

# Primera evidencia de Aprendizaje

Materia: Práctica Profesionalizante

**Nombre completo:** Florencia Carolina Andrada Valentini

**DNI:**44788525

**Carrera:** Tec en ciencias de datos e IA.

**Comision:** A

**Repositorio:** <https://github.com/Flor3ncia-Andr4d4/EV1-P.PROFECIONALIZANTE>

**Nombre del docente:** Accietto Dianela.

# Contexto del Cliente

## Primer vistazo a problemas críticos

El principio del problema radica en la implementación de un chatbot de IA generativa denominado *MediBot*, diseñado para automatizar la atención inicial de pacientes y la gestión de turnos. A pesar de contar con ocho meses de desarrollo, el sistema no logró resolver los principales *pain points* identificados, generando mayores complicaciones operativas y afectando de manera negativa la experiencia del paciente.

## Situación de la clínica

La Clínica MediCare es un centro médico privado ubicado en Córdoba, con una plantilla de 150 empleados. En el área de atención al paciente se registraban las siguientes condiciones:

- Recepción de más de 800 llamadas diarias para consultas y turnos.
- Tiempo de espera promedio de 12 minutos.
- Solo 3 operadoras asignadas de forma exclusiva a la atención telefónica.
- El 85% de las consultas eran repetitivas (horarios, ubicación, preparación de estudios).
- Alta rotación del personal de *call center* debido al nivel de estrés laboral.

## Análisis

Del análisis de la situación se desprende que el área de atención telefónica presentaba una cobertura insuficiente. Con 150 empleados en total y únicamente 3 dedicados al *call center*, la cobertura frente al volumen de llamadas (800 diarias) era del 18,75%. Esto implica que, en promedio, cada empleado debía responder 5,3 llamadas en simultáneo, lo que resulta insostenible.

Este desbalance se reflejaba en un tiempo de espera promedio de 12 minutos, afectando directamente la calidad del servicio y la satisfacción del usuario. Además, se observaron los siguientes problemas críticos:

- Dotación de personal inadecuada para la demanda existente.
- Elevada proporción de consultas repetitivas que podrían ser automatizadas.

- Alta rotación de personal por estrés, generando costos adicionales de capacitación y pérdida de conocimiento operativo.

En conjunto, estos factores evidencian una estructura deficiente que impactó negativamente tanto en la eficiencia operativa como en la experiencia de los pacientes.

## Objetivos vs. Resultados

### Análisis de objetivos vs. resultados

El proyecto estableció cuatro objetivos principales: reducir el tiempo de espera, automatizar consultas básicas, disminuir los costos operativos y mejorar la satisfacción del paciente. Sin embargo, los resultados obtenidos evidencian que ninguno de los objetivos fue alcanzado y que, en algunos casos, los indicadores mostraron un empeoramiento significativo.

En primer lugar, el tiempo de espera promedio, que debía reducirse a menos de 3 minutos, se incrementó a 15 minutos, reflejando que la atención se tornó más lenta que antes de la implementación. La automatización de consultas básicas solo alcanzó un 23%, muy por debajo del 70% proyectado, lo que demuestra que la estrategia tecnológica aplicada fue insuficiente.

Respecto a los costos operativos, en lugar de disminuir un 40% como se esperaba, aumentaron un 15%, lo que indica una gestión ineficiente de los recursos. Finalmente, la satisfacción de los pacientes, medida a través del *Net Promoter Score* (NPS), descendió de la meta planteada de 8.5/10 a un 4.2/10, mostrando un impacto negativo directo en la experiencia del usuario.

## Análisis Técnico

### Tecnología implementada

El sistema *MediBot* fue desarrollado como un chatbot personalizado en React, con integración mediante APIs REST al sistema de gestión hospitalaria (HIS). La base de conocimiento utilizada se limitó a 500 preguntas frecuentes (FAQs). El modelo de IA empleado fue GPT-3.5 Turbo, con un entrenamiento básico.

### 1. Capacidades insuficientes del chatbot

El modelo se encontraba entrenado con una base reducida y generalista, sin incluir terminología médica especializada ni contemplar escenarios complejos. Esto impedía dar respuesta a consultas específicas o inusuales.

Además, la ausencia de memoria de contexto provocaba que cada interacción fuera independiente, dificultando la continuidad en conversaciones largas o de seguimiento.

## **2. Integración parcial con el sistema**

El chatbot carecía de acceso a historiales médicos y a la disponibilidad real de turnos, lo que limitaba su utilidad práctica. En consecuencia, muchas interacciones debían derivarse nuevamente a los operadores humanos.

La falta de un mecanismo de escalamiento inteligente generaba frustración en los usuarios, ya que no existía claridad sobre cuándo podían ser atendidos por personal humano.

## **3. Limitaciones en comprensión del lenguaje**

El procesamiento de lenguaje natural (PLN) resultó insuficiente para reconocer expresiones locales, jerga cordobesa o variaciones en las consultas cotidianas. Esta limitación redujo significativamente la naturalidad y efectividad de la interacción.

## **4. Impacto en los resultados**

Las limitaciones técnicas señaladas se tradujeron en consecuencias directas sobre los indicadores clave del proyecto: incremento en los tiempos de espera, bajo nivel de automatización, insatisfacción del paciente y sobrecarga del personal humano.

En síntesis, la tecnología aplicada no estaba preparada para abordar el volumen ni la complejidad de las consultas en el ámbito sanitario, lo cual se reflejó en el fracaso general del proyecto.

# **Timeline de Implementación**

## **Análisis del *timeline* de implementación (con prueba Beta destacada)**

**Kick-off y desarrollo inicial (enero – marzo 2024).** Durante esta etapa se definieron los requerimientos básicos del sistema y se desarrolló un MVP (*Producto Mínimo Viable*) con una base inicial de 200 preguntas frecuentes. Las pruebas realizadas fueron exclusivamente internas, sin participación de usuarios reales, lo que limitó la capacidad de detectar problemas de usabilidad o de comprensión del chatbot.

**Integración con HIS (abril 2024).** En abril, se integró el chatbot con el sistema hospitalario (HIS). Sin embargo, surgieron problemas de compatibilidad que restringieron su funcionamiento. La conexión resultó incompleta y no permitió acceder a información crítica, como la disponibilidad de turnos en tiempo real.

**Prueba Beta (mayo 2024).** En esta fase se realizaron pruebas con un grupo piloto de 50 pacientes. Los resultados mostraron baja precisión en las respuestas, errores frecuentes y limitaciones técnicas significativas. Esta instancia fue clave, ya que permitió identificar que el chatbot no estaba preparado para un lanzamiento masivo. Sin embargo, los hallazgos críticos no fueron corregidos de manera adecuada, constituyendo un punto de inflexión en el proyecto.

**Lanzamiento oficial (junio 2024).** Pese a las señales de alerta detectadas en la fase Beta, el sistema fue puesto en producción a gran escala. El lanzamiento oficial ocasionó una sobrecarga inmediata del sistema, un aumento en las quejas de los usuarios y una degradación significativa en la calidad del servicio. Este hecho demuestra que la decisión de avanzar al *go-live* fue prematura y se realizó ignorando riesgos conocidos.

**Crisis y desactivación parcial (julio – agosto 2024).** Durante los meses posteriores, los problemas se intensificaron y no fueron resueltos a tiempo. La clínica se vio obligada a reducir las funcionalidades del chatbot y limitar su uso, con el objetivo de evitar mayores perjuicios a la experiencia del paciente. Esta decisión marcó la desactivación parcial del proyecto, a la espera de soluciones técnicas más robustas.

