



Análisis de caso GenIA fallido: MediBot

Información del estudiante

Nombre y Apellido

Jonathan, Molina

DNI

32.623.833

Carrera

TSCDIA

Comisión

B - viernes

Repositorio

[¡Ir al repositorio!](#)

Docente

Diana, Accietto

A) Investigación profunda del caso

Causas raíz identificadas

Aspectos técnicos

Se considera que la tecnología seleccionada para el desarrollo del chatbot, en principio, fue adecuada; sin embargo, no hubo un diseño e implementación correctos. No se realizó una integración profunda con el sistema hospitalario, no se garantizó la memoria de contexto y el fine-tuning fue escaso, lo cual generó que el modelo no tuviera conocimiento del lenguaje local ni de la terminología médica requerida. Por otra parte, se falló en el mecanismo de escalamiento para derivar la consulta a operadores humanos cuando era necesario.

La combinación de estas falencias generó un resultado insatisfactorio.

Aspectos organizacionales

Se observó que, en términos organizacionales, el proyecto se vio afectado por la falta de coordinación entre los distintos departamentos (IT, personal médico y administrativo), siendo esto un aspecto fundamental para garantizar un resultado acorde a los requerimientos. Por otra parte, se puede deducir que hubo una definición poco clara de los roles y las responsabilidades, ya que no se estableció quién debía supervisar la validación médica de las respuestas del bot, quién coordinaba la integración con el sistema hospitalario ni quién gestionaba la experiencia de los pacientes, lo que contribuyó a fallas operativas y a la sobrecarga del personal humano.

Por último, se puede inferir que hubo una ausencia de políticas internas capaces de orientar la adaptación del chatbot a los procesos existentes, ya que no se establecieron lineamientos claros para la sincronización con el sistema de turnos, la validación de la información médica, el manejo de datos sensibles ni la derivación de consultas complejas a operadores humanos. Esta falta de definición de políticas contribuyó a que se presentasen inconsistencias en la operación del bot y afectó la experiencia de los pacientes.

Gestión de proyecto

Se observó que en términos de gestión de proyecto, la planificación fue insuficiente como así también no se tuvo una visión global de lo que se esperaba del chatbot.

Se puede inferir que esto pudo tener que ver con plazos poco realistas o bien no hubo un buen manejo/planificación de tareas de los distintos roles intervinientes.

Tampoco las pruebas piloto realizadas fueron suficientes y evidentemente no se definieron criterios claros de aceptación, ni planes de mitigación de riesgos, esto trajo como resultado que no se identificaron fallas críticas antes de lanzar el chatbot.

Factores humanos

Se observó que principalmente los pacientes eran adultos mayores (65%) y dado que este grupo etario presenta brechas tecnológicas tanto de acceso como de uso, no se los tuvo en cuenta al momento de diseñar la interfaz de usuario para que esta sea más amigable. Esto decantó en que los niveles de insatisfacción aumentasen.

B) ¿Qué pudo prevenirse?

Problemas prevenibles con mejor planificación, testing o metodología

Muchos de los problemas enfrentados por el proyecto MediBot eran prevenibles mediante una planificación más rigurosa, pruebas piloto exhaustivas y la aplicación de metodologías ágiles. La falta de definición clara del alcance, plazos poco realistas y pruebas limitadas contribuyeron directamente a que el bot no cumpliera con los requerimientos del cliente.

Red flags ignoradas, riesgos no evaluados y mejores prácticas no aplicadas

Entre las red flags ignoradas se encuentran la insuficiencia de la base de conocimiento, la falta de integración profunda con el sistema hospitalario y la ausencia de mecanismos de escalamiento a operadores humanos. No se evaluaron correctamente riesgos críticos como la entrega de información médica incorrecta, la sobrecarga del personal de soporte y la dificultad de interacción para pacientes adultos mayores. Además, no se aplicaron buenas prácticas como pruebas piloto con retroalimentación de usuarios reales, monitoreo en tiempo real y validación de contenido médico por profesionales calificados.

Plan de prevención de riesgos específico

Para proyectos similares, se recomienda definir desde el inicio roles y políticas claras, implementar pruebas piloto progresivas con retroalimentación constante de pacientes y profesionales, monitorear indicadores de desempeño del bot en tiempo real y establecer protocolos de escalamiento y validación de contenido médico antes del lanzamiento oficial. También resulta fundamental anticipar la capacitación del personal y diseñar la experiencia de usuario considerando a los pacientes con mayor dificultad para interactuar con la tecnología.

C) ¿Qué no pudo prevenirse?

Aspectos del fracaso inherentes a limitaciones tecnológicas

Algunos problemas del proyecto MediBot no pudieron prevenirse debido a las limitaciones actuales de la tecnología GenAI utilizada. Por ejemplo, el modelo GPT-3.5, incluso con fine-tuning, tiene restricciones en comprensión de jerga local, manejo de contextos complejos y especialización en terminología médica. Estas limitaciones sistémicas impidieron que el bot pudiera responder con precisión todas las consultas, aun cuando se aplicaran buenas prácticas de desarrollo y capacitación.

Restricciones presupuestarias y externas

Además, restricciones presupuestarias realistas limitaron la posibilidad de entrenar modelos más avanzados, contratar especialistas en NLP médico o integrar el bot de manera más profunda con el HIS. Factores externos, como cambios en la disponibilidad de turnos, variaciones en la demanda de consultas y características demográficas de los pacientes (mayoría adultos mayores), también contribuyeron al impacto negativo sin estar bajo el control directo del equipo.

