



Curso: Inteligencia Artificial – Generación de Prompts

Comisión: 84180

Apellido y Nombre: Abarca Patricia

Proyecto: Asistente IA para Respuesta Automática de Consultas
Frecuentes

Presentación del problema a abordar

El **Instituto de Obra Médico Asistencial (IOMA)** es la obra social pública de la Provincia de Buenos Aires, responsable de administrar la cobertura de salud de los empleados estatales y sus familias. En la regional de IOMA en Saladillo —que nuclea 12 delegaciones y más de 100 empleados— una parte sustancial de la jornada laboral de los agentes se destina a atender consultas internas y repetitivas sobre plazos, requisitos, documentación y circuitos administrativos. Estas consultas se canalizan por correo institucional, mensajería de trabajo, llamadas y notas, y en su mayoría remiten a normativa vigente y procedimientos ya documentados en circulares, protocolos y comunicaciones oficiales.

Aun con la información disponible, se observan **dificultades estructurales y operativas**:

- **Volumen y repetición:** gran parte de las consultas son variaciones de las mismas 10–15 FAQs internas.
- **Dispersión y versionado:** la normativa está repartida en múltiples documentos y cadenas de correo; no siempre es claro cuál es la versión vigente.
- **Búsqueda ineficiente:** localizar el apartado exacto demanda tiempo y depende del conocimiento previo del agente.
- **Conocimiento tácito:** la resolución de dudas descansa en “referentes” informales; cuando no están disponibles, el flujo se demora.
- **Riesgo de inconsistencias:** respuestas distintas para una misma consulta según quién responda.
- **Impacto operativo:** cuellos de botella en horarios pico, re-trabajos y postergación de tareas estratégicas.
- **Impacto en la experiencia interna:** tiempos de espera variables para delegaciones y equipos, con potencial de reclamos y escalaciones internas.

Ejemplos típicos de consulta interna: plazos entre carátula y arribo a Nivel Central; criterios para auditorías; requisitos para autorizaciones (medicación de alto costo, internaciones programadas, prótesis); documentación respaldatoria; y aplicación de convenios.

Indicadores para línea de base (a relevar). La línea de base es la medición inicial (Semana 0) del desempeño actual sin el asistente; servirá como comparación para evaluar las mejoras del MVP.

- Tiempo promedio de respuesta por consulta interna (minutos desde ingreso hasta primera respuesta útil).
- % de re-preguntas en 7 días sobre el mismo tema (misma delegación y FAQ).
- % de respuestas con corrección o ampliación (casos que requirieron ajuste posterior).
- Volumen semanal de consultas por canal (correo, mensajería, teléfono) y, cuando sea posible, por tema/FAQ.
- Cumplimiento de SLA internos: % fuera de SLA y desvío promedio en minutos respecto del objetivo definido.

(Estas mismas métricas se repetirán en Semanas 2–4 para comparar contra la línea de base.)

Alcance del problema: se centra en consultas internas de gestión administrativa y aplicación normativa por parte de agentes IOMA. No sustituye instancias de auditoría médica, decisiones clínicas ni validaciones que requieran intervención profesional especializada. No está dirigido a afiliados ni a prestadores.

**SLA significa Service Level Agreement o Acuerdo de Nivel de Servicio. Es un compromiso medible (entre áreas o con usuarios) sobre qué se entrega y en qué plazo/calidad*

Desarrollo de la propuesta de solución

Enfoque general

Se implementará un sistema de prompts optimizados para un asistente de IA que responda, en lenguaje natural, las consultas internas de agentes sobre procedimientos, plazos y normativa de IOMA. La solución es independiente de la plataforma y puede funcionar en la intranet institucional, en entornos colaborativos o en mensajería interna, con acceso restringido a personal autorizado.

Alcance inicial (MVP – producto mínimo viable): se abordarán 30–40 FAQs internas (plazos, auditorías, autorizaciones, documentación) y se crearán 2–3 diagramas de procesos clave. Este alcance permite probar rápido, medir impacto y ajustar los prompts antes de escalar, sin integraciones complejas.