## Problemas Críticos Identificados

### Análisis de problemas críticos identificados

El conjunto de problemas detectados refleja que el proyecto no estaba preparado para operar de manera segura, confiable ni eficiente. Los impactos pueden clasificarse en tres grandes categorías:

#### 1. Riesgos para la seguridad y el cumplimiento legal

- La falta de encriptación en el almacenamiento de datos sensibles representó una violación de los principios básicos de privacidad y seguridad de la información.
- La entrega de información médica incorrecta o desactualizada puso en riesgo la salud de los pacientes y expuso a la clínica a posibles responsabilidades legales.
- La ausencia de validación profesional sobre las respuestas del chatbot reflejó la carencia de un control de calidad esencial en un entorno sanitario.

## 2. Limitaciones técnicas que afectaron la operación

- El sistema presentaba una latencia elevada, con tiempos de respuesta de entre 8 y 12 segundos, lo que redujo la fluidez de la interacción.
- Se registraron caídas frecuentes del sistema (15 interrupciones en el primer mes), afectando la continuidad del servicio.
- La desincronización en la información de turnos generó errores en la gestión de citas, provocando confusión y retrasos adicionales.

## 3. Deficiencias en la experiencia del usuario

- La interfaz resultaba poco amigable para adultos mayores, quienes representaban el 65% de la base de pacientes.
- El chatbot no estaba capacitado para reconocer situaciones de urgencia médica ni para derivarlas a un profesional humano, lo que implicaba un riesgo crítico en contextos sensibles.
- Las respuestas rígidas y robóticas generaron frustración en los usuarios, especialmente en consultas relacionadas con temas delicados de salud.

# Impacto en el Cliente

## Análisis del impacto en el cliente

Los indicadores muestran que la implementación de *MediBot* tuvo un efecto negativo significativo sobre la experiencia y la confianza de los pacientes.

1. **Disminución de la satisfacción del cliente (-47%).** La caída en la satisfacción, junto con la pérdida de 340 pacientes que cambiaron de prestador médico, evidencia que los usuarios no confiaron en el servicio debido a errores, demoras y respuestas inexactas del chatbot.
2. **Incremento de costos operativos (+15%).** En lugar de generar ahorros, el proyecto ocasionó costos adicionales de aproximadamente USD 23.000 en soporte técnico y recursos humanos, derivados de problemas técnicos, escalamiento de llamadas y mantenimiento del sistema.

3. **Aumento del tiempo de espera (+25%).** El tiempo de espera se incrementó, reflejando que la automatización no alivió la carga de trabajo del personal. Por el contrario, las operadoras humanas debieron atender un 60% más de llamadas derivadas, lo que incrementó la sobrecarga laboral.
4. **Reducción de nuevos pacientes (-12%).** La disminución en la captación de nuevos pacientes evidencia que la mala experiencia de usuario y la reputación online afectaron negativamente la percepción del servicio, como se observa en la caída de la calificación de Google de 4.3 a 2.8 estrellas.
5. **Riesgo legal.** Se registraron tres reclamos formales por información médica incorrecta, evidenciando que los fallos del chatbot no solo afectaron la experiencia del paciente, sino que también generaron responsabilidades legales para la clínica.

## Investigación Profunda del Caso

### 1. Causas raíz del fracaso

#### Aspectos técnicos

Al inicio, la clínica enfrentaba una alta demanda de llamadas que no fue gestionada adecuadamente. Aunque contaba con 150 empleados, solo tres secretarías atendían las 800 llamadas diarias, lo que generaba una cobertura de apenas 18,75%. Esta insuficiencia de personal evidencia deficiencias en la organización técnica inicial. Para intentar mejorar la atención, se contactó a desarrolladores y se decidió implementar un chatbot de inteligencia artificial (IA).

#### Aspectos organizacionales

Desde la perspectiva organizacional, la falta de planificación y coordinación en la gestión de llamadas era notoria. Esta carencia generó gastos innecesarios y redujo la eficiencia del servicio. La organización no priorizó la estructura administrativa ni la asignación adecuada de recursos humanos frente a la alta demanda de consultas.

#### Gestión de proyecto

La gestión del proyecto mostró deficiencias en la administración de empleados y recursos, así como en la planificación de tiempos y tareas. No se definió una hoja de ruta clara para la implementación del chatbot ni se asignaron responsabilidades de manera efectiva, afectando el desempeño global del proyecto.

#### Factores humanos

- **Dotación insuficiente de personal:** La concentración de llamadas en solo tres operadoras aumentó la carga de trabajo y el riesgo de errores, generando retrasos y afectando la calidad de atención.
- **Alta carga de consultas repetitivas:** El 85% de las llamadas eran consultas repetitivas, provocando desgaste mental y disminución de motivación del personal al no poder enfocarse en tareas de mayor valor.
- **Estrés laboral y rotación de personal:** La presión por atender múltiples llamadas simultáneamente, sin soporte tecnológico adecuado, derivó en alta rotación y pérdida de experiencia.
- **Impacto en desempeño y satisfacción del usuario:** La sobrecarga del personal generó errores, tiempos de espera elevados y una experiencia de usuario deficiente, reforzando la percepción de un servicio ineficiente.

## 2. Objetivos vs. resultados

### Declaración de objetivos principales

Aunque los objetivos declarados eran correctos, no se consideró la necesidad de establecer objetivos intermedios para alcanzarlos de manera efectiva. La falta de metas específicas para grupos de usuarios como adultos mayores limitó la accesibilidad y eficiencia de la herramienta.

### Aspectos organizacionales

No se organizaron de manera adecuada los objetivos ni se anticiparon posibles fallas, lo que afectó la efectividad del proyecto. La ausencia de planificación estratégica impidió prever ajustes necesarios durante la implementación.

### Gestión de proyecto

No existió una hoja de ruta clara. Se evidenciaron fallas críticas en planificación, seguimiento de hitos, asignación de recursos y control de calidad. Los resultados deteriorados demostraron la necesidad de:

- Revisión inmediata de estrategias.
- Ajustes en la planificación.
- Redistribución de recursos.



- Mejoras en los procesos de atención.

## **Factores humanos**

La sobrecarga de trabajo, el estrés y la falta de capacitación frente a la automatización fueron determinantes en el incumplimiento de objetivos. Es clave mejorar la asignación de personal, la capacitación y el soporte tecnológico para evitar impactos negativos en futuros proyectos.

## **3. Análisis técnico**

### **3.1 Capacidades insuficientes del chatbot**

El chatbot inicialmente seleccionado estaba obsoleto y no contaba con terminología médica, lo que limitó su capacidad para ofrecer soluciones precisas. La base de conocimiento de 500 preguntas frecuentes estaba mal definida, impidiendo conversaciones efectivas.

### **3.2 Integración parcial con el sistema**

El bot no estaba conectado al historial médico ni a la disponibilidad de turnos, lo que generó incertidumbre en los usuarios. La mala configuración de APIs REST con el sistema hospitalario reflejó falta de planificación y coordinación técnica, afectando la interoperabilidad y la experiencia del paciente.

### **3.3 Limitaciones en comprensión del lenguaje**

El chatbot no comprendía completamente el lenguaje humano ni la jerga local. La ausencia de un relevamiento de expresiones y pruebas con usuarios representativos redujo la naturalidad de la interacción y la efectividad del servicio.

### **3.4 Impacto en los resultados**

El chatbot no estaba preparado para atender adecuadamente a adultos mayores ni manejar el volumen y la complejidad de consultas reales. Esto incrementó los tiempos de espera, disminuyó la automatización efectiva y generó insatisfacción en los pacientes.

## **4. Timeline de implementación**

- **Kick-off y desarrollo inicial (enero–marzo 2024):** Se definieron requerimientos básicos y se construyó un MVP con 200 FAQs. Las pruebas fueron internas, sin retroalimentación de usuarios reales, generando brechas de validación.