Diferenciación entre problemas evitables y limitaciones tecnológicas

Si bien algunos fallos fueron prevenibles mediante mejor planificación, testing y protocolos internos, otros correspondieron a limitaciones sistémicas de la tecnología y al contexto externo. La distinción permite reconocer que, aunque la metodología y gestión del proyecto podían mejorar muchos resultados, ciertas restricciones de GenAI y condiciones del entorno eran inevitables con los recursos y herramientas disponibles en ese momento.

D) ¿Cómo trabajaría en el futuro?

Para futuras implementaciones de GenAI en salud se propone un enfoque más estructurado.

En primer lugar, establecer un framework de viabilidad que combine análisis técnico, regulatorio, clínico y económico, de modo que antes de iniciar el desarrollo quede claro si el proyecto es realista.

En segunda instancia, aplicar un testing escalonado: pruebas controladas en entornos simulados, luego pilotos reducidos y recién después una implementación amplia, asegurando métricas claras en cada etapa.

Los criterios de éxito o fallo deben incluir precisión clínica, impacto en la seguridad del paciente, aceptación de profesionales de salud y sostenibilidad financiera.

Por otra parte, se recomienda una gestión del cambio activa con formación continua, comunicación transparente y espacios de retroalimentación con médicos y usuarios para reducir la resistencia.

Por último, se sugiere un plan de contingencia con protocolos de rollback, redundancias técnicas y apoyo humano alternativo en caso de que la IA falle.

Todo esto debe ir acompañado de un monitoreo continuo y flexible, considerando los avances tecnológicos esperados en interpretabilidad, interoperabilidad y regulaciones específicas de IA en salud.

E) Lecciones personales aprendidas

De este caso me llevo varias lecciones que considero clave para mi desarrollo profesional.

En primer lugar, comprendí la importancia de entender y detectar las condiciones y necesidades reales del proyecto: presupuesto, tiempos, infraestructura y capacidad de los equipos.

También detecté que se identificaron ciertas limitaciones técnicas demasiado tarde y eso me demostró que es fundamental establecer pautas de evaluación más realistas y que sean progresivas para evitar comprometer recursos de gran escala.

Otra lección importante es que la implementación de la IA, especialmente en el sector salud, no puede desligarse del plano de la ética y de la responsabilidad profesional. Más allá de los resultados técnicos, lo que está en juego son decisiones que afectan directamente a personas, lo que implica ser consiente de ello en el ejercicio de la profesión, con rigurosidad en la validación de datos, en la contrucción de algoritmos eficientes y en la comunicación de riesgos.

En cuanto a las competencias a desarrollar, me gustaría ganar experiencia que me permita tener una visión general de todas las necesidades que atraviesa un proyecto: desde la gestión y la planificación hasta el diseño y la implementación. Creo que, independientemente del rol que me pueda tocar ocupar dentro de una empresa, esa perspectiva global me daría la capacidad de aportar de manera más sólida al equipo, colaborar de forma efectiva y al mismo tiempo cerciorarme de que cada etapa se cumpla según lo previsto, en pos de garantizar los resultados y la calidad final.

En lo personal, este caso cambia mi forma de encarar un proyecto ya puedo dejar de pensarlos como un desafío exclusivamente técnico, sino también como un proceso integral en el que intervienen múltiples actores, con responsabilidades compartidas y consecuencias que van más allá de la innovación.

F) Recomendaciones estratégicas

Para la organización afectada, las métricas de seguimiento deberían centrarse en la calidad del servicio y la seguridad de los pacientes, midiendo la tasa de diagnósticos correctos, la eficiencia en los tiempos de respuesta y la satisfacción de los usuarios. Como indicadores de alerta temprana, resultan clave las caídas en la precisión del sistema, el aumento de quejas recurrentes y los retrasos en la atención. En cuanto a protocolos de escalamiento, es necesario que existan canales de reporte inmediato a los responsables técnicos y clínicos, junto con planes de contingencia que garanticen la continuidad del servicio sin comprometer la atención médica.

En relación con los profesionales del sector, es fundamental que adopten un enfoque de formación continua en el uso de IA aplicada a la práctica clínica. La participación en capacitaciones con



simulaciones y casos prácticos puede servir como métrica de preparación y como refuerzo de la confianza en la herramienta. Como señales de alerta, deben observarse la falta de actualización de conocimientos o la resistencia al uso de la tecnología. El protocolo de escalamiento debería incluir tutorías, programas de apoyo y refuerzo de la capacitación, asegurando que los equipos clínicos cuenten con respaldo suficiente para integrar adecuadamente estas herramientas en su práctica cotidiana.

Para los proveedores de tecnología GenAI, resulta clave priorizar la transparencia y la explicabilidad de los modelos, ofreciendo documentación clara y procesos de mejora continua basados en métricas de desempeño real. Al mismo tiempo, las métricas de seguimiento deberían enfocarse en la estabilidad del sistema, el cumplimiento normativo y los tiempos de soporte. Como alertas tempranas se deben contemplar errores recurrentes, vulnerabilidades de seguridad y brechas en la trazabilidad de decisiones. Frente a estos casos, los protocolos de escalamiento deben prever la activación de equipos especializados de respuesta rápida, la comunicación inmediata con las organizaciones usuarias y la publicación transparente de las correcciones realizadas.

Finalmente, los reguladores del sector salud deben avanzar en la creación de marcos normativos específicos para la aplicación de IA, incorporando métricas de seguridad, criterios de validación clínica y estándares internacionales. Como indicadores de alerta temprana, se deben considerar las irregularidades detectadas en auditorías, los incumplimientos de protocolos y la reiteración de denuncias vinculadas a un mismo sistema. En cuanto a protocolos de escalamiento, resulta necesario establecer mecanismos claros que obliguen a reportar incidentes, definan responsabilidades entre organizaciones, proveedores y profesionales, y, en situaciones críticas, permitan suspender de manera temporal el uso de la tecnología hasta garantizar su confiabilidad.