Criterios de calidad de respuesta (Conjunto de reglas para asegurar que cada respuesta sea correcta, consistente y verificable)

- **Exactitud normativa** (alineada a la versión vigente). Verificar fecha de vigencia y citar la norma/convenio correspondiente para evitar errores.
- **Claridad y brevedad** ($\leq 120-150$ palabras, o pasos numerados cuando aplique). Oraciones simples, sin rodeos; usar pasos si describe un procedimiento.
- **Trazabilidad** (indicar norma/convenio de respaldo cuando corresponda). Incluir referencia explícita (norma, convenio o ID interno) para que otros puedan auditar.
- **Consistencia de estilo** (tono institucional y formal). Usar verbos de obligación (“debe”, “corresponde”), evitar opiniones y jerga.
- **Derivación responsable si falta información o el caso es excepcional** (indicando dato faltante y área interna). Señalar que falta, a quién se deriva y el plazo objetivo (SLA).

Formato de salida esperado (forma en que se debe presentar la respuesta (estructura y estilo) para que sea clara y usable.)

- **Respuesta breve con fuente citada cuando aplique.** Un párrafo que cierre el tema y la línea “Fuente: [norma/ID] (vigente)”.
- **Uso de pasos/checklist para procedimientos o documentación.** Numerar etapas o listar documentos, un ítem por línea.
- **Derivación clara al área correspondiente en casos no cubiertos por la base.** Frase corta del tipo: “Falta X. Derivar a Y por Z. SLA: N horas.”

Base de conocimiento

El asistente utilizará una base única de referencia (single source of truth) con normativa y respuestas estandarizadas, diseñada para claridad, trazabilidad y búsqueda eficiente.

Cobertura (que incluye):

- Normativa interna: circulares y disposiciones de la Dirección Central.
- Protocolos/Manual operativo: guías vigentes para agentes.
- Convenios: acuerdos institucionales (p. ej., IOMA–CECID) y anexos relevantes.
- Comunicaciones internas: notas/oficios a delegaciones con criterios aplicables.
- FAQs internas validadas: pregunta estandarizada + respuesta + fuente.
- Plantillas y diagramas: modelos de respuesta y bases para texto→imagen.
- Referencias públicas oficiales (cuando aporten precisión): boletines/sitios oficiales explícitamente identificados.

Estructura práctica para esta etapa (simple y demostrable):

- Planilla única (Google Sheets/Excel) con una columna Tipo para distinguir Normativa, Protocolo, Convenio, Comunicación, FAQ o Plantilla.

Metadatos mínimos por registro (columnas):

- Título | Tipo | ID/N° | Vigencia (fecha) | Fuente/Enlace | Palabras clave (3–7) | Extracto/Resumen (1–3 líneas).
- Respuesta validada (solo en FAQs; texto breve con plazos y condiciones).

Ejemplo (FAQ):

- Título/Pregunta: Plazo entre carátula y arribo a Nivel Central
- Tipo: FAQ
- Respuesta validada: “El plazo máximo es de 72 horas. La auditoría en terreno debe realizarse dentro de 48 horas o, en su defecto, auditoría telefónica debidamente justificada.”
- Fuente/Enlace: Convenio IOMA–CECID (ID/URL interna)
- Vigencia: 2025-01-01
- Palabras clave: plazo, carátula, Nivel Central, auditoría
- Extracto: Plazo 72 h; auditoría 48 h o telefónica justificada.

Ejemplo (Normativa):

- Título: Circular 123/2025 – Autorizaciones de alto costo
- Tipo: Normativa
- ID/N°: CIR-123-2025
- Vigencia: 2025-03-01
- Fuente/Enlace: Archivo interno / link oficial
- Palabras clave: alto costo, autorización, auditoría
- Extracto: Define documentación mínima y plazos por etapa.

Modalidades de respuesta

Para cumplir con los requisitos del curso y, a la vez, sostener precisión operativa, el asistente trabajará con dos modalidades.

- La modalidad texto→texto prioriza respuestas breves, normadas y accionables.
- La modalidad texto→imagen convierte procesos en visualizaciones simples (flujos, timelines, etc).

En ambos casos, el contenido surge de la base única de referencia.

Nota sobre fuentes :

Todas las respuestas se elaboran exclusivamente con la Base de conocimiento definida en 2.2 (Normativa, Protocolos, Convenios, Comunicaciones y FAQs validadas, con metadatos de vigencia y última actualización). Cuando aporte precisión, se podrá complementar con referencias públicas oficiales explícitamente identificadas (p. ej., boletines/sitios oficiales: dominios .gob.ar, .gov, etc.).

a) Texto → Texto

Propósito: ofrecer respuestas claras, verificables y reutilizables.