- **Integración con HIS (abril 2024):** La conexión con el sistema hospitalario presentó problemas de compatibilidad y no se establecieron mecanismos de contingencia, comprometiendo la estabilidad del proyecto.
- **Prueba Beta (mayo 2024):** Con 50 pacientes, se detectaron errores y limitaciones técnicas. Sin embargo, los hallazgos no se tradujeron en acciones correctivas, evidenciando falta de gobernanza y control de calidad.
- **Lanzamiento oficial (junio 2024):** Se ejecutó a pesar de los problemas detectados en la beta, generando sobrecarga del sistema, mala experiencia de usuario y múltiples quejas.
- **Crisis y desactivación parcial (julio–agosto 2024):** La falta de resolución oportuna obligó a limitar funciones y reducir el uso del bot, afectando la confianza de los usuarios y sobrecargando al personal humano.

## 5. Problemas críticos identificados

### Aspectos técnicos

- Falta de investigación y empatización con el usuario.
- Uso insuficiente de la tecnología y preparación tecnológica limitada.

### Aspectos organizacionales

- Riesgos de seguridad y cumplimiento legal por ausencia de protocolos de validación y encriptación.
- Limitaciones técnicas que afectaron operación y experiencia del usuario.

### Gestión de proyecto

- Falta de control de calidad, supervisión de riesgos y planificación de capacidad.
- Decisiones apresuradas impulsadas por cronogramas sin priorizar la estabilidad técnica.

## **6. Impacto en el cliente**

### **Aspectos técnicos**

- La satisfacción del cliente cayó un 47%, con pérdida de 340 pacientes.
- Los costos operativos aumentaron un 15%, equivalentes a USD 23,000.
- El tiempo de espera incrementó un 25%.
- Reducción en captación de nuevos pacientes en 12%.
- Riesgo legal: tres reclamos por información médica incorrecta.

### **Aspectos organizacionales**

- Falta de planificación, coordinación y gestión de recursos humanos.
- Déficit en control de riesgos y gobernanza.
- Impacto en reputación y procesos internos.

### **Gestión de proyecto**

- Deficiencias en planificación, control de calidad y seguimiento de hitos.
- Incremento de costos y sobrecarga operativa.
- Necesidad de reforzar monitoreo de indicadores críticos y protocolos de mitigación de riesgos.

### **Factores humanos**

- Sobrecarga laboral, estrés, intervención constante y manejo de errores críticos.
- Impacto directo en experiencia del paciente, calidad de atención y satisfacción.

## **7. Análisis de stakeholders, decisiones clave y puntos de inflexión**

### **Stakeholders involucrados**

- **Pacientes:** Usuarios finales del chatbot, con necesidades particulares como adultos mayores que requerían interfaces accesibles y comprensión clara de la herramienta. Su feedback fue limitado durante las fases iniciales, lo que afectó el diseño y la efectividad del sistema.
- **Personal operativo (secretarías y operadores):** Soportaron la carga de trabajo real y compensaron las deficiencias del chatbot, enfrentando estrés, sobrecarga y errores derivados de la intervención manual constante.
- **Equipo técnico y desarrolladores:** Encargados de implementar el chatbot, definir la base de conocimiento y asegurar la integración con HIS. Las limitaciones técnicas del modelo inicial y la falta de coordinación con otras áreas fueron críticas.
- **Gestión de la clínica:** Tomó decisiones estratégicas sobre presupuesto, cronograma y alcance del proyecto. Su enfoque priorizó fechas de lanzamiento sobre validación de calidad y preparación operativa.
- **Reguladores o supervisores legales:** Intervinieron indirectamente al existir riesgos de responsabilidad legal por información médica incorrecta y falta de protocolos de seguridad de datos.

### Decisiones clave tomadas

- Implementar un chatbot de IA generativa para automatizar la atención de llamadas repetitivas y gestión de turnos, sin un análisis exhaustivo de capacidad técnica y necesidades del usuario final.
- Desarrollar un MVP limitado a 200–500 preguntas frecuentes sin terminología médica ni casos complejos, lo que restringió la utilidad del sistema.
- Avanzar al lanzamiento oficial a pesar de los hallazgos críticos de la prueba beta, priorizando el cronograma sobre la calidad técnica y la experiencia del usuario.
- No establecer mecanismos de contingencia ni protocolos de escalamiento, trasladando la carga operativa al personal humano.

### Puntos de inflexión

- **Kick-off y desarrollo inicial:** La decisión de realizar pruebas únicamente internas generó una brecha de validación que afectó todas las etapas posteriores.

- **Prueba Beta:** La falta de acciones correctivas frente a errores detectados marcó un punto crítico; se perdió la oportunidad de mejorar el chatbot antes del go-live.
- **Lanzamiento oficial:** Constituyó el punto de no retorno: el sistema se expuso a sobrecarga, frustración de usuarios y crisis operativa, evidenciando la insuficiencia de planificación y coordinación entre áreas.
- **Crisis y desactivación parcial:** Este punto consolidó los impactos negativos sobre personal, pacientes y reputación institucional, resaltando la importancia de incorporar validación temprana, protocolos de contingencia y retroalimentación constante de stakeholders en futuros proyectos.

## B) ¿Qué pudo prevenirse?

### 1. Mejor planificación

**Problemas detectados.** El tiempo de espera promedio aumentó de 3 a 15 minutos. La automatización de consultas básicas solo alcanzó 23% frente al 70% esperado. Los costos operativos aumentaron un 15% en lugar de reducirse un 40%. La satisfacción del paciente cayó de 8.5/10 a 4.2/10.

**Análisis.** La planificación inicial no contempló:

- Estimación precisa del volumen de llamadas y capacidad de atención del personal.
- Estrategias de escalabilidad para horarios pico.
- Presupuestos y análisis de riesgos financieros por fallas técnicas o retrabajos.
- Fases piloto y pruebas progresivas antes de despliegues masivos.
- Mecanismos de retroalimentación temprana para ajustar expectativas y recursos.

### Prevención.

- Simulaciones de flujo de llamadas y análisis de capacidad de personal.
- Prototipos piloto y pruebas escalonadas de automatización.
- Planificación de contingencias para desvío temporal a agentes humanos o externos.

- Indicadores de alerta temprana y monitoreo continuo de KPIs operativos.

## 2. Testing

**Problemas detectados.** Cuellos de botella en tiempos de espera, bajo nivel de automatización de consultas, incremento inesperado de costos operativos y caída de satisfacción del paciente.

**Análisis.** El testing fue insuficiente en:

- Pruebas de carga y flujo completo para simular escenarios reales.
- Pruebas de integración y regresión con sistemas existentes (HIS, bases de datos).
- Pruebas de experiencia de usuario para detectar problemas de interacción, claridad y accesibilidad.
- Uso de datos reales para evaluar precisión del chatbot.

### Prevención.

- Implementar pruebas unitarias, de integración y regresión iterativas.
- Simulaciones progresivas de carga de usuarios y flujo de llamadas.
- Tests de UX con pacientes de distintos perfiles.
- Monitoreo de KPIs durante pruebas y producción para detectar desviaciones antes de afectar a usuarios.

## 3. Metodología

**Problemas detectados.** Los objetivos no estaban vinculados con entregables concretos. Hubo falta de revisiones periódicas, ausencia de gestión de riesgos y control de calidad, además de carencia de ciclos de retroalimentación y aprendizaje continuo.

**Análisis.** La metodología utilizada no contempló:

- Iteraciones cortas y entregas parciales validadas.

- Uso de OKRs o KPIs claros para medir avance y detectar desviaciones.
- Integración de gestión de riesgos y control de calidad en cada fase.
- Retroalimentación continua de usuarios y stakeholders.

