

Guadalupe Mendoza Despontin

DNI: 45934026

Ciencia de Datos e Inteligencia Artificial

Repositorio:

<https://github.com/Guadamendoza/Practica-profesionalizante>

Docente: Daniela Accietto

Caso GenAI fallido

Para automatizar la gestión de pacientes y sus turnos la clínica “MediCare” optó por el lanzamiento de un chatBot, al cual lo definimos como un software de inteligencia artificial que simula tener una conversación humana para agilizar el trabajo del personal dentro de una organización. Sin embargo sus resultados no fueron los esperados. Por lo tanto, en este documento se analizará en profundidad el fracaso del proyecto, identificando sus causas, lecciones aprendidas y posibles mejoras futuras.

I. **Análisis de causas**

Para analizar a profundidad el caso de la clínica MediCare es importante comprender la información obtenida. Como mencionamos, el objetivo del bot era automatizar la gestión de pacientes y turnos, ya que presentaban problemas en la administración de llamadas (800 al día), alto tiempo de espera, preguntas repetitivas y la constante rotación del personal por estrés. A pesar de una inversión de \$85000 y 8 meses de desarrollo los resultados fueron totalmente negativos.

En cuanto a los **factores técnicos**, la principal causa de la falla fue un **entrenamiento insuficiente del chatbot**. Su efectividad depende directamente de la calidad del entrenamiento, que se basa en el procesamiento del lenguaje natural (PLN). La poca eficacia del bot se vio reflejada en su base de conocimiento limitada, no incluía terminología médica ni comprendía la jerga local. Además su falta de integración profunda impidió que accediera a información crucial como el historial médico y la disponibilidad de turnos. Como resultado, el objetivo principal del bot, que era gestionar consultas y turnos de manera efectiva, no pudo cumplirse.

Con respecto a las **causas organizacionales y de gestión del proyecto** podemos basarnos en la línea del tiempo. En abril y mayo se ignoraron advertencias sobre la falta de compatibilidad y la baja precisión del bot en las pruebas beta. Sin duda, el punto de inflexión más importante fue la decisión de seguir adelante con el proyecto sin antes resolver estos errores.

Además es importante destacar que se ignoraron a los **stakeholders claves**. En el caso de MediCare los adultos mayores eran los stakeholders externos más importantes ya que constituían el 65% de la base de pacientes. Para este grupo, la interfaz era confusa. También el 18% de las respuestas eran directamente incorrectas, lo que se agravó por la **falta de validación profesional** ya que las respuestas no eran revisadas por el personal médico.

En conclusión, el fracaso del chatbot de MediCare fue producto de una gestión apresurada y deficiente. Se ignoraron advertencias clave lo que produjo fallas técnicas, como el entrenamiento insuficiente y falta de datos complejos. Además la falta de organización con el personal médico produjo la presencia de respuestas incorrectas y la confusión de los usuarios finales más importantes, las personas mayores. Estos errores concluyeron con la pérdida de 340 pacientes.

II. ¿Qué pudo prevenirse?

En su mayoría, los errores del proyecto pudieron haberse prevenido con una mejor gestión de proyecto, un liderazgo adecuado para intervenir en la toma de decisiones y una comunicación efectiva dentro de la clínica.

Hablando específicamente, la **violación de privacidad**, pudo haberse evitado implementando una política de seguridad y capacitaciones obligatorias de la misma para que los protocolos se cumplan correctamente dentro de la clínica. Medidas como el cifrado de datos, auditorías regulares para identificar vulnerabilidades y la destrucción segura de información innecesaria son prácticas básicas que no se aplicaron y que hubieran reducido significativamente el riesgo de incidentes legales y reputacionales.

Por otra parte, los **problemas de experiencia de usuario**. La confusión para los adultos mayores pudo haberse evitado si se hubiera involucrado a este grupo en las pruebas beta y se hubieran definido correctamente los perfiles de usuario final desde la fase kick-off del proyecto. Además, la falta de reconocimiento médico del bot pudo haberse evitado implementando un sistema híbrido que deriva automáticamente al paciente a un profesional cuando la consulta es demasiado compleja.

