# Plan de prueba de 30 días de un agente de IA de atención al cliente

Lo que va entre corchetes se sustituye por los datos de la empresa. Las cifras de los criterios son un ejemplo.

EMPRESA
[Nombre de la empresa]

AGENTE DE IA
[Herramienta]

DEL
[fecha] al [fecha]

LA PRUEBA, EN CIFRAS

| 30 | 1 | 30 | 4 |
|---|---|---|---|
| Días | Canal | Preguntas de prueba | Métricas |

## 01 Alcance de la prueba

Del [fecha] al [fecha]. Canal [chat de la web / correo / WhatsApp]. Horario en que responde la IA [todo el día / fuera de horario]. Clientes [todos / un tipo de cliente].

Temas que lleva la IA [lista]. Temas que pasan siempre a una persona [lista]. Gestiones que la IA puede hacer sola [lista]. Gestiones que esperan a una persona [lista].

## 02 Quién hace qué

| Papel | Persona | Qué hace |
|---|---|---|
| **Decide** | [Nombre] | Aprueba el alcance el día 1 y decide el día 30 |
| **Lleva la prueba** | [Nombre] | Revisa a diario, corrige y prepara el informe |
| **Configura la IA** | [Nombre] | Documentación, directrices y respuestas guiadas |
| **Técnico** | [Nombre] | Acciones sobre la API y sus permisos |

| Papel | Persona | Qué hace |
|---|---|---|
| Legal | [Nombre] | Valida qué datos ve la IA, si hay datos sensibles |

## 03 Semana 1. Preparar sin clientes

- [ ] Sacar treinta preguntas reales de las conversaciones de los últimos tres meses.
- [ ] Escribir la respuesta correcta y la fuente de cada una.
- [ ] Escribir las directrices y los temas que pasan siempre a una persona.
- [ ] Marcar qué gestiones esperan a una persona.
- [ ] Pasar las treinta preguntas y anotar el resultado.
- [ ] Corregir documentación y directrices y volver a pasarlas.
- [ ] Medir la línea base del equipo (resolución, CSAT y tiempos del último mes).

## 04 Semana 2. Un canal, revisión diaria

- [ ] Encender la IA solo en el canal acordado.
- [ ] Leer cada día todas las conversaciones que pasaron a una persona.
- [ ] Leer cada día una muestra de las que resolvió la IA.
- [ ] Anotar cada fallo en el registro, con su tipo y su arreglo.
- [ ] Arreglar en el sitio correcto y repetir las preguntas de prueba.

## 05 Semana 3. Corregir lo que se repite

- [ ] Agrupar los fallos por tipo. Falta documentación, falta un dato del cliente, una directriz mal escrita o una acción que falla.
- [ ] Escribir los artículos de las preguntas que la IA no supo responder.
- [ ] Revisar los traspasos por motivo y ajustar las reglas de escalado.
- [ ] Repetir las preguntas de prueba después de cada cambio.

## 06 Semana 4. Medir y decidir

- [ ] Sacar las cuatro métricas de las semanas 2 a 4.
- [ ] Compararlas con la línea base y con los criterios del apartado 9.
- [ ] Preparar una página para quien decide.
- [ ] Reunión de decisión el día 30.

## 07 Preguntas de prueba

| Pregunta real | Tipo | Respuesta correcta y fuente | Semana 1 | Semana 4 |
|---|---|---|---|---|
| [Pregunta] | Frecuente | [Respuesta. Artículo] | | |
| [Pregunta] | Necesita la cuenta | [Respuesta. Dato o acción] | | |
| [Pregunta] | Sin respuesta escrita | Lo dice y pasa a una persona | | |
| [Pregunta] | Tema que va al equipo | Pasa a una persona con un resumen | | |
| [Pregunta] | Pide una persona | Pasa siempre | | |

## 08 Métricas

Resolución, las que la IA cerró sin que entrara nadie del equipo entre todas las que respondió. Traspasos, las que pasaron a una persona, con su motivo. CSAT, el de la IA aparte del del equipo. Revisadas correctas, las correctas entre las que leyó una persona.

| Métrica | Línea base | Semana 2 | Semana 3 | Semana 4 |
|---|---|---|---|---|
| Tasa de resolución de la IA | No aplica | | | |
| Traspasos, por motivo | No aplica | | | |
| CSAT de la IA y del equipo | [CSAT del equipo] | | | |
| Respuestas revisadas correctas | No aplica | | | |

## 09 Criterios para ampliar o parar

Se escriben y se firman el día 1. Las cifras son un ejemplo que la empresa sustituye por las suyas.

| Decisión | Se toma si |
|---|---|
| **Ampliar a otro canal** | Correctas por encima del [90 %] dos semanas seguidas, ningún dato inventado en temas sensibles, CSAT de la IA no peor que el del equipo y todo el que pidió una persona llegó a una |
| **Seguir un mes igual** | Correctas entre el [75 %] y el [90 %], con los fallos concentrados en temas sin documentación |
| **Parar** | Un dato inventado sobre dinero, contrato o datos personales que se repite después de corregirlo, o un cliente que pidió una persona y no llegó |

## 10 Registro de revisión

| Fecha | Conversación | Qué pasó | Tipo de fallo | Arreglo y dónde |
|---|---|---|---|---|
| [fecha] | [enlace] | | | |
| [fecha] | [enlace] | | | |
| [fecha] | [enlace] | | | |

## 11 Decisión del día 30

Decisión [ampliar / seguir / parar]. Siguiente canal o tema [cuál]. Lo que falta por arreglar y quién [lista y nombres]. Próxima revisión [fecha].