



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—, el área de Afiliaciones concentra gran parte de las consultas internas sobre altas, bajas y modificaciones de afiliados obligatorios, voluntarios y colectivos, junto con sus grupos familiares. Los agentes reciben de manera recurrente preguntas sobre documentación, plazos, circuitos internos y criterios de aplicación normativa, que ya se encuentran descriptos en circulares y protocolos, pero cuya consulta resulta poco ágil en la práctica.

Se identifican dificultades específicas:

- Volumen y repetición: gran parte de las consultas son variaciones de las mismas 8–10 FAQs (altas de convivientes, incorporación de hijos, bajas por fallecimiento, cambios de condición, incorporación de voluntarios o colectivos).
- Dispersión de normativa: las circulares y notas internas están distribuidas en múltiples correos o archivos, lo que complica ubicar la versión vigente.
- Búsqueda ineficiente: localizar requisitos exactos (ej. documentación mínima) demanda tiempo y experiencia previa del agente.
- Dependencia de referentes: la resolución descansa en personal con mayor trayectoria; su ausencia genera demoras.
- Riesgo de inconsistencias: respuestas distintas para un mismo caso según quién atienda.
- Impacto operativo: tiempos de espera variables y postergación de trámites sensibles para afiliados.

Ejemplos típicos de consulta interna en Afiliaciones refieren a los plazos y pasos para afiliarse a un conviviente, los requisitos para incorporar hijos menores o estudiantes, la documentación exigida ante cambios de estado civil o domicilio, el procedimiento para dar de baja a un afiliado por fallecimiento, así como los criterios y la documentación necesaria para gestionar las altas de afiliados voluntarios y colectivos.

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 y plantillas estandarizadas para un asistente de IA que responda, en lenguaje natural, las consultas internas del personal de Afiliaciones sobre altas, bajas y modificaciones de afiliados obligatorios, voluntarios y colectivos, incluyendo sus grupos familiares. La solución podrá ser utilizada en cualquier plataforma (intranet, entornos colaborativos o mensajería interna), con acceso restringido a personal autorizado.

El asistente opera sobre una base única de referencia con control de versiones (normativa y procedimientos de Afiliaciones con vigencia, fuente e ID), y utiliza ampliación de contexto controlada (búsqueda granular de fragmentos relevantes) para cada respuesta. Este enfoque asegura exactitud normativa, trazabilidad total de fuentes y una política explícita: cuando un dato no consta en la base, se lo indica y se deriva responsablemente al área correspondiente con SLA.

Alcance inicial (MVP – producto mínimo viable): Se priorizará 30–40 FAQs de Afiliaciones (p. ej., alta de conviviente, incorporación de hijos, cambios de estado civil o domicilio, baja por fallecimiento, altas de voluntarios y colectivos) y se diseñarán 2–3 diagramas de procesos clave (flujo de alta, timeline de plazos y checklist de documentación). El MVP permitirá probar rápido, medir impacto frente a la línea de base (tiempos de respuesta, re-preguntas, concordancia normativa) y ajustar contenidos y 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)

1- Exactitud normativa (BASE obligatoria para “checklist”)

- Todo ítem del campo "checklist" debe salir **exclusivamente** de <<BASE>>.
- Si un dato no figura en la BASE, usar la leyenda **“No consta en la normativa adjunta”** (sin inventar).
- El "cierre" debe mantener la trazabilidad: **“Fuente: base de conocimiento vigente{ + ID/Nº si consta}”**.

2- Claridad y brevedad (límites por campo)

- "checklist": **3–8 ítems**, cada ítem ≤ **18 palabras**.
- "terminos_clave": **3–6 entradas** con formato **“Término: definición breve”**, cada una ≤ **16 palabras**.
- "objetivo_y_buenas_practicas": **3–6 entradas**, cada una ≤ **20 palabras**.
- Español, tono institucional, oraciones simples; sin jerga.

3- Trazabilidad y consistencia

- Priorizar siempre la BASE cuando exista contradicción con conocimiento externo.
- Mantener estilo uniforme (verbos de obligación: “debe”, “corresponde”, “verificar”).
- Evitar opiniones, supuestos y PII.

4- Derivación responsable (política explícita)

- Cuando aparezca **“No consta...”**, el "cierre" debe incluir derivación y **SLA**:
“Fuente: base de conocimiento vigente · Derivar a Afiliaciones (Delegación/Región) — SLA: 24 h hábiles.”