Durante el desarrollo, hubo señales de **alerta claras que fueron ignoradas**. Por ejemplo, en las pruebas beta ya se detectaban problemas de precisión y compatibilidad, pero se decidió avanzar con el lanzamiento sin resolverlos. Además, no se evaluaron correctamente los riesgos legales asociados a la entrega de información médica incorrecta, ni se previó el impacto en la reputación online y la pérdida de pacientes.

La **falta de validación profesional** también fue un problema evitable. La ausencia de revisión médica se reflejó en la entrega de información errónea, la cual generó reclamos y puso en riesgo la salud de los pacientes.

En conclusión, la mayoría de los errores del proyecto MediBot eran prevenibles con una mejor gestión de riesgos, planificación y comunicación.

Plan de riesgos:

Objetivo:

Evitar futuras fallas en el software garantizando una correcta gestión del proyecto.

1. Identificación de riesgos:

A. Riesgos de seguridad:

- Violación de privacidad de datos médicos
- Entrega de información incorrecta

B. Riesgos técnicos:

- Latencia alta (8-12 segundos)
- Caídas del sistema (15 interrupciones en el primer mes)
- Información desactualizada

C. Riesgos de experiencia de usuario:

- Interfaz confusa
- No derivaba a los pacientes correctamente
- Respuestas robóticas
- Mala valoración online

D. Riesgos legales:

- Reclamos por información médica incorrecta

E. Riesgos económicos:

- Pérdida de 340 pacientes
- Aumento del 15% de costos operativos

2. Evaluación de riesgos: Se clasifican según su impacto: (alto-medio-bajo)

- Riesgos de seguridad: Alto
- Riesgos técnicos: Alto
- Riesgos de experiencia de usuario: Medio-Alto
- Riesgos legales: Alto
- Riesgos económicos: Alto

3. Plan de prevención:

Seguridad:

- Encriptación de datos
- Revisar el cumplimiento de la norma ISO 27001
- Capacitar al personal técnico y médico
- Realizar auditorías internas

Técnicos:

- Entrenar al bot con terminología médica especializada
- Revisión y actualización de la base de conocimiento
- Mejorar la conexión con sistemas HIS para acceso a historiales médicos y turnos en tiempo real
- Incorporar modelos que entiendan jerga local
- Pruebas con usuarios reales para ajustar fallas
- Implementar pruebas de testing

Experiencia de usuario:

- Rediseñar la interfaz teniendo en cuenta las necesidades de los usuarios finales
- Implementar un chatbot híbrido que derive casos complejos a los profesionales
- Capacitar al personal para revisar los mensajes regularmente

Económicos:

- Establecer soporte humano para informar fallas
- Implementar pruebas beta con usuarios reales
- Implementar pruebas piloto

Legales:

- Revisar la información entregada a los pacientes
- Implementar la encriptación de datos
- Capacitar al personal sobre el manejo de datos sensibles e información errónea

III. ¿Que no pudo prevenirse?

Por otra parte, algunas fallas del proyecto no eran completamente prevenibles. La IA generativa no es perfecta y tiene limitaciones que afectan su desempeño.

Los modelos de IA como el utilizado en MediBot tienen barreras tecnológicas, no comprenden contextos complejos ni conversaciones muy prolongadas. En el caso de MediBot, el modelo tenía dificultades

para interpretar correctamente la jerga local, su base de conocimiento no incluía terminología médica específica ni casos edge. Por lo tanto, no es sorpresa que muchas de sus respuestas sean incorrectas, es simple, si los datos no son de buena calidad los resultados tampoco y muchas veces la inteligencia artificial si no sabe, inventa.

Podemos compararlo con otro caso de un chatbot fallido dentro del campo de la medicina. Tessa, un chatbot de una ONG estadounidense dedicada a los trastornos alimenticios, que fue suspendido después de enviar respuestas incorrectas, como consejos de pérdida de peso que no pertenecían al programa oficial de la organización. Esta es una prueba de que con entrenamiento y supervisión, un chatbot de IA generativa no puede garantizar respuestas 100% correctas sin la intervención de un profesional humano.