#### **Prevención.**

- Implementar metodologías ágiles (Scrum/Kanban) con sprints cortos.
- Definir OKRs y dashboards de seguimiento de indicadores clave.
- Establecer gestión de riesgos formal, revisión periódica y control de calidad.
- Crear ciclos de aprendizaje y mejora continua con feedback de usuarios reales.

#### **4. Red flags ignoradas**

**Problemas detectados.** Se presentó incremento inesperado del tiempo de espera, cobertura limitada de automatización, incremento de costos operativos y caída en satisfacción del paciente.

**Análisis.** Se ignoraron señales de alerta tempranas tales como:

- Indicadores de desempeño insuficiente durante pruebas internas.
- Limitaciones técnicas del chatbot (memoria de contexto, manejo de conversaciones largas).
- Discrepancia entre recursos planificados y necesidades reales.
- Feedback de usuarios piloto mostrando errores y baja satisfacción.

#### **Prevención.**

- Establecer pruebas piloto y métricas iterativas antes del go-live.
- Redefinir objetivos basados en factibilidad técnica.
- Simulaciones de costos y recursos para anticipar desviaciones.

- Monitoreo de KPIs de satisfacción y tiempo de espera, con alertas tempranas y corrección inmediata.

## 5. Plan de prevención de riesgos

Para proyectos similares, se recomienda:

- **Evaluación inicial detallada:** volumen de usuarios, capacidad operativa, recursos humanos y tecnológicos.
- **Pruebas piloto iterativas:** validar procesos, automatización y experiencia de usuario.
- **Metodología ágil:** entregas incrementales, seguimiento de KPIs y revisiones periódicas.
- **Gestión de riesgos formal:** identificar, evaluar, mitigar y monitorear riesgos técnicos, operativos y financieros.
- **Monitoreo y feedback continuo:** KPIs de rendimiento, satisfacción y costos, con alertas tempranas y acciones correctivas.
- **Planes de contingencia:** recursos humanos, soporte externo y escalamiento gradual para eventos inesperados.

**Conclusión.** Un enfoque integral combinando planificación estratégica, testing exhaustivo, metodología ágil y gestión de riesgos, complementado con detección temprana de red flags, habría prevenido los problemas críticos observados y asegurado que los objetivos del proyecto se cumplieran efectivamente.

## 6. Impacto en el cliente

**Problemas detectados.** Satisfacción disminuida (-47%), incremento de costos operativos (+15%), aumento del tiempo de espera (+25%), reducción de nuevos pacientes (-12%) y riesgo legal (reclamos por información incorrecta).

### Análisis y prevención.

- **Satisfacción:** implementar pruebas piloto y retroalimentación continua; capacitar al personal en soporte híbrido.



- **Costos:** presupuestos detallados por fase; plan de optimización de recursos; indicadores de control financiero.
- **Tiempo de espera:** simulaciones de carga; priorización y derivación automatizada; ajuste progresivo de usuarios atendidos.
- **Nuevos pacientes:** monitoreo de reputación online; campañas de información; plan de mejora continua.
- **Riesgo legal:** validación profesional de contenido médico; protocolos legales; capacitación en gestión de reclamos.

**Conclusión.** Una planificación integral y proactiva habría permitido mantener satisfacción, eficiencia operativa, reputación y cumplimiento legal, mitigando impactos negativos en el cliente.

## C) ¿Qué NO Pudo Prevenirse?

### 2. Objetivos vs Resultados

#### 1. Restricciones Presupuestarias Realistas

Al evaluar los resultados del proyecto frente a los objetivos iniciales, se observa que parte del desempeño limitado puede atribuirse a restricciones presupuestarias reales que limitaban la capacidad del equipo para implementar soluciones completas. Por ejemplo:

La automatización de consultas básicas solo alcanzó un 23% frente al 70% esperado. Esto sugiere que los recursos financieros disponibles no permitieron entrenar modelos más complejos ni incorporar personal técnico adicional para optimizar la integración del chatbot.

El aumento del 15% en costos operativos puede reflejar la necesidad de cubrir gastos inesperados de soporte técnico o mantenimiento, pero también evidencia que la inversión inicial no contempló márgenes de error realistas para un proyecto de esta escala.

**Conclusión:** La limitación presupuestaria realista impuso un techo de desempeño que el equipo no podía superar, incluso con planificación y ejecución óptimas.

#### 2. Factores Externos Fuera del Control del Equipo

Algunos resultados negativos parecen estar relacionados con factores externos imposibles de controlar completamente:

Incremento en el tiempo de espera a 15 minutos podría haber sido parcialmente causado por una demanda de usuarios superior a la proyectada o cambios inesperados en la afluencia de llamadas debido a campañas de salud externas o emergencias locales.

La satisfacción del paciente reflejada en un NPS de 4.2/10 puede haber sido afectada por factores externos, como expectativas elevadas del público o eventos sociales/culturales que influyen en la percepción del servicio.

**Conclusión:** Aunque el equipo implementó estrategias de planificación, testing y metodologías ágiles, ciertos factores externos como volumen de usuarios inesperado y percepción social no pudieron ser prevenidos directamente.

### 3. Problemas Evitables

Algunos de los resultados negativos sí hubieran podido mitigarse:

La baja cobertura de automatización y la sobrecarga del personal indican que hubo fallas en planificación, pruebas de capacidad y escalamiento inteligente, lo que representa problemas que sí podían prevenirse mediante mejores simulaciones, fases piloto más robustas y ajustes iterativos.

El aumento de costos operativos también podría haberse limitado mediante seguimiento financiero más detallado y planes de contingencia.

**Conclusión:** Esta categoría evidencia que no todos los resultados fallidos pueden atribuirse a factores inevitables; parte del desempeño crítico deriva de fallas de ejecución prevenibles.

### 4. Limitaciones Sistémicas

Existen elementos del sistema y la organización que impidieron alcanzar los objetivos a pesar de la planificación:

La integración parcial del chatbot con los sistemas existentes (HIS, bases de datos de turnos) limitó la capacidad de automatizar consultas y derivar eficientemente casos a operadores humanos, aumentando el tiempo de espera.

Procesos internos rígidos, dependencia de recursos humanos limitados y estructuras jerárquicas lentas dificultaron la implementación de ajustes rápidos frente a problemas emergentes.

**Conclusión:** Estas limitaciones sistémicas representan barreras estructurales que no podían ser completamente resueltas por el equipo en el corto plazo y contribuyeron al incumplimiento de los objetivos.

### 5. Limitaciones Tecnológicas Genuinas

Finalmente, algunos resultados negativos se derivan de limitaciones tecnológicas genuinas del momento y de las herramientas disponibles:

Modelos de lenguaje y chatbots con memoria de contexto limitada y capacidad de manejar solo un número reducido de FAQs (500) restringieron la cobertura real de automatización, haciendo imposible alcanzar el 70% proyectado sin esperar avances tecnológicos adicionales.

Falta de sofisticación en reconocimiento de jerga local, variaciones lingüísticas y escenarios clínicos complejos implicó que la experiencia del usuario no pudiera mejorar completamente, afectando la satisfacción y la eficiencia operativa.

**Conclusión:** Estas limitaciones tecnológicas genuinas establecieron un límite natural al

rendimiento del proyecto que estaba fuera del control del equipo, independientemente de planificación o metodología.

### 3. Análisis Técnico

#### 1. Restricciones Presupuestarias Realistas

La limitación de la base de conocimiento a 500 FAQs básicas probablemente reflejaba un techo presupuestario que impedía entrenar un modelo más amplio y completo o contratar expertos en contenido médico para enriquecer la base de datos.