Formatos frecuentes: párrafo breve (definición + plazo), pasos numerados (procedimiento) y checklist (documentación). Se aplican los criterios de calidad (exactitud normativa, brevedad, trazabilidad, tono institucional y derivación responsable si algo no consta).

Ejemplo aplicado (Afiliación de conviviente del afiliado obligatorio)

❖ Párrafo breve (≤150 palabras)

Responde en un párrafo breve (≤150 palabras), tono institucional y uso interno, a: ¿Cómo afiliar a un conviviente de un afiliado obligatorio?

Usá exclusivamente la base de conocimiento ya definida (CONTEXTO). Cuando aporte precisión, podés complementar con referencia pública oficial del organismo (sin URL). Si un dato no consta en la base ni en la referencia pública, indicá “No consta en la normativa adjunta” y propone derivación responsable a Afiliaciones (Delegación/Región) — SLA: 24 h hábiles.

Incluí: criterio principal, pasos clave, plazos (si los hubiera) y condiciones.

Cerrá con: “Fuente: base de conocimiento vigente {+ referencia pública oficial [organismo, fecha] si se usó}.”

❖ **Pasos numerados (procedimiento)**

Explica en pasos numerados (máx. 6) el procedimiento para afiliar a un conviviente de un afiliado obligatorio, indicando el plazo por etapa cuando corresponda.

Basarse solo en el CONTEXTO. Si falta un paso o plazo, marca “No consta en la normativa adjunta” y derivá a Afiliaciones (Delegación/Región) – SLA: 24 h hábiles.

Cerrá con: “Fuente: normativa vigente (según contexto) · Ámbito: uso interno · Exclusiones: casos clínicos/PII”.

❖ **Checklist (documentación)**

Generá un checklist de documentación requerida para afiliar a un conviviente de un afiliado obligatorio, un ítem por línea con aclaraciones breves (p. ej., “vigente”, “original/copia”, “legalizado” si aplica).

Usá exclusivamente el CONTEXTO. No infieras documentos: si un requisito no consta, indicálo y derivá a Afiliaciones (Delegación/Región) – SLA: 24 h hábiles.

Cerrá con: “Fuente: normativa vigente (según contexto)”.

(Las llaves {} indican el tramo opcional solo si realmente se usó una referencia pública oficial, y siempre sin URL.)

b) Texto → Imagen

Propósito: reducir ambigüedad y acelerar la comprensión mediante visualizaciones estandarizadas.

Formatos frecuentes: diagrama de flujo (etapas y áreas), timeline (plazos por etapa), mapa de roles y checklist visual.

Fuentes: usar exclusivamente la base de conocimiento ya definida (CONTEXTO). Cuando aporte precisión, se puede complementar con referencia pública oficial (sin URL; indicar organismo y fecha).

Pie de figura recomendado: “Fuente: base de conocimiento vigente{ + referencia pública oficial [organismo, fecha] si se usó} · Uso interno (IOMA)”.

Ejemplo aplicado (Afiliación de conviviente del afiliado obligatorio)

❖ Diagrama de flujo (circuito del trámite)

Diseñá un diagrama de flujo para afiliar a un conviviente de un afiliado obligatorio (uso interno, agentes IOMA) con etapas, áreas intervinientes y plazos por etapa (si los hubiera).

Basate solo en la base de conocimiento ya definida (CONTEXTO); si un dato no consta, marcá “No consta en la normativa adjunta”.

Si el dato fuera imprescindible para avanzar, derivá a Afiliaciones (Delegación/Región) — SLA: 24 h hábiles.

Título: “Afiliación de conviviente — Circuito interno (IOMA)”.

Nota al pie: “Fuente: base de conocimiento vigente{ + referencia pública oficial [organismo, fecha] si se usó} · Uso interno”.

❖ **Timeline de plazos (si aplica)**

Generá un timeline con los plazos del trámite (p. ej., validaciones, alta en sistema), basado exclusivamente en el CONTEXTO.

Cada hito debe indicar plazo y condición (p. ej., “con documentación completa”).

Si falta un plazo, consigná “No consta en la normativa adjunta” y derivá a Afiliaciones (SLA: 24 h hábiles).

Nota: “Plazos sujetos a condiciones del trámite. Fuente: base de conocimiento vigente{ + referencia pública oficial [organismo, fecha] si se usó}”.