Sin embargo, podemos destacar que el presupuesto no era el suficiente. Un ejemplo de proyecto exitoso es AIME, un chatbot médico desarrollado para asistencia avanzada. Su costo estimado según [HYENA.AI](#) oscila entre **\$100,000 y \$200,000 USD**, e incluso más. (Hyena.ai, 2024). AIME incluye funcionalidades clave que aseguran su eficacia, tales como:

- **Capacidades de PNL** (procesamiento del lenguaje natural)
- **Comprobación de síntomas y apoyo al diagnóstico**, (el chatbot analiza los síntomas y recomienda posibles afecciones)
- **Gestión de datos de los pacientes** (tiene acceso a su historial médico y conversaciones con el bot)
- **Programación de citas** (tiene acceso a la disponibilidad en tiempo real)
- **seguridad y cumplimiento de las normas sanitarias** (el cumplimiento de las normativas, como la HIPAA, garantiza la seguridad de los datos y la privacidad del paciente)

Comparando con el proyecto MediBot, que contaba con un presupuesto total de **\$85,000 USD**, es evidente que los recursos eran insuficientes para incluir funcionalidades similares a AIME. Esta limitación en el presupuesto probablemente impactó en las fallas del proyecto, como el entrenamiento del modelo con terminología médica específica, la integración profunda con los sistemas HIS, la capacidad de mantener conversaciones complejas. Como consecuencia, MediBot presentó latencia alta, caídas frecuentes del sistema, errores en la información proporcionada y una mala experiencia de usuario.

En conclusión, aunque MediBot falló por falta de entrenamiento, pruebas de testing y mejor organización entre los técnicos y la clínica, muchas de las limitaciones son propias de la tecnología de IA disponible en ese momento. Esto refleja la necesidad de combinar la IA con la supervisión profesional.

IV. Trabajo para el futuro: Metodología mejorada para implementaciones futuras de GenAI en salud

A. Framework de Evaluación de Viabilidad.

Antes de iniciar cualquier proyecto de GenAI en salud, es fundamental evaluar su viabilidad técnica, económica, legal y organizacional. Propongo el siguiente framework:

Análisis de las necesidades de stakeholders:

- Identificar los problemas reales a resolver y los usuarios finales (pacientes, médicos, administrativos).
- Involucrar a todos los stakeholders desde el inicio, a los mayores.

Evaluación técnica:

- Evaluar si calidad y cantidad de datos son suficientes para entrenar el modelo (ej: historiales médicos, jerga local, palabras técnicas).
- Evaluar la integración con HIS en tiempo real.
- Validar la seguridad de los datos.

Evaluación legal:

- Verificar el cumplimiento de la normativa de protección de datos.
- Validar si la información proporcionada a los pacientes es la correcta.

Evaluación económica:

- Analizar si el presupuesto es suficiente para desarrollar un proyecto completo y complejo
- Contemplar costos de mantenimiento

B. Proceso de testing escalonado:

Pruebas unitarias y de integración:

- Evaluar cada componente por separado (NLP, integración HIS, seguridad de datos).

Pruebas beta:

- Involucrar a los usuarios finales, especialmente a los adultos mayores y personal médico

- Evaluar la precisión del bot.

Validación profesional:

- Todas las respuestas deben ser revisadas y aprobadas por profesionales de la salud.

Pruebas piloto:

- Implementar el sistema en un entorno controlado, con monitoreo constante y posibilidad de rollback inmediato.

Monitoreo post-lanzamiento:

- Seguimiento en tiempo real de métricas clave (errores, satisfacción, tiempos de respuesta, incidentes de seguridad).

C. Criterios de éxito/fallo:

Criterios de éxito:

- Precisión > 95% en respuestas médicas básicas (validadas por médicos).
- Reducción del 40% en llamadas repetitivas.
- Aumento de la puntuación online.
- Cero violaciones de privacidad reportadas.