La implementación de una integración más profunda con el HIS y un sistema de escalamiento inteligente también requiere recursos financieros significativos, incluyendo servidores, licencias, APIs de pago y personal especializado.

**Conclusión:** Parte de las limitaciones técnicas eran consecuencia de restricciones presupuestarias reales que el equipo no podía superar, incluso con buena planificación.

#### 2. Factores Externos Fuera del Control del Equipo

La jerga local cordobesa y variaciones lingüísticas no fueron contempladas inicialmente, y ajustar el modelo a estos matices requiere datos que, en muchos casos, no están disponibles públicamente o requieren colaboración de terceros (médicos, lingüistas). Esto representa un factor externo que estaba fuera del control inmediato del equipo.

Cambios en las necesidades de los usuarios, como el aumento en la complejidad de consultas médicas o emergencias sanitarias, no podían preverse completamente.

**Conclusión:** El equipo enfrentó factores externos que limitaban la capacidad de su tecnología para adaptarse a interacciones humanas complejas y cambiantes.

#### 3. Problemas Evitables

Integración parcial con el sistema HIS: si bien la integración completa puede ser costosa, una planificación más robusta del flujo de información y escalamiento podría haber reducido los tiempos de espera y la frustración del usuario.

La falta de memoria de contexto del chatbot genera conversaciones desconectadas. Este problema podría mitigarse parcialmente mediante ajustes de diseño, como mantener un registro temporal de interacciones durante la sesión.

Pruebas piloto insuficientes: implementar escenarios de prueba que incluyan variaciones lingüísticas y consultas complejas podría haber permitido identificar estas limitaciones antes del despliegue masivo.

**Conclusión:** Algunos problemas técnicos fueron prevenibles con mejor planificación, pruebas y diseño iterativo, aunque no completamente eliminables.

#### 4. Limitaciones Sistémicas

La arquitectura del sistema estaba diseñada para un flujo parcial y manual de derivación de consultas, lo que genera cuellos de botella inevitables en situaciones de alta demanda.

La dependencia del personal humano para consultas complejas es una limitación sistémica que no puede eliminarse sin cambiar el modelo operativo completo del centro de atención.

**Conclusión:** Las limitaciones de integración y escalamiento reflejan barreras estructurales dentro del sistema que el equipo técnico no podía modificar rápidamente.

#### 5. Limitaciones Tecnológicas Genuinas

La capacidad del modelo GPT-3.5 Turbo tiene límites inherentes: no maneja contextos largos ni consultas complejas de manera confiable. Esto hace que el chatbot no pueda resolver por sí solo el volumen ni la complejidad de las consultas médicas.

La comprensión limitada del lenguaje natural, especialmente en dialectos regionales o terminología especializada, es una restricción tecnológica genuina que no puede superar un equipo pequeño en el corto plazo.

La base de conocimiento limitada y la falta de acceso a historiales médicos eran barreras tecnológicas que afectaron directamente la automatización y la satisfacción del paciente.

**Conclusión:** Estas limitaciones reflejan la frontera tecnológica disponible en el momento y constituyen factores que no pudieron prevenirse por el equipo, independientemente de planificación o metodología.

#### 4. Timeline de Implementación

##### 1. Restricciones Presupuestarias Realistas

La fase inicial (Kick-off y desarrollo del MVP) se limitó a 200 FAQs, claramente reflejando restricciones de presupuesto para construir una base de conocimiento más amplia o entrenar un modelo más robusto.

La corrección de problemas detectados en la prueba beta habría requerido recursos adicionales significativos (tiempo de desarrollo, expertos en integración y pruebas con usuarios reales). Estas restricciones probablemente influyeron en la decisión de no retrasar el lanzamiento oficial, aun frente a señales de alerta.

**Conclusión:** Parte de los errores en el timeline se derivan de limitaciones financieras que el equipo no podía superar, impactando directamente en la capacidad de prevención de fallos.

##### 2. Factores Externos Fuera del Control del Equipo

Los problemas de compatibilidad durante la integración con el HIS fueron parcialmente imprevisibles, dado que los sistemas hospitalarios suelen variar en estándares y configuraciones, algo que el equipo no puede controlar ni anticipar completamente.

La respuesta de los pacientes durante la prueba beta también es un factor externo: los errores de precisión y la frustración generada dependen de cómo los usuarios interactúan con el sistema, y estos patrones no siempre se pueden simular o prever con exactitud.

**Conclusión:** Algunas fallas críticas del timeline, especialmente la magnitud de los errores durante la beta y el impacto del lanzamiento, estaban fuera del control directo del equipo.

### **3. Problemas Evitables**

El lanzamiento oficial a pesar de la señal de alerta en la beta fue un claro problema prevenible. La decisión de avanzar sin corregir fallas detectadas refleja mala gestión de riesgos, no una limitación tecnológica.

La falta de un plan de escalamiento gradual también contribuyó a la sobrecarga del sistema. Se podría haber mitigado mediante un despliegue progresivo o pruebas con un grupo piloto más amplio.

La crisis que obligó a desactivar parcialmente el bot también podría haberse reducido con una planificación de contingencia más sólida, incluyendo protocolos de rollback y soporte intensivo para usuarios.

**Conclusión:** Estos problemas fueron en gran parte evitables con decisiones de gestión más prudentes y un enfoque iterativo en el timeline.

### **4. Limitaciones Sistémicas**

La arquitectura del proyecto y la dependencia del equipo para resolver problemas detectados tardíamente muestran limitaciones sistémicas en la planificación de proyectos tecnológicos en entornos hospitalarios.

El cronograma ajustado y la presión por cumplir fechas de lanzamiento reflejan un sistema de gestión que no incorporó suficientemente la incertidumbre tecnológica ni la necesidad de iteración basada en pruebas beta.

**Conclusión:** Estas limitaciones son sistémicas, propias de cómo se planificó y gestionó el proyecto, y contribuyeron a que ciertos problemas no pudieran prevenirse plenamente.

### **5. Limitaciones Tecnológicas Genuinas**

Durante la prueba beta se evidenció la baja precisión del bot y errores en respuestas, limitaciones inherentes al modelo y a la base de conocimiento reducida.

Estas deficiencias tecnológicas, especialmente en comprensión del lenguaje y manejo de contextos, no podían resolverse completamente dentro del timeline previsto sin aumentar significativamente la inversión en desarrollo y entrenamiento.

**Conclusión:** Parte de los problemas técnicos del timeline reflejan restricciones tecnológicas genuinas que el equipo no podía superar con los recursos y plazos disponibles.

## **5. Problemas Críticos Identificados**

### **1. Restricciones Presupuestarias Realistas**

La falta de validación profesional y mecanismos de control de calidad probablemente se relaciona con un presupuesto limitado. Contratar expertos en medicina para supervisar las respuestas del bot o implementar sistemas de revisión de contenido habría requerido recursos adicionales.

La ausencia de optimizaciones que redujeran latencia y caídas frecuentes también puede estar asociada a restricciones financieras que limitaron el acceso a servidores más potentes o arquitecturas escalables.

**Conclusión:** Parte de los problemas críticos fueron consecuencia de limitaciones presupuestarias reales, que el equipo no podía superar sin inversión extra.

### **2. Factores Externos Fuera del Control del Equipo**

La interacción con pacientes reales trajo desafíos que no se podían prever completamente: adultos mayores con dificultades tecnológicas, variaciones en el lenguaje o expresiones locales, y situaciones de urgencia que el bot no reconocía.

La exposición a riesgos legales derivados de información médica incorrecta también depende de factores externos: regulaciones locales, responsabilidad civil y reacción de pacientes ante errores, elementos que el equipo no podía controlar directamente.

**Conclusión:** Algunos impactos, especialmente en experiencia de usuario y cumplimiento legal, eran difíciles de anticipar y prevenir, independientemente del esfuerzo del equipo.

