

Ensamblaje del agente en Retell AI

Datos de esta lección

Fase	F2 — Capa de telefonía y generación de voz
Módulo	M2.3 — Ensamblaje del agente en Vapi / Retell
Puntos de profesionalidad al superar el test	60

Objetivo práctico

Al terminar esta lección tendrás una cuenta de Retell AI creada y un primer agente configurado con LLM, voz y prompt, probado mediante llamada de prueba.

Resultado que obtendrás

Un agente de Retell funcional, con tu Prompt Maestro cargado, probado y respondiendo correctamente.

Dónde encaja dentro del sistema final

Retell AI es la segunda plataforma de orquestación de referencia de este curso, alternativa a Vapi. Este curso enseña ambas con la misma profundidad para que elijas la más adecuada a cada proyecto o cliente.

Conocimientos previos

- Prompt Maestro completo (módulo 1.3).
- Voz probada en ElevenLabs (módulo 2.2).
- Ensamblaje en Vapi (lección anterior, como referencia comparativa).

Herramientas necesarias

- Cuenta de Retell AI.
- Tu Prompt Maestro personalizado.

Posible coste

Según la documentación oficial de precios de Retell AI verificada el 30 de agosto

de 2026, el coste por minuto se compone de infraestructura de voz (aproximadamente 0,055 USD/min), texto a voz (entre 0,015 y 0,040 USD/min según el proveedor), LLM (desde aproximadamente 0,003 USD/min en modelos económicos hasta cifras muy superiores en modelos de alta gama) y telefonía (aproximadamente 0,015 USD/min, variable según el país). Retell ofrece 10 USD de crédito gratuito para empezar a probar sin compromiso inicial, y 20 llamadas concurrentes gratuitas por cuenta.

Advertencia de seguridad

Sigue el protocolo del módulo 2.1. La documentación oficial de precios de Retell consultada no menciona una función de límite de gasto configurable nativa: aplica las medidas externas del módulo 2.1 si vas a hacer pruebas extensas.

Modelo mental

Igual que con Vapi, piensa en Retell como en otra centralita telefónica completa, con sus propios mandos y disposición, pero cumpliendo exactamente la misma función: conectar cerebro, oído y voz en un único agente coherente.

Explicación

Según la documentación oficial de Retell AI verificada el 30 de agosto de 2026, un agente se crea desde el panel («Agents → Create an Agent»), pudiendo elegir entre un modo de «Single prompt» (conversación libre a partir de un único prompt, más parecido a lo trabajado en este curso) o «Conversational flow» (un flujo determinista por pasos, más rígido y orientado a procesos muy estructurados). El motor de respuesta («Retell LLM») se configura de forma independiente y se asocia al agente; la voz se elige entre varios proveedores disponibles, incluido ElevenLabs.

Retell permite además configurar «Custom Functions» (funciones personalizadas): cuando el LLM decide invocar una, Retell envía una petición POST a la URL que hayas configurado, con el nombre de la función y sus parámetros — el mismo mecanismo de Webhook estudiado en el módulo 1.2, aplicado aquí al Tool Calling (módulo 3.4). Retell también ofrece un sistema de Webhooks generales de cuenta, para recibir notificaciones de eventos (inicio de llamada, fin de llamada, entre otros) sin necesidad de consultar la API repetidamente.

Vocabulario nuevo

Custom Function (Retell)

Piénsalo así: El mismo recado que el telefonista envía al departamento correspondiente mediante el buzón (Webhook).

Función personalizada que el LLM de un agente Retell puede invocar durante la conversación, enviando una petición POST a una URL configurada por ti.

Ejemplo: Una función «reservar_cita» que envía los datos de la cita a un Webhook de n8n.

Preparación

Ten a mano tu Prompt Maestro personalizado y tu clave de API de ElevenLabs.

Procedimiento paso a paso

1. **Contexto:** Vas a crear tu cuenta de Retell AI y tu primer agente.

Acción de teclado: Ve a la dirección oficial retellai.com, regístrate siguiendo el protocolo del módulo 2.1, y dentro del panel recorre encabezados con **H** hasta «Agents». Localiza «Create an Agent» con **B** o **Tab**.

Respuesta esperada de NVDA: NVDA debería anunciar un texto similar a «Create an Agent, botón».

Qué significa: Se te pedirá elegir entre un agente de voz o un agente de chat; para este curso, elige agente de voz.

Acción siguiente: Elige el modo «Single prompt», más adecuado para el Prompt Maestro trabajado en la Fase 1.

2. **Contexto:** Vas a configurar el modelo de respuesta (Retell LLM), la voz y el prompt.

Acción de teclado: Localiza el editor del «Retell LLM» asociado al agente y pega tu Prompt Maestro completo en el campo de prompt principal con **Ctrl** + **V**. Después, recorre el selector de voz y elige tu voz de ElevenLabs.

Respuesta esperada de NVDA: NVDA anunciará cada campo y selector conforme los recorras.

Qué significa: Igual que en Vapi, estás ensamblando en un único agente las piezas trabajadas en los módulos anteriores.