5- Ampliación de contexto controlada

- "terminos_clave" y "objetivo_y_buenas_practicas" pueden usar conocimiento general **solo** si **no contradice** la BASE.
- Si hay duda, prevalece la BASE.

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

```
{  
  "checklist": ["...", "..."],  
  "terminos_clave": ["Término: definición breve", "...", "..."],  
  "objetivo_y_buenas_practicas": ["Buena práctica: detalle breve", "...", "..."],  
  "cierre": "Fuente: base de conocimiento vigente"  
}
```

Base de conocimiento

El asistente de Afiliaciones trabajará sobre una base única de referencia, donde el eje central lo constituyen las FAQs internas validadas (pregunta estandarizada + respuesta + fuente). Estas FAQs concentran los casos más frecuentes de altas, bajas y modificaciones de afiliados obligatorios, voluntarios y colectivos, junto con su grupo familiar.

La ampliación de contexto será posible únicamente de manera controlada, incorporando fragmentos relevantes de normativa, protocolos, convenios o comunicaciones internas cuando aporten precisión. En situaciones puntuales, también se podrán usar referencias públicas oficiales (p. ej., boletines o sitios .gob.ar), siempre de forma explícita y complementaria a la base principal.

Este enfoque asegura que cada respuesta parta de un núcleo seguro y validado (FAQs), con posibilidad de enriquecerse con contexto adicional para sostener la exactitud normativa, la trazabilidad de la fuente y la derivación responsable cuando un dato no conste.

Estructura práctica para esta etapa (simple y demostrable):

Planilla única en formato CSV/Excel, que concentra toda la información de Afiliaciones.

Columna **Tipo** para distinguir entre: FAQ, Normativa, Convenio, Comunicación, Protocolo o Plantilla.

El **núcleo central** son las **FAQs validadas**, y el resto de categorías funcionan como **contexto de apoyo** cuando aporta precisión.

Metadatos mínimos por registro (columnas):

- **Título / Pregunta**
- **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)

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 para las consultas más frecuentes del área de **Afiliaciones**.

Formato de salida esperado:

- **Checklist:** listado breve de requisitos o pasos, construido exclusivamente a partir de la base de conocimiento. Si un dato no consta, debe aparecer la leyenda “No consta en la normativa adjunta”.
- **Términos clave:** definiciones breves de los conceptos principales vinculados a la consulta (ejemplo: conviviente, vigencia, grupo familiar).
- **Objetivo y buenas prácticas:** recomendaciones concretas para asegurar la correcta aplicación de la normativa, el uso de la base y la derivación responsable.
- **Cierre:** siempre incluir la referencia de trazabilidad —“Fuente: base de conocimiento vigente”— y, en caso de que falte información, la derivación correspondiente a Afiliaciones con el plazo de SLA definido.

Criterios aplicados:

- **Exactitud normativa:** usar únicamente la base de conocimiento y su contexto controlado.
- **Brevedad:** ítems concisos, definiciones cortas, sin rodeos.
- **Trazabilidad:** toda salida debe cerrar con la fuente.
- **Tono institucional:** oraciones formales, uso de verbos de obligación.
- **Derivación responsable:** cuando un dato no consta, indicar derivación explícita a Afiliaciones (Delegación/Región) con SLA: 24 h hábiles.

b) Texto → Imagen

Propósito: reducir la ambigüedad y acelerar la comprensión de los trámites de **Afiliaciones** mediante visualizaciones estandarizadas, basadas en la base de conocimiento.

Formatos de salida esperados:

Diagrama de flujo: muestra etapas del trámite, áreas intervinientes y plazos (cuando consten).

- **Timeline:** visualiza los plazos de cada hito del trámite (ej. validación, alta en sistema), indicando condiciones asociadas.
- **Checklist visual o Mapa de roles:** lista de requisitos/documentación o identificación de responsables, revisores y validadores por etapa.

Criterios aplicados:

- Usar **exclusivamente la base de conocimiento** como fuente principal.
- **Ampliar contexto solo de manera controlada**, con normativa o referencias oficiales explícitas (sin URL, indicando organismo y fecha).
- Cuando un dato **no conste**, debe marcarse explícitamente con la leyenda *“No consta en la normativa adjunta”*.
- Si el dato fuera imprescindible para completar el flujo, se debe **derivar a Afiliaciones (Delegación/Región) con SLA: 24 h hábiles**.