### **3. Problemas Evitables**

La falta de cifrado y medidas de seguridad es un problema prevenible, dado que existen estándares de seguridad informática ampliamente conocidos. La ausencia de estas medidas refleja una deficiencia de diseño y planificación.

La interfaz poco amigable y respuestas robóticas podrían haberse mejorado mediante pruebas de usabilidad más amplias o diseño centrado en el usuario.

**Conclusión:** Mientras que algunos problemas fueron inevitables, otros sí podrían haberse prevenido con mejores prácticas de desarrollo seguro y diseño de UX.

### **4. Limitaciones Sistémicas**

La combinación de fallas técnicas, mala experiencia de usuario y riesgos legales refleja limitaciones sistémicas: la estructura de planificación del proyecto no contemplaba iteraciones suficientes ni supervisión profesional constante, lo que impidió anticipar cómo el sistema funcionaría en el mundo real.

La sobrecarga del personal humano y la incapacidad del bot para manejar situaciones complejas también revelan deficiencias en el sistema organizativo, donde no se asignaron recursos de soporte adecuados.

**Conclusión:** Las limitaciones sistémicas fueron un factor clave que hizo que ciertos problemas fueran inevitables en la operación real del sistema.

## 5. Limitaciones Tecnológicas Genuinas

La latencia alta, desincronización y caídas frecuentes representan limitaciones tecnológicas genuinas del entorno en que se implementó el bot, incluyendo infraestructura y modelo de IA con memoria de contexto limitada.

La incapacidad de reconocer urgencias o comprender lenguaje local también refleja deficiencias inherentes al modelo GPT-3.5 Turbo y a la base de conocimiento limitada.

**Conclusión:** Algunos problemas críticos no podían prevenirse completamente, ya que derivan de limitaciones técnicas reales del software y la infraestructura disponible.

## 6. Impacto en el Cliente

### 1. Restricciones Presupuestarias Realistas

El aumento de costos operativos (+15%) y la sobrecarga del personal humano reflejan limitaciones financieras que restringieron la capacidad del proyecto para implementar soluciones más robustas o infraestructura tecnológica escalable.

La falta de inversión en iteraciones adicionales de prueba de usuario o en optimización del chatbot también se vincula con presupuesto limitado, que condicionó la prevención de errores de atención y tiempos de espera elevados.

**Conclusión:** Parte del impacto negativo sobre clientes se derivó de restricciones presupuestarias inevitables, que limitaron mejoras preventivas.

### 2. Factores Externos Fuera del Control del Equipo

La reacción de los pacientes, su percepción de la experiencia y la reputación online son factores externos que el equipo no podía controlar completamente. La caída de calificación en Google y la reducción de nuevos pacientes (-12%) dependen de cómo los usuarios perciben y difunden su experiencia.

Los reclamos legales por información médica incorrecta también incluyen factores externos: normativa sanitaria, interpretación de responsabilidad legal y reacción de los pacientes ante errores, que no se podían prever con exactitud.

**Conclusión:** Varios impactos negativos, especialmente en confianza, reputación y riesgo legal, eran difíciles de anticipar y prevenir.

### 3. Problemas Evitables

La mala experiencia del usuario y respuestas imprecisas del chatbot podrían haberse mitigado con pruebas de usuario más exhaustivas, mejoras en la interfaz y escalamiento inteligente de consultas.

La sobrecarga del personal humano y el incremento de llamadas de escalamiento son problemas parcialmente prevenibles, si se hubieran planificado mejor los recursos de soporte y la automatización gradual.

**Conclusión:** Algunos impactos negativos sí podrían haberse prevenido mediante mejores prácticas de UX, planificación y gestión de escalamiento.

#### **4. Limitaciones Sistémicas**

La combinación de aumento de tiempo de espera (+25%), costos adicionales y pérdida de nuevos pacientes refleja deficiencias sistémicas en la planificación del proyecto, en la gestión de riesgos y en la alineación entre tecnología, operaciones y experiencia de usuario.

La incapacidad de anticipar cómo los errores del chatbot afectarían directamente la satisfacción y la retención de pacientes evidencia una limitación sistémica de control preventivo sobre los resultados del cliente.

**Conclusión:** Las limitaciones del sistema organizativo contribuyeron significativamente a la magnitud del impacto en clientes.

#### **5. Limitaciones Tecnológicas Genuinas**

El chatbot no preparado para consultas complejas, falta de memoria de contexto y errores en respuestas son limitaciones tecnológicas genuinas que impactaron directamente en la satisfacción y confianza del cliente.

La sobrecarga de operadores humanos y la lentitud del sistema reflejan restricciones inherentes a la tecnología implementada, como capacidad de procesamiento, integración con HIS y capacidades de automatización del modelo GPT-3.5 Turbo.

**Conclusión:** Algunos problemas que afectaron a los clientes no podían prevenirse completamente debido a las limitaciones técnicas reales del sistema.

### **D) Cómo Trabajarías en el Futuro**

La implementación fallida de MediBot evidencia la necesidad de un enfoque más riguroso, escalonado y adaptativo para futuros proyectos de IA conversacional en el sector salud. A continuación, se detalla una propuesta integral basada en los problemas técnicos, operativos y estratégicos identificados.

#### **1. Framework de Evaluación de Viabilidad**

- **Técnica**
  - Evaluar compatibilidad con sistemas de gestión de turnos y registros médicos electrónicos (HIS, EHR, CRM).
  - Probar la capacidad de la IA para entender consultas médicas comunes y complejas sin errores críticos.



- Ampliar la base de conocimiento con terminología médica, jerga local y casos clínicos complejos.
- Mejorar la memoria de contexto para mantener conversaciones continuas y consultas de seguimiento.
- Garantizar seguridad y privacidad de datos sensibles (cumplimiento HIPAA/GDPR).
- **Financiera**
  - Comparar costos de desarrollo vs. beneficios esperados, incluyendo ahorro de tiempo y reducción de errores.
  - Incluir presupuesto para mantenimiento, actualizaciones y entrenamiento continuo de la IA.
- **Operativa**
  - Verificar disponibilidad de personal capacitado para supervisión, soporte y escalamiento.
  - Definir claramente qué consultas pueden automatizarse y cuáles requieren intervención humana.
- **Ética y Legal**
  - Implementar auditorías de decisiones automatizadas para detectar sesgos clínicos o errores.
  - Establecer responsabilidad clara en caso de recomendaciones incorrectas o errores médicos.

## 2. Proceso de Testing Escalonado

- **Pruebas Unitarias:** Evaluar cada módulo de IA (interpretación de lenguaje, asignación de turnos, respuestas a FAQs).
- **Pruebas de Integración:** Asegurar correcta interacción entre módulos y conexión con HIS y CRM.
- **Pruebas de Usuario Final:** Personal médico y administrativo simula escenarios reales para validar efectividad y escalamiento.

- **Piloto Limitado:** Implementación en una sección de la clínica para evaluar desempeño, tiempo de respuesta y satisfacción del paciente.
- **Monitoreo Continuo:** Revisar métricas críticas: tiempo de respuesta, resolución de consultas, errores recurrentes y NPS.

### 3. Criterios de Éxito y Fallo

- **Éxito**
  - Reducción del tiempo de espera  $\geq 50\%$ .
  - Automatización de consultas básicas con precisión  $\geq 80\%$ .
  - Satisfacción del paciente (NPS  $\geq 8/10$ ).
  - Disminución de errores administrativos y sobrecarga del personal.
- **Fallo**
  - Incremento de tiempos de espera o consultas incorrectamente gestionadas.
  - Retroceso en la satisfacción del paciente (NPS  $< 6/10$ ).
  - Necesidad constante de intervención humana debido a errores del bot.

