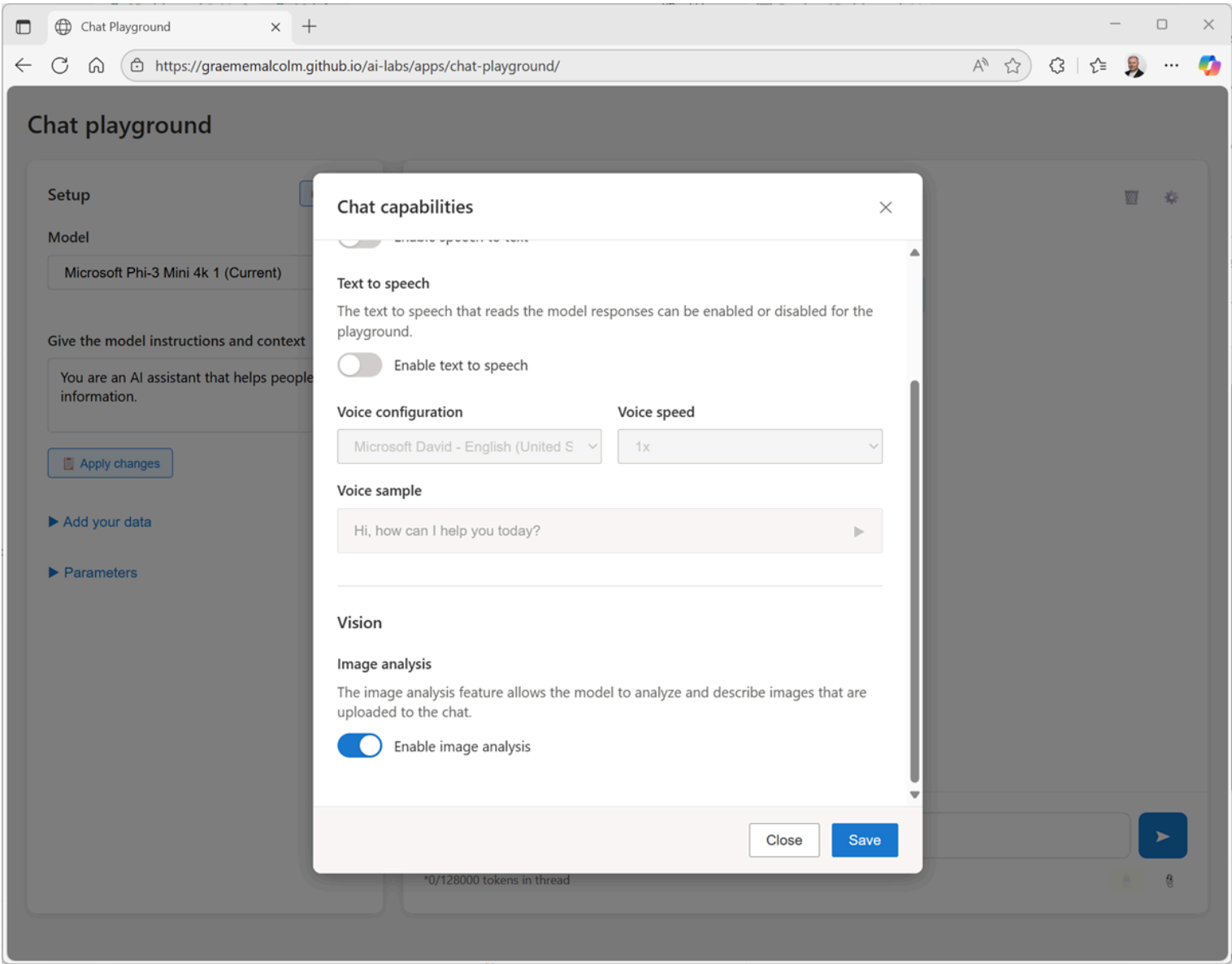
Comencemos conversando con un modelo de IA generativa. En este ejercicio, utilizaremos el modelo de visión artificial basado en CNN **MobileNetV3 para la clasificación de imágenes y el modelo de lenguaje natural Microsoft Phi 3 Mini** para interpretar las imágenes y responder preguntas sobre ellas.

! **Nota** : Los modelos se ejecutarán en su navegador, en su ordenador. El rendimiento puede variar según la memoria disponible y el ancho de banda de su red. Si su navegador no admite los modelos WebLLM, se activará un modo alternativo con funcionalidad reducida.

1. En un navegador web, abra el [Chat Playground](#) en `https://aka.ms/chat-playground` .
2. Espere a que el modelo se descargue e inicialice.

! **Consejo** : La primera vez que abras el chat de pruebas, la descarga del modelo puede tardar unos minutos. Las descargas posteriores serán más rápidas.

3. Cuando el modelo esté listo, en una nueva pestaña del navegador, descargue [images.zip](#) a `https://aka.ms/ai-images` su ordenador local.
4. Extrae el archivo descargado en una carpeta local para ver los archivos que contiene. Estos archivos son las imágenes que analizarás con IA.
5. Regrese a la pestaña del navegador que contiene el área de chat y, en la parte superior derecha del panel del **historial de chat** , utilice el botón **Configuración** ( ⚙ ) para ver las opciones de capacidades de chat.
6. En la sección **Visión** , habilite **el análisis de imágenes** . Luego, espere a que se descargue e inicialice el modelo de visión por computadora antes de guardar los cambios.