Trazabilidad en la visualización:

- **Título:** debe ser claro, indicando el trámite y su ámbito (ej. “Afiliación de conviviente – Circuito interno”).
- **Nota al pie / leyenda:** siempre incluir *“Fuente: base de conocimiento vigente{ + referencia pública oficial [organismo, fecha] si se usó} · Uso interno (IOMA)”*.
- En timelines de plazos, aclarar: *“Plazos sujetos a condiciones del trámite”*.

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.

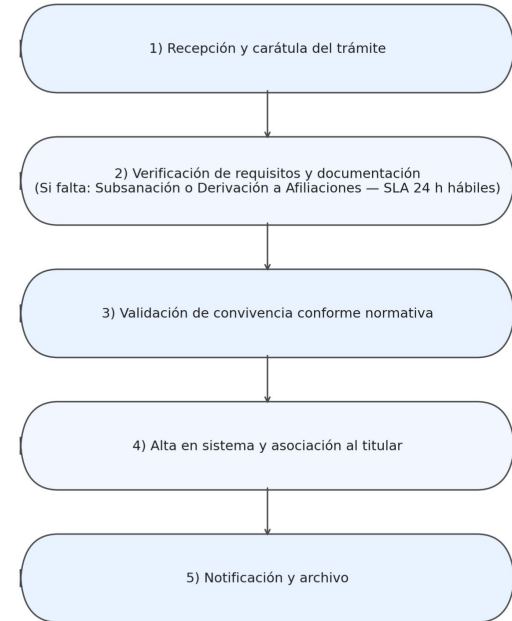
Ejemplo de Entregable:

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”.

Afiliación de conviviente — Circuito interno (IOMA)



Anotar plazos por etapa solo si constan en la base.
Si un dato no consta, rotular: “No consta en la normativa adjunta”.

Fuente: base de conocimiento vigente · Uso interno (IOMA)

Camino de escalado (Post MVP y Curso)

Una vez validadas las 30–40 FAQs de Afiliaciones y los 2–3 diagramas clave (concordancia normativa $\geq 90\%$ y resolución en primer contacto $\geq 70\%$), la cobertura podrá extenderse gradualmente a otros temas de gestión interna de IOMA, manteniendo siempre el alcance a agentes y no a afiliados ni prestadores.

La ampliación incluirá:

- Nuevas FAQs internas de otras áreas administrativas (ej. autorizaciones, auditorías, documentación general).
- Plantillas específicas por área (Delegación, Autorizaciones, Mesa de Entradas).
- Glosario de términos frecuentes para unificar lenguaje institucional.
- Anexo de privacidad y límites de uso, recordando que no cubre criterio clínico ni datos sensibles, y que el contenido es orientativo, no vinculante.

Las métricas también se ampliarán para medir:

- Reducción de consultas repetitivas por correo/teléfono.
- Porcentaje de autoatención interna.
- Resolución en primer contacto.
- Tiempo de resolución promedio.
- Satisfacción del agente (CSAT interno).
- El despliegue quedará condicionado a contar con una base actualizada, mensajes validados y una capacitación breve a las delegaciones sobre uso y límites del asistente.

Justificación de la viabilidad del proyecto

Viabilidad Técnica

El proyecto es factible utilizando herramientas gratuitas o educativas, sin necesidad de infraestructura propia ni desarrollo a medida.

Texto → Texto: se implementa con modelos como ChatGPT (plan gratuito/edu), alimentados con una base de conocimiento en formato CSV que contiene las FAQs validadas de Afiliaciones con sus metadatos de vigencia, fuente e identificador. La salida estructurada en JSON garantiza consistencia, trazabilidad y derivación responsable cuando falta información.

Texto → Imagen: para generar diagramas y visualizaciones simples (flujos, timelines, checklists visuales) se emplean herramientas gratuitas como diagrams.net/draw.io, Mermaid, Excalidraw o Canva, utilizando los mismos datos del CSV como insumo.

Independencia de plataforma: la solución puede ejecutarse en intranet, entornos colaborativos (SharePoint/OneDrive o Google Drive) o en mensajería interna (Teams/Slack), con acceso restringido a personal autorizado.

Automatización mínima: para registrar nuevas FAQs o consultas no cubiertas alcanza con usar Power Automate, Zapier/Make o Google Apps Script, que permiten actualizar el CSV de manera controlada y mantener la trazabilidad de los cambios.