Acción siguiente: Guarda la configuración del agente.

3. **Contexto:** Quieres probar el agente sin usar todavía un número de teléfono real.

Acción de teclado: Busca la opción de prueba del agente (suele llamarse «Test» o mostrarse como un botón de llamada dentro del propio editor) y actívala con `Enter`.

Respuesta esperada de NVDA: NVDA anunciará el inicio de la prueba; deberías escuchar la primera respuesta del agente.

Qué significa: Esta prueba te permite verificar el comportamiento del agente sin coste de telefonía real.

Acción siguiente: Haz al menos tres preguntas de prueba, incluida una fuera del alcance de tu Prompt Maestro.

Posibles diferencias de interfaz

Retell distingue entre «Single prompt» y «Conversational flow» como dos formas distintas de construir un agente; este curso usa siempre «Single prompt» por su compatibilidad directa con la estructura de Prompt Maestro de la Fase 1. Si tu versión del panel presenta estas opciones con nombres ligeramente distintos, busca por los conceptos («prompt único» frente a «flujo por pasos»).

Errores frecuentes y diagnóstico

No encuentras dónde pegar el Prompt Maestro completo.

Cómo localizarlo: En el modo «Conversational flow», el prompt se reparte en varios nodos en lugar de un único campo.

Solución: Confirma que has elegido el modo «Single prompt» al crear el agente; si ya lo creaste con el modo equivocado, créalo de nuevo con el modo correcto.

Comprobación final

Realiza una prueba de conversación con el agente y confirma que responde según las reglas de tu Prompt Maestro, incluida la prohibición de inventar datos. Si el comportamiento es correcto, la comprobación es positiva.

Qué acabas de conseguir

Tienes un segundo agente completamente funcional, en una plataforma alternativa a Vapi, lo que te permite comparar ambas antes de decidir cuál usar en un proyecto real.

Ejercicio práctico

Compara, con las mismas preguntas de prueba, el comportamiento de tu asistente de Vapi (lección anterior) y tu agente de Retell. Anota cualquier diferencia relevante de tono, latencia o precisión.

Resumen de lo imprescindible

- Retell ofrece dos modos de construcción de agente: Single prompt y Conversational flow; este curso usa Single prompt.
- El «Retell LLM» es el motor de respuesta que se asocia al agente y contiene el prompt principal.
- Las Custom Functions envían peticiones POST a una URL configurada, igual que un Webhook.
- Se puede probar el agente sin número de teléfono real antes de avanzar al módulo 2.4.

No necesitas aprender todavía

Todavía no necesitas configurar Custom Functions ni un número de teléfono real: se trabajan en los módulos 3.4 y 2.4.

Estoy preparado para continuar si puedo...

- Crear un agente en Retell AI con LLM, voz y prompt configurados.
- Elegir el modo «Single prompt» y justificar por qué es el adecuado para este curso.
- Probar el agente y verificar su comportamiento ante preguntas fuera de alcance.

Test de comprobación

Para registrar el resultado y obtener Puntos de profesionalidad, realiza este test desde la versión web de la lección. Las opciones se muestran aquí sin señalar la respuesta correcta.

1. Vas a construir un agente en Retell usando el Prompt Maestro completo desarrollado en la Fase 1, sin dividirlo en pasos rígidos. ¿Qué modo de agente deberías elegir?
 - a. Conversational flow, porque siempre es la opción más avanzada.
 - b. Single prompt, porque permite cargar el Prompt Maestro completo en un único campo, de forma directamente compatible con lo trabajado en la Fase 1.
 - c. Ambos modos son idénticos y da igual cuál se elija.
2. Configuraste una Custom Function en Retell para que el agente registre una cita en tu sistema externo, pero al hacer la prueba nunca llega ninguna petición a ese sistema. Según esta lección, ¿qué revisarías primero?
 - a. Que la URL configurada en la Custom Function sea correcta y esté accesible, ya que Retell envía una petición POST a esa dirección al invocar la función.
 - b. La voz asociada al agente, porque un fallo en las funciones siempre indica un problema de proveedor de voz.
 - c. El modo elegido al crear el agente (Single prompt frente a Conversational flow), que no influye en las Custom Functions.
3. Comparas dos llamadas de prueba en Retell y observas que el coste por minuto varía notablemente según el modelo de LLM elegido, aunque la voz y la telefonía sean las mismas. ¿Por qué, según lo explicado en esta lección?
 - a. Porque el coste por minuto de Retell se compone de varios elementos independientes (infraestructura de voz, TTS, LLM y telefonía), y el precio del LLM varía mucho entre modelos económicos y modelos de alta gama.
 - b. Porque Retell cobra una tarifa fija de orquestación que nunca depende del modelo elegido.
 - c. Es un error de facturación, ya que todos los modelos de LLM cuestan exactamente lo mismo en Retell.

Fuentes

- [Retell AI Pricing — Retell AI \(consultado el 2026-08-30\)](#)

- [Retell AI — Custom Function — Retell AI \(consultado el 2026-08-30\)](#)

Última actualización de este contenido: **2026-08-30**.