❖ **Checklist visual / Mapa de roles**

Crea un checklist visual de requisitos y validaciones o un mapa de roles (Responsable / Revisor / Validador por etapa).

Usá sólo el CONTEXTO; si un rol o requisito no está definido, indicá “No consta en la normativa adjunta” y, si es necesario para continuar, derivá a Afiliaciones (SLA: 24 h hábiles).

Leyenda: “Fuente: base de conocimiento vigente{ + referencia pública oficial [organismo, fecha] si se usó} · Uso interno (IOMA)”.

(Las llaves {} indican el tramo opcional solo si realmente se usó una referencia pública oficial, y siempre sin URL.)

Prueba inicial de prompts

Prompt 1 (texto→texto, párrafo breve).

Responde en un párrafo breve (≤150 palabras), tono institucional y uso interno, a: ¿Cómo afiliar a un conviviente de un afiliado obligatorio?

Usá exclusivamente la base de conocimiento ya definida (CONTEXTO); si algún dato no consta, indicá “No consta en la normativa adjunta” y derivá a Afiliaciones (Delegación/Región) — SLA: 24 h hábiles. Incluí: criterio principal, pasos clave, plazos (si existieran) y condiciones.

Salida de ejemplo (IA).

“Para afiliar a un conviviente de un afiliado obligatorio, el agente debe verificar la convivencia acreditable y reunir la documentación que exija la normativa vigente (identificaciones, constancia de domicilio y formulario/declaración correspondiente). Con la presentación completa, se registra la solicitud, se valida la condición y se gestiona el alta en el sistema. Cuando existan plazos definidos por etapa, se aplican conforme a la norma y se notifica el resultado. Si faltara algún requisito, se solicita subsanación o se deriva a Afiliaciones (SLA: 24 h hábiles). Fuente: base de conocimiento vigente {+ referencia pública oficial [organismo, fecha] si se usó}.”

Prompt 2 (texto→texto, pasos numerados).

Explica en pasos numerados (máx. 6) el procedimiento para afiliar a un conviviente de un afiliado obligatorio, indicando el plazo por etapa cuando corresponda. Basate solo en el CONTEXTO; si un paso o plazo no consta, márcalo y derivá a Afiliaciones (Delegación/Región) – SLA: 24 h hábiles.

Salida de ejemplo (IA).

1. Recepción y carátula del trámite.
2. Verificación de requisitos y documentación (checklist).
3. Validación de la condición de convivencia conforme norma.
4. Alta en sistema y asociación al titular.
5. Notificación y archivo.

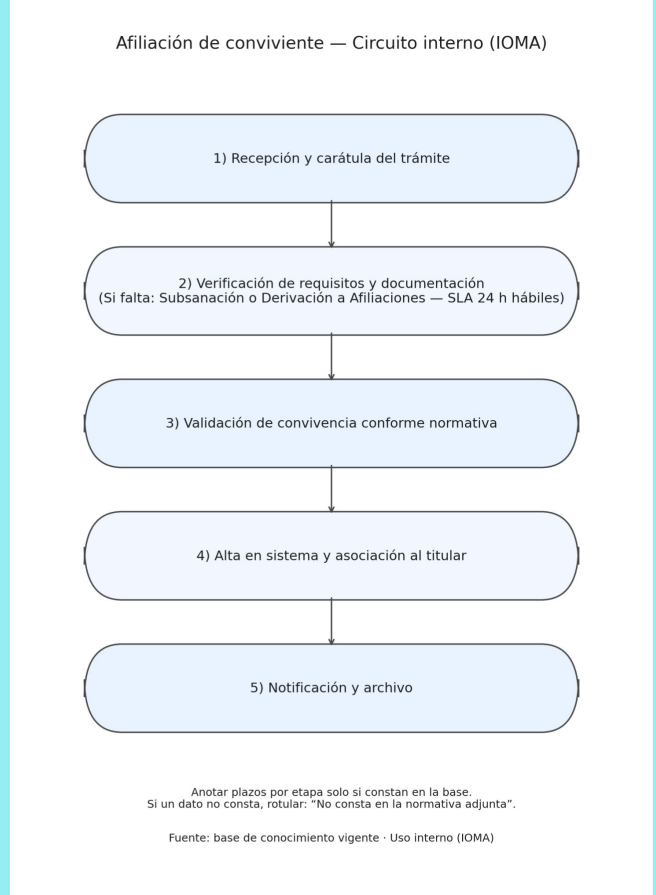
Fuente: base de conocimiento vigente {+ referencia pública oficial [organismo, fecha] si se usó} · Ámbito: uso interno · Exclusiones: criterio clínico/PII.