Criterios de fallo:

- Tasa de respuestas incorrectas > 10% en pruebas beta.
- Aumento de quejas > 20% en el primer trimestre.
- Latencia > 8 segundos en >15% de interacciones.
- Falta de aprobación regulatoria en auditorías.

D. Estrategias de gestión de cambio:

Capacitación para el personal:

- Entrenamiento específico para personal médico, administrativo y usuarios finales.
- Simulacros de uso y resolución de incidentes.

Enfoque híbrido:

- Mantener siempre la opción de derivar a un humano en casos complejos o de insatisfacción del usuario.

Feedback y mejora Continua:

- Canales abiertos para reportar errores, sugerencias y experiencias negativas.

Revisión periódica de métricas y ajustes iterativos.

E. Plan de contingencia:

Auditorías y revisiones:

- Auditorías mensuales de seguridad, calidad de datos y cumplimiento legal.

Mitigación de riesgos:

- Reserva del 20% del presupuesto para ajustes imprevistos.

Adaptación a avances tecnológicos:

- cláusulas en el contrato para actualizar modelos (ej: migrar a GPT-5 si está disponible, modelos multimodales).

Plan de recuperación:

- Procedimientos para restaurar datos, corregir errores y comunicar incidentes a los pacientes y autoridades.

Backup humano:

- Siempre estará disponible un equipo de soporte humano para tomar el control en caso de fallo del sistema.

V. Lecciones aprendidas

Mi experiencia con MediBot me dejó varias lecciones importantes, no solo para futuros proyectos de inteligencia artificial en el sector salud, sino para cualquier proyecto en un entorno crítico.

Personalmente el análisis de MediBot me permitió comprender que la IA no puede ni debe reemplazar al ser humano. Propongo la importancia de un enfoque híbrido, donde la supervisión humana tome un rol más importante. Asimismo, aprendí que tarde o temprano el modelo necesitará ajustes, sobre todo cuando ingresan nuevas tecnologías al mercado. Estos avances nos obligan a mantenernos en constante movimiento para poder mejorar y garantizar el éxito de los proyectos.

En cuanto a la gestión, la planificación es un paso muy importante. Es fundamental identificar las necesidades de los usuarios finales, ya que son ellos quienes interactuarán con la herramienta y dependerán de su correcto funcionamiento. Además, es necesario establecer procesos de testing para detectar y corregir errores antes de la implementación completa, como también, definir correctamente los criterios de éxito y fracaso para tener los mejores resultados posibles.

Investigando comprendí que es muy importante adoptar ciertas competencias. Las habilidades de gestión de proyectos son fundamentales para saber desarrollar cualquier tipo de proyectos, ya sea médicos, empresariales, educativos, etc. Permite identificar necesidades, establecer objetivos y metas claras, anticipar riesgos, y analizar tanto los éxitos como los fracasos. Además de la gestión, la inteligencia artificial es una habilidad fundamental. Su avance es rápido y requiere una actualización constante para aplicar técnicas y herramientas que optimicen los resultados. Por último, los conocimientos legales y éticos son fundamentales en el proceso, para garantizar la correcta implementación de los resultados, siguiendo las leyes para garantizar la seguridad e integridad de la organización y los clientes.

VI. Recomendaciones adicionales y específicas.

Para la organización afectada:

Recomendaciones:

- Auditorías periódicas del chatbot para corregir su funcionamiento y adaptarlo a las nuevas tecnologías.
- Capacitación al personal técnico y médico para evitar información errónea.
- Ajustar un presupuesto más alto para desarrollar un modelo más complejo y eficiente.
- Integrar HIS para la gestión eficiente de disponibilidad y turnos.

Métricas de Seguimiento:

- Recuperación de Confianza: Calificación en Google de 2.8 a 4.0.
- Reducción de quejas: Disminución del 47% de aumento en quejas a un -10% respecto al nivel inicial.

Indicadores de alerta temprana:

- Aumento de quejas en un 15%.
- Aumento de reclamos legales en el primer trimestre.
- Aumento de valoraciones negativas online.