### 4. Estrategias de Gestión del Cambio

- Capacitación progresiva del personal en el uso del chatbot.
- Comunicación clara sobre funciones, limitaciones y beneficios del sistema.
- Canales de feedback para reportar problemas y sugerencias de mejora.
- Incentivos al personal que adopte la herramienta y participe en su optimización.

### 5. Plan de Contingencia

- Procedimientos manuales alternativos para consultas críticas o fallas del bot.
- Estrategia de rollback para desactivar MediBot sin afectar la operación clínica.

- Equipo de soporte técnico y médico listo para intervención inmediata.
- Alertas tempranas y monitoreo de métricas clave para identificar fallos rápidamente.

## **6. Consideración de Avances Tecnológicos**

- Integración con IA multimodal: voz, texto e imagen para diagnósticos preliminares.
- Modelos clínicos especializados entrenados en datasets locales para reducir sesgos.
- Funcionalidad predictiva para optimizar programación de turnos y anticipar picos de atención.
- Actualización dinámica de la base de conocimiento según interacciones reales y feedback de usuarios.
- Monitoreo en tiempo real de métricas de desempeño para ajuste continuo del sistema.

## **E) Lecciones Aprendidas Personales**

El análisis del caso MediBot me permitió comprender de manera crítica los desafíos y riesgos asociados a la implementación de sistemas de IA en entornos sensibles como la salud. Uno de los aprendizajes más relevantes es que la tecnología, por sí sola, no garantiza eficiencia ni mejora de la experiencia del usuario; su efectividad depende de una planificación integral que considere factores técnicos, operativos y humanos de manera simultánea.

La experiencia con MediBot evidencia que un enfoque centrado únicamente en cumplir plazos de desarrollo y objetivos tecnológicos puede generar consecuencias negativas graves, desde sobrecarga operativa hasta riesgos legales y disminución de la confianza de los pacientes. Esto me lleva a reconsiderar mi enfoque profesional: en proyectos futuros priorizaré la evaluación de viabilidad antes de implementar soluciones de IA, asegurando que exista compatibilidad técnica con los sistemas existentes, disponibilidad de personal capacitado para supervisión y mecanismos claros de escalamiento ante situaciones críticas.

Otro aprendizaje central es la importancia de la iteración y las pruebas escalonadas. El caso mostró que un lanzamiento masivo sin pruebas piloto suficientemente robustas puede resultar en indicadores de desempeño críticos deteriorados, como tiempos de espera más largos, baja automatización y reducción de la satisfacción de los usuarios. Esto refuerza la necesidad de un enfoque gradual, donde se realicen pruebas unitarias, de integración y de usuario final,

acompañadas de un monitoreo continuo que permita identificar y corregir problemas antes de afectar a los pacientes.

La reflexión crítica sobre esta experiencia también me lleva a adoptar principios claros para futuros proyectos de IA: la ética y la responsabilidad profesional deben ser pilares innegociables, garantizando que la IA nunca comprometa la seguridad del paciente ni viole normativas de privacidad y protección de datos; el diseño debe estar centrado en el usuario, considerando diversidad de perfiles y necesidades; y la transparencia debe guiar la comunicación sobre las capacidades, limitaciones y procedimientos de escalamiento del sistema.

Desde un punto de vista de competencias profesionales, el caso evidencia que no basta con habilidades técnicas aisladas. Para liderar proyectos de IA efectivos, es necesario desarrollar capacidades en gestión de proyectos, integración tecnológica, análisis crítico de resultados, coordinación de equipos multidisciplinarios y ética aplicada a la IA. La supervisión humana no solo es un requisito operativo, sino también una responsabilidad profesional que protege a los usuarios y asegura la confiabilidad del sistema. El aprendizaje ético del caso es contundente: los errores de la IA pueden tener consecuencias tangibles sobre la salud, la confianza y la reputación de la organización, por lo que asumir responsabilidad y mantener una auditoría constante del desempeño del sistema es esencial.

En conclusión, MediBot me enseñó que el éxito de un proyecto de IA no depende únicamente de la sofisticación tecnológica, sino de la capacidad de integrar planificación rigurosa, evaluación de riesgos, experiencia del usuario, supervisión profesional y ética en todas las decisiones. Para futuros proyectos, mi enfoque será holístico y crítico: implementar soluciones escalonadas, auditar constantemente el desempeño de la IA, priorizar la seguridad y satisfacción del usuario, y mantener transparencia y responsabilidad en todas las etapas del proyecto. Este aprendizaje fortalece mi visión profesional y establece un marco de acción que busca no solo eficiencia operativa, sino también confianza, seguridad y valor real para los usuarios.

## **F) Recomendaciones Estratégicas**

### **Introducción**

La implementación de sistemas GenAI en el sector salud representa una oportunidad transformadora, pero también conlleva riesgos éticos, técnicos y comunicacionales que deben ser abordados con rigurosidad. El caso fallido analizado evidencia la necesidad de una planificación integral que contemple no solo la eficiencia operativa, sino también la seguridad del paciente, la validación médica y la protección de datos sensibles. Este documento presenta recomendaciones específicas para los principales actores involucrados: la organización afectada, los profesionales de la salud, los proveedores tecnológicos y los reguladores,

integrando métricas, protocolos de escalamiento y reflexiones críticas basadas en experiencias reales y estándares internacionales.

## 1. Recomendaciones para la Organización Afectada (Clínica MediCare)

La clínica debe asumir un rol activo en la integración de GenAI, priorizando la calidad humana y la seguridad institucional.

### Acciones clave:

- **Revisión de procesos internos:** Realizar un mapeo detallado de los flujos de atención para identificar cuellos de botella y definir qué tareas pueden ser automatizadas sin comprometer la experiencia del paciente. Según el MIT (2023), el 40% de los errores en IA médica provienen de una integración deficiente con procesos existentes.
- **Fase piloto controlada:** Implementar pruebas escalonadas con grupos reducidos (10–20% de pacientes) permite detectar fallas tempranas y ajustar el sistema antes del despliegue total. El Hospital Mount Sinai en EE.UU. aplicó este enfoque con éxito en su chatbot de triage.
- **Sistema híbrido humano + IA:** Toda interacción debe tener un fallback hacia un operador humano en menos de 90 segundos si el bot no resuelve la consulta, basado en estándares de experiencia del paciente (CX) en salud digital.
- **Capacitación del personal:** Entrenar al call center y médicos en protocolos de interacción con IA garantiza continuidad sin fricciones. La falta de formación fue uno de los factores críticos en el caso fallido de Babylon Health (Reino Unido).

### Métricas de seguimiento:

- Tiempo promedio de espera en llamadas < 5 min
- Tasa de resolución de consultas por IA  $\geq 60\%$  en 6 meses
- NPS  $\geq 7.5/10$  en encuestas de pacientes

### Indicadores de alerta temprana:

- Incremento en reclamos > 5% mensual
- Escalamiento de llamadas a humanos > 40%

- Tiempo de respuesta > 10 minutos

#### **Protocolos de escalamiento:**

- Activación automática del fallback si la IA supera el 30% de respuestas fallidas en un día
- Comité interno de crisis con médico, técnico y legalista

#### **Reflexión crítica:**

La automatización no debe reemplazar el vínculo humano en salud, sino complementarlo. La clínica debe evitar la deshumanización del servicio y garantizar que la tecnología amplifique, no sustituya, la empatía profesional.

## **2. Recomendaciones para Profesionales del Sector Salud**

Los médicos y enfermeros son garantes de la calidad clínica y deben participar activamente en el diseño y validación de sistemas GenAI.