Ejemplo de Entregable:

Prompt 3 (texto→imagen, diagrama de flujo).

Diseñá un diagrama de flujo del trámite de afiliación de conviviente con etapas, áreas intervinientes y plazos por etapa (si existieran). Usá sólo el CONTEXTO; cuando un dato no conste, indicarlo.

Entregable esperado. Título: “Afiliación de conviviente — Circuito interno (IOMA)”. Pie: “Fuente: base de conocimiento vigente {+ referencia pública oficial [organismo, fecha] si se usó} · Uso interno”.



Camino de escalado (Post MVP y Curso)

Una vez validadas las 30–40 FAQs internas y los 2–3 diagramas (concordancia normativa $\geq 90\%$ y resolución en primer contacto $\geq 70\%$), la cobertura se extenderá de forma gradual a afiliados y prestadores mediante un canal asistido (p. ej., micrositio/FAQ pública o formulario guiado).

Se incorporarán plantillas específicas por público (afiliado/prestador), derivaciones formales (Mesa de Entradas, Delegación, Autorizaciones) y un glosario de términos frecuentes. Se publicará un anexo de privacidad y límites de uso (sin criterio clínico y sin PII; contenido orientativo, no vinculante) con avisos claros de derivación ante casos clínicos o datos sensibles. Las métricas se ampliarán para medir reducción de correos/llamadas externas, porcentaje de autoatención, resolución en primer contacto, tiempo de resolución y satisfacción del usuario (CSAT), con seguimiento quincenal durante el piloto externo.

El despliegue quedará condicionado a base actualizada, mensajes validados y una capacitación breve a delegaciones sobre uso y límites del asistente.

Justificación de la viabilidad del proyecto

Viabilidad Técnica

El proyecto es factible con herramientas gratuitas o educativas y sin infraestructura propia. Para texto→texto se puede usar Chat GPT (plan gratuito/edu) con una base de conocimiento ya definida cargada en una planilla (Google Sheets o Excel) que incluya vigencia y fuente. Para texto→imagen bastan opciones sin costo como diagrams.net/draw.io, Mermaid (diagramas desde texto), Excalidraw o Canva para maquetar infografías.

La solución es independiente de plataforma y puede operar en intranet, en entornos colaborativos (SharePoint/OneDrive o Google Drive) o en mensajería interna (Teams/Slack), con acceso restringido.

Si se requiere una automatización mínima (alta de nuevas FAQs, registro de consultas no cubiertas), alcanzan Power Automate, Zapier/Make o Google Apps Script. Todo esto permite mantener una base única de referencia con trazabilidad sin desarrollar software a medida.

Viabilidad Económica

Durante la etapa del curso, el costo es nulo al priorizar planes gratuitos/educativos y un MVP acotado (30–40 FAQs + 2–3 diagramas). En un uso productivo, el gasto se controla con prompts breves, plantillas reutilizables, reutilización de respuestas para preguntas repetidas y contexto recortado (solo el fragmento relevante de la base).

Se pueden fijar límites de uso por usuario/equipo y monitorear el consumo con la misma planilla.

Como referencia, se propone un tope por consulta \leq USD 0,01; el costo mensual se estima con la fórmula:

$$((tokens_in + tokens_out) / 1000) \times precio_1K \times n^\circ_de_consultas.$$

Aquí, tokens_in es lo que envías por consulta (prompt + contexto) y tokens_out lo que responde el modelo; precio_1K es el valor por 1.000 tokens y n°_de_consultas es cuántas preguntas harás. Ejemplo rápido: si por consulta usás 250 tokens de entrada y 300 de salida (total 550 \rightarrow 0,55 “miles”), y el precio es USD 0,02 por 1.000 tokens, cuesta $0,55 \times 0,02 =$ USD 0,011 por consulta; con 2.000 consultas, serían USD 22.

Con estas prácticas y el uso de herramientas visuales gratuitas para los diagramas, el proyecto se mantiene predecible y económico.