Protocolo de escalamiento:

- Desactivar el bot si las respuestas siguen siendo confusas o erróneas. Contactar con los técnicos para desarrollar nuevos ajustes
- Convocar al comité de ética si los reclamos aumentan un 5%

Para profesionales del sector:

Recomendaciones:

- Capacitación acerca de los límites de la inteligencia artificial en el ámbito clínico.
- Comprender el funcionamiento del chatbot.
- Participar en la fase del entrenamiento del chatbot validando los datos y las respuestas que se desarrollan.
- Participación activa en las pruebas beta y aportación de términos específicos de la medicina.

Métricas de seguimiento:

- Porcentaje alto de profesionales que participan activamente en el diseño del chatbot aportando conocimiento técnico.
- Número de errores o mejoras sugeridas por el personal médico durante las pruebas.

Indicadores de Alerta Temprana:

- Falta de participación de más del 50% del personal en las fases de diseño.
- Aparición de errores en la información a pesar de la participación de profesionales de la salud.

Protocolos de Escalamiento:

- Cualquier respuesta incorrecta detectada por un profesional debe ser reportada inmediatamente al equipo técnico.

Para proveedores:

Recomendaciones:

- Desarrollar modelos especializados en salud.
- Entrenar con conjuntos de datos médicos válidos.

Métricas de validación:

- 80% de respuestas válidas y completas.

- Latencia de respuestas baja.

Indicadores de Alerta Temprana:

- Errores en las pruebas internas.
- Fallas en la integración con sistemas HIS durante el desarrollo.

Protocolos de Escalamiento:

- Reentrenar el modelo con la aparición de errores críticos en las pruebas beta.
- Frenar la implementación si las respuestas erróneas superan el 20%

Para Reguladores del Sector Salud:

Recomendaciones:

- Exigir estándares de encriptación y gestión de la privacidad de datos de pacientes.
- Desarrollar guías y normativas claras que aborden específicamente el uso de la IA generativa en servicios de salud, diferenciándose de otros tipos de IA.
- Crear protocolos para definir la responsabilidad legal en caso de errores de la IA que causen daño al paciente.
- Financiar la investigación sobre la aplicación segura y ética de la IA en salud y promover la formación de profesionales en estas áreas.

Métricas de seguimiento:

- Reducción del número de reclamos legales o incidentes de seguridad relacionados con IA.

Indicadores de alerta temprana:

- Aumento de incidentes de seguridad o privacidad de datos reportados por clínicas.
- Aumento de reclamos legales por violación de datos sensibles.

Protocolos de Escalamiento:

- Revisión de las normativas en caso de reportar un alto porcentaje de fallas en las implementaciones.
- Frenar temporalmente las implementaciones en caso de reportar un alto porcentaje de reclamos legales e interferir en el desarrollo.

Conclusión:

El caso de MediBot demuestra que la IA por sí sola no garantiza soluciones: sin planificación rigurosa, validación médica y participación activa de usuarios y profesionales, la tecnología puede empeorar los problemas. El éxito en salud no depende solo de la inversión, sino de una metodología centrada en el paciente, con supervisión humana, seguridad de datos y mejora continua.

BIBLIOGRAFIA:

“Costo de desarrollo de un chatbot de atención médica como AMIE de Google: “ Hyena information. 16 jul 2024

<https://www.hyena.ai/cost-to-develop-a-healthcare-chatbot-like-googles-amie/>

“¿Qué son los chatbots híbridos?” Botpenguin

<https://botpenguin.com/glossary/hybrid-chatbot>

“7 malos ejemplos de chatbot: fallos de los chatbots y cómo evitarlos”. Trengo. 19 sept 2024

<https://trengo.com/es/blog/bad-chatbot-examples>

“Las limitaciones de la IA generativa, según Generative AI”. Lingaro Group

<https://lingarogroup.com/blog/the-limitations-of-generative-ai-according-to-generative-ai>

“La guía definitiva de los chatbots PNL en 2025.” Botpress. 27 jun 2025

<https://botpress.com/es/blog/nlp-chatbot>