#### **Acciones clave:**

- **Validación del corpus médico:** Revisar FAQs, protocolos y terminología del chatbot. Según la OMS (2021), la validación humana es obligatoria en sistemas de IA que interactúan con pacientes.
- **Formación continua:** Cursos obligatorios en ética, sesgos algorítmicos y uso seguro de IA en salud para fortalecer la alfabetización digital y reducir riesgos clínicos.
- **Protocolos claros de uso:** Definir qué consultas puede atender la IA y cuáles requieren intervención humana (ej.: síntomas graves). Esta distinción fue clave en el éxito del sistema Ada Health en Alemania.

#### **Métricas de seguimiento:**

- % de contenido médico validado por especialistas = 100%
- Participación de profesionales en capacitaciones  $\geq 90\%$

#### **Indicadores de alerta temprana:**

- Casos de información médica incorrecta reportados > 1/mes
- Falta de actualización del corpus > 3 meses

#### **Protocolos de escalamiento:**

- Derivación automática al área médica ante síntomas críticos
- Reporte semanal de errores al comité científico

#### **Reflexión crítica:**

La IA no puede reemplazar el juicio clínico. La participación activa de los profesionales es esencial para evitar que el sistema reproduzca sesgos, errores o simplificaciones peligrosas en contextos complejos.

### **3. Recomendaciones para Proveedores de Tecnología GenAI**

Los desarrolladores deben adaptar sus modelos al dominio médico local y garantizar una integración técnica robusta.

#### **Acciones clave:**

- **Modelos entrenados en contexto clínico:** Usar terminología médica local y ejemplos reales de consultas hospitalarias. El caso de MedPaLM (Google Health) demostró que el entrenamiento específico mejora la precisión en un 20%.
- **Integración con sistemas HIS:** Conectividad bidireccional para turnos, historiales y disponibilidad en tiempo real. Esto evita duplicación de datos y mejora la experiencia del paciente.
- **Diseño inclusivo:** Interfaces accesibles para adultos mayores, con voz, botones grandes e instrucciones simples. Según la Fundación Mayo Clinic, el 30% de los errores de uso en salud digital provienen de interfaces poco amigables.
- **Módulo de escalamiento inteligente:** Detectar frustración del usuario y transferir a humano.

#### **Métricas de seguimiento:**

- Latencia del sistema < 2 segundos por respuesta

- Uptime del sistema  $\geq 99.5\%$
- Precisión de respuestas del chatbot  $\geq 85\%$

#### **Indicadores de alerta temprana:**

- Caídas recurrentes > 2 por semana
- Tiempo de inactividad > 1 hora mensual
- Rechazo del bot > 20%

#### **Protocolos de escalamiento:**

- Línea directa de soporte 24/7 para hospitales
- Contratos SLA con penalizaciones si no se cumplen métricas críticas

#### **Reflexión crítica:**

La tecnología debe ser confiable, inclusiva y transparente. Los proveedores deben asumir responsabilidad ética y técnica, evitando soluciones genéricas que ignoren el contexto clínico real.

## **4. Recomendaciones para Reguladores del Sector Salud**

Los organismos reguladores deben establecer marcos normativos específicos para IA en salud, garantizando seguridad, transparencia y protección de datos.

#### **Acciones clave:**

- **Normativa específica:** Requisitos mínimos de seguridad, encriptación y validación médica. La Unión Europea trabaja en el AI Act, que clasifica la IA médica como “alto riesgo”.
- **Certificación obligatoria:** Chatbots médicos deben pasar pruebas de calidad antes de interactuar con pacientes, similar a los dispositivos médicos.
- **Protección de datos sensibles:** Supervisar cumplimiento de la Ley 25.326 en Argentina y promover auditorías periódicas.



- **Supervisión continua:** Crear un registro nacional de proyectos de IA en salud, con reportes trimestrales y auditorías externas.

#### **Métricas de seguimiento:**

- % de plataformas certificadas antes de operar = 100%
- Cumplimiento de normas de privacidad  $\geq 95\%$
- Número de incidentes de seguridad reportados

#### **Indicadores de alerta temprana:**

- Reportes de fallas graves > 3 en un mes
- Vulneraciones de datos personales > 1 incidente confirmado

#### **Protocolos de escalamiento:**

- Suspensión inmediata del uso de la IA en caso de riesgo a la seguridad del paciente
- Multas y sanciones a proveedores que no cumplan con certificaciones

#### **Reflexión crítica:**

La regulación debe ser proactiva, no reactiva. El Estado tiene la responsabilidad de proteger a los pacientes y garantizar que la innovación tecnológica no avance sin controles éticos y clínicos adecuados.

La implementación de GenAI en salud exige una mirada multidimensional, donde cada actor asuma su rol con responsabilidad, visión preventiva y compromiso ético. Las recomendaciones presentadas buscan corregir errores del caso fallido y construir un marco de referencia para futuras integraciones tecnológicas que respeten la dignidad del paciente, la calidad clínica y la sostenibilidad institucional. La clave está en la colaboración estratégica, la validación continua y la capacidad de anticipar riesgos antes de que se conviertan en crisis.

## Bibliografía

### Implementación de IA en salud

- Hospital Italiano de Buenos Aires. (2024). *Frameworks y guías para la implementación de IA en salud*. Recuperado de <https://www.hospitalitaliano.org.ar/hiba/es/news/frameworks-y-guias-para-la-implementacion-de-ia-en-salud>
- La Nación. (2024, 26 de noviembre). *El dilema de los médicos argentinos frente al avance de la IA*. Recuperado de <https://www.lanacion.com.ar/economia/IA/el-42-de-los-medicos-argentinos-teme-que-la-inteligencia-artificial-reemplace-su-criterio-clinico-nid26112024/>
- Young, P., Gil Uliarte, M. L., Papiccio, C. D., & Ernst, G. (2024). Plataforma avanzada de inteligencia artificial para la gestión de pacientes. *Revista Médica de Chile*, 152(9). Recuperado de [https://www.scielo.cl/scielo.php?script=sci\\_arttext&pid=S0034-98872024000901023](https://www.scielo.cl/scielo.php?script=sci_arttext&pid=S0034-98872024000901023)

### Riesgos éticos y legales de GenAI en medicina

- Bellido Diego-Madrado, R. A., et al. (2025). Aspectos éticos y legales del uso de nuevas tecnologías en la Medicina. *Revista Ocronos*, 8(2), 814. Recuperado de <https://revistamedica.com/aspectos-eticos-legales-nuevas-tecnologias-medicina/>
- Organización Mundial de la Salud (OMS). (2021). *Ética y gobernanza de la inteligencia artificial para la salud*. Recuperado de <https://www.who.int/es/publications/i/item/9789240029201>
- Comisión Europea. (2024). *AI Act: Un enfoque europeo para la inteligencia artificial*. Recuperado de <https://digital-strategy.ec.europa.eu/en/policies/a-i-act>

### Gestión de proyectos y metodologías

- Project Management Institute (PMI). (2024). *Guía de los fundamentos para la dirección de proyectos (Guía del PMBOK) - Séptima Edición*. Project Management Institute.
- Scrum Alliance. (2024). *The Scrum Guide*. Recuperado de <https://scrumguides.org/>

### Procesamiento de Lenguaje Natural (NLP)

- Prahallad, R., Sanyal, S., et al. (2024). *Med-PaLM: A Large Language Model for Medical Applications*. Recuperado de <https://ai.googleblog.com/2024/02/med-palm-large-language-model-for.html>

### Diseño y Experiencia de Usuario (UX)

- Nielsen Norman Group. (2024). *UX Research on Older Adults: 12 Common Misconceptions*. Recuperado de <https://www.nngroup.com/articles/ux-older-adults/>
- Norman, D. (2013). *The Design of Everyday Things*. Basic Books.