Viabilidad en Tiempo

MVP (Producto Mínimo Viable) en 4 semanas

- ➔ Semana 1 — Base de conocimiento. Curado de 30–40 FAQs internas, estructura de planilla (campos y metadatos de vigencia/fuente), carga inicial y pequeño glosario.
- ➔ Semana 2 — Contenidos del asistente. Redacción de prompts texto→texto (plantillas y estilo) y 2–3 diagramas texto→imagen; criterios de calidad y formato de salida.
- ➔ Semana 3 — Pruebas y ajustes. Piloto con usuarios internos (2–3 delegaciones), registro de incidencias, ajustes de contenido y de prompts.
- ➔ Semana 4 — Cierre operativo. Documentación breve de uso (playbook), capacitación express (30') y tablero simple de métricas (línea de base vs. post-MVP).

Supuestos y alcance

Cubre: consultas internas de agentes IOMA (gestión administrativa, plazos, documentación y aplicación normativa) basadas en la base de conocimiento ya definida.

No cubre: atención a afiliados o prestadores, criterio/decisiones clínicas y tratamiento de PII o datos de salud.

Condiciones de operación: horario hábil, documentación completa cuando el proceso lo requiera; referencias públicas oficiales solo para complementar lo ya presente en la base.

Métricas de éxito del MVP (objetivos y medición)

- Concordancia normativa $\geq 90\%$ (muestra auditada contra base).
- Tiempo de respuesta \downarrow 30–50% vs. línea de base (Semana 0).
- Re-preguntas $\leq 15\%$ sobre el mismo tema en 7 días.
- Cobertura lograda: ≥ 30 –40 FAQs + 2–3 diagramas publicados.
- Resolución en primer contacto $\geq 70\%$ (para las FAQs priorizadas).
- Cumplimiento de SLA interno $\geq 80\%$ (p. ej., respuesta ≤ 60 min hábiles).

Medición: línea de base en Semana 0; seguimiento en Semanas 2–4 con planilla (canal, tema, tiempo, re-pregunta, SLA).

Riesgos y mitigaciones

Riesgos = posibles eventos que pueden afectar alcance, tiempo, costo o calidad del proyecto.

Mitigaciones = acciones preventivas/correctivas para reducir probabilidad de que ocurran o bajar su impacto si pasan.

- Cambios normativos. Mitigar con metadatos de vigencia/fecha, control de versiones y actualización rápida de FAQs afectadas (marcar la anterior como “no vigente”).
- Ambigüedad / faltan datos. Responder lo seguro, hacer una repregunta concreta o derivar al área responsable con SLA; registrar el caso para mejorar la base.
- Calidad / consistencia de fuentes. Priorizar documentos oficiales; si no hay evidencia, indicar “No consta en la normativa adjunta” y evitar suposiciones.
- Adopción del usuario. Mini-capacitación de 30', plantillas claras y ejemplo de buenas prácticas; canal para feedback y ajustes rápidos.
- Volumen mayor al previsto. Definir límites de uso, reutilizar respuestas frecuentes y priorizar FAQs críticas; escalar contenidos por etapas si es necesario.



Impacto esperado

Reducción de tiempos de respuesta.

Atención casi inmediata a consultas recurrentes y menor tiempo operativo del equipo. Meta MVP: ↓ 30–50% del tiempo promedio por consulta vs. línea de base (Semana 0), con nueva medición en Semanas 2–4.

Uniformidad y trazabilidad de la información.

Respuestas alineadas a normativa vigente y siempre con fuente citada. Meta MVP: ≥ 90% de concordancia normativa (muestra auditada) y 100% de respuestas con fuente. Retrabajos por corrección: ↓ ≥ 50%.

Optimización de la comunicación interna.

Menor volumen de correos/llamadas por FAQs y menos re-preguntas. Meta MVP: ↓ 30–50% de consultas por canales tradicionales (email/teléfono) en tópicos priorizados y ≥ 70% de resolución en primer contacto del asistente.

Liberación de carga operativa.

Más tiempo para tareas de mayor valor (casos complejos, seguimiento de SLA, capacitación). Meta MVP: horas recuperadas/semana estimadas como

Horas recuperadas \approx ($\#$ de consultas resueltas \times minutos ahorrados) / 60. Objetivo orientativo: 8–12 h/semana por región.

Cómo se medirá el impacto

- Línea de base (Semana 0): tiempo promedio por consulta, $\#$ de consultas por canal, % de re-preguntas.
- Post-MVP (Semanas 2–4): mismas métricas + % de consultas resueltas por el asistente y concordancia normativa (muestra auditada).