Utilizar visión artificial para analizar imágenes

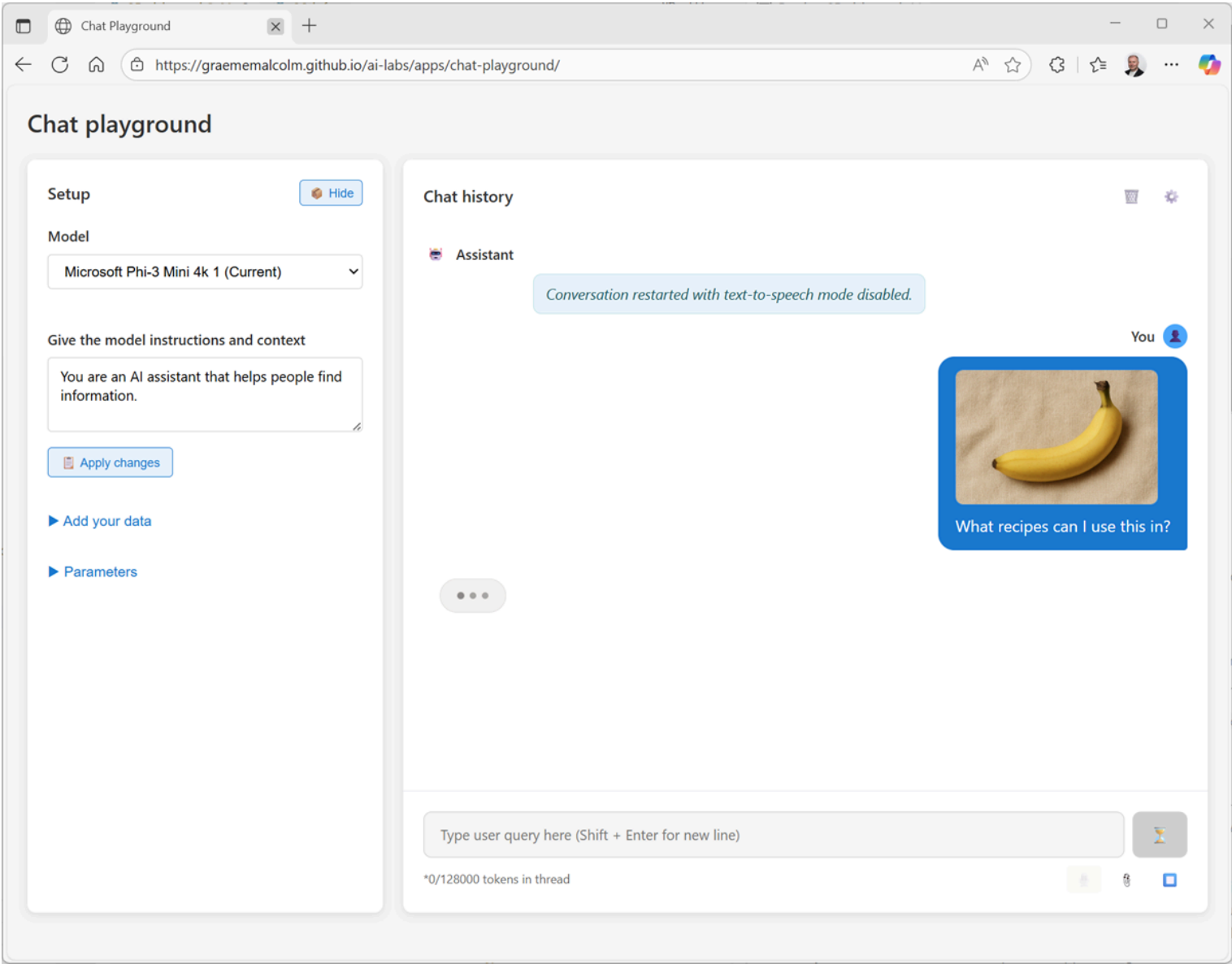
Resumen

Debajo de la interfaz de chat, está habilitado un botón **para subir imágenes** ( 🖼️ ).

7. Haz clic en el botón **"Subir imagen"** y busca para seleccionar una de las imágenes que has extraído en tu ordenador.

Se añade una miniatura de la imagen al área de entrada de la solicitud.

8. Introduzca un mensaje como el siguiente `What recipes can I use this in?` . La imagen se incluye en el mensaje.



El modelo MobileNetV3 se utiliza para determinar el sujeto probable de la imagen, y los resultados de dicho análisis se incluyen en la solicitud al modelo de lenguaje Phi. El resultado debe ser una respuesta que utilice la información de la imagen para responder a la pregunta.

9. Envía propuestas que incluyan las otras imágenes, como por 

How should I cook this?

 ejemplo:

What desserts could I make with this?

## Resumen

En este ejercicio, exploraste el uso de la visión por computadora con un modelo de IA generativo en un entorno de chat interactivo.

Algunos modelos son *multimodales* y admiten de forma nativa entradas basadas en imágenes; otras soluciones, en cambio, se basan en el uso de un modelo independiente de clasificación de imágenes o detección de objetos. Azure AI Foundry admite una variedad de modelos multimodales, así como modelos de visión de terceros en su catálogo. Además, puede usar el servicio Azure AI Content Understanding para analizar imágenes.