Este esquema asegura que el asistente funcione de forma segura, escalable y sostenible, priorizando la base en CSV como núcleo y permitiendo complementarse con normativa o contexto sólo cuando corresponda.

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 escenario productivo, el gasto se mantiene bajo y predecible gracias a:

Prompts compactos (\approx 336 tokens promedio).

Respuestas controladas (\approx 250 tokens máximo).

Uso del CSV como base, lo que evita enviar grandes volúmenes de contexto.

Reutilización de plantillas y respuestas frecuentes.

Monitoreo de consumo en la misma planilla (FAQ consultada, tokens usados, costo).

Costo real medido:

Una consulta cuesta aprox. USD 0.0002.

Con 100–120 consultas semanales (\approx 400–480 al mes), el costo mensual estimado es de USD 0.08 a USD 0.10.

Esto se encuentra muy por debajo del tope de referencia (\leq USD 0.01 por consulta) y permite sostener el proyecto de manera económica y escalable, sin depender de infraestructura propia ni licencias costosas.

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 de **Afiliaciones de IOMA**.
- Temas de gestión administrativa: **altas, bajas y modificaciones** de afiliados obligatorios, voluntarios y colectivos, incluyendo sus grupos familiares.
- Plazos, documentación y aplicación normativa, siempre basados en la **base de conocimiento en CSV** (FAQs validadas como núcleo).

No cubre:

- Atención directa a afiliados o prestadores.
- Criterio o decisiones clínicas.
- Tratamiento de datos sensibles o PII vinculados a la salud.

Condiciones de operación:

- Respuestas en **horario hábil**.
- Procesos condicionados a la **documentación completa** cuando corresponda.
- Referencias públicas oficiales (ej. sitios .gob.ar) se utilizan **solo como complemento**, nunca como sustituto de la base.

Métricas de éxito del MVP

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

Medición:

- Línea de base en **Semana 0** (tiempo de respuesta, volumen de consultas, re-preguntas).
- Seguimiento en **Semanas 2–4**, con planilla que registre canal, tema, tiempo, re-pregunta y cumplimiento de SLA.

Riesgos y mitigaciones

Riesgo: Cambios normativos

- *Impacto:* puede invalidar respuestas ya publicadas.
- *Mitigación:* mantener metadatos de vigencia y fecha, control de versiones y actualización rápida de las FAQs afectadas, marcando la anterior como “no vigente”.

Riesgo: Ambigüedad o falta de datos en la base

- *Impacto:* respuestas incompletas o inconsistentes.
- *Mitigación:* responder solo lo seguro, marcar “No consta en la normativa adjunta”, derivar al área responsable con SLA y registrar el caso para mejorar la base.

Riesgo: Calidad y consistencia de fuentes

- *Impacto:* errores o contradicciones en la información.
- *Mitigación:* priorizar siempre documentos oficiales; evitar suposiciones cuando no exista evidencia suficiente.

Riesgo: Adopción del usuario

- *Impacto:* baja utilización del asistente por parte de los agentes.
- *Mitigación:* capacitación breve de 30', plantillas claras, ejemplos de buenas prácticas y canal de feedback para ajustes rápidos.

Riesgo: Volumen de consultas mayor al previsto

- *Impacto:* incremento de costos o saturación del sistema.
- *Mitigación:* fijar límites de uso por usuario/equipo, reutilizar respuestas frecuentes, priorizar FAQs críticas y escalar contenidos por etapas.



Impacto esperado

- **Reducción de tiempos de respuesta**

Atención casi inmediata a consultas recurrentes, con menor carga operativa 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 reducción de re-preguntas. **Meta MVP:** ↓ 30–50% de consultas en canales tradicionales (correo/teléfono) en tópicos priorizados y ≥ 70% de resolución en primer contacto.

- **Liberación de carga operativa**

Más tiempo disponible para tareas de mayor valor (casos complejos, seguimiento de SLA, capacitación). **Meta MVP:** horas recuperadas por semana estimadas como: Horas recuperadas ≈ (# de consultas resueltas × minutos ahorrados) / 60.

Objetivo orientativo: 8–12 horas/semana por región.

Cómo se medirá el impacto

En la **línea de base (Semana 0)** se registrará el tiempo promedio por consulta, el número de consultas por canal y el porcentaje de re-preguntas.

En el **post-MVP (Semanas 2–4)** se medirán esas mismas métricas junto con el porcentaje de consultas resueltas por el asistente y la concordancia normativa sobre una muestra auditada.