



UNIVERSIDAD TECNOLÓGICA CENTROAMERICANA

FACULTAD DE INGENIERÍA

PROYECTO DE INVESTIGACIÓN

**ESTUDIO SOBRE LA APLICABILIDAD Y USABILIDAD DE LOS CHATBOTS EN LAS
ACTIVIDADES DEL INGENIERO CLÍNICO EN UN ENTORNO HOSPITALARIO**

PREVIO A LA OBTENCIÓN DEL TÍTULO

INGENIERO EN BIOMÉDICA

PRESENTADO POR:

21811087 KARLYN ELIZA CASTELLANOS ARIAS

21851108 VALERIA MONTSERRAT COTO QUIJADA

ASESOR: MANUEL GAMERO

CAMPUS SAN PEDRO SULA;

OCTUBRE, 2023

DEDICATORIA

A mi familia, quienes siempre han sido mi mayor apoyo y fuente de inspiración. Gracias por creer en mí y por estar a mi lado en cada etapa de esta larga travesía.

-Karlyn Eliza Castellanos Arias

Con mucho amor, para Dios, mis padres y mis hermanos, por ser los cimientos de mi vida.

-Valeria Montserrat Coto Quijada

AGRADECIMIENTOS

Expreso mi gratitud a Dios por permitirme alcanzar todas las metas que me he propuesto y por cuidar tanto de mi familia como de mí a lo largo de este extenso trayecto. También, agradezco profundamente a mis padres por respaldar cada elección y proyecto que he emprendido; su amor incondicional ha sido un pilar fundamental que ha facilitado superar cada desafío. A mis hermanos, les agradezco por estar a mi lado constantemente y por el cariño que siempre me brindan. Además, quiero reconocer a mis amigos y compañeros, quienes han estado presentes en cada etapa, fortaleciéndome día a día. Mis maestros merecen mi agradecimiento por su contribución a mi formación académica y profesional, y por recordarme siempre que tengo la capacidad de mejorar constantemente.

-Karlyn Eliza Castellanos Arias

A Dios, por todo lo que soy y todo lo que seré. A mi padre, Jorge Coto, cuyas enseñanzas me formaron y que sigo atesorando en mi corazón día a día. A mi madre, Sandra Quijada, por ser mi mayor aspiración y enseñarme a ser la mejor versión de mí. A mis hermanos, Jorge y Xavier Coto, por influenciar mi vida de la mejor manera posible y motivarme a seguir su ejemplo. A mis sobrinas, Jimena y Layla Coto, por llenar de alegría mi vida. Gracias familia por amarme, guiarme y apoyarme a través de todos mis proyectos. Las palabras nunca serán suficientes para expresar mi gratitud.

A Reyna Rivas, por su acompañamiento a lo largo de mi carrera universitaria. A José Barahona, por alentarme y apoyarme en todo momento. Gracias por su cariño y por ser mis aliados en esta etapa.

Finalmente, a todas las personas que me aman y me han amado, pues hoy, me debo a ustedes.

-Valeria Montserrat Coto Quijada

A nuestro asesor, Manuel Gamero, por su instrucción y guía a lo largo de este proyecto. A todos los ingenieros que nos brindaron su tiempo para permitirnos realizar esta investigación. Gracias por su colaboración. Por último, a nosotras mismas, por todo el esfuerzo realizado y plasmado en este proyecto.

-Karlyn Castellanos & Valeria Coto

EPÍGRAFE

"Nada puede atenuar la luz que brilla desde adentro."

Maya Angelou

RESUMEN EJECUTIVO

Los chatbots en el ámbito de la salud representan aplicaciones de inteligencia artificial (IA) diseñadas con la finalidad de proporcionar orientación médica, apoyo emocional, información relacionada con la salud y respuestas a consultas vinculadas al bienestar y la atención médica a través de diálogos en línea. Estos sistemas de chatbot encuentran aplicabilidad en diversos contextos dentro del campo de la salud y la atención médica, con la capacidad de conferir una serie de ventajas, aunque al mismo tiempo enfrentan retos y consideraciones críticas. En este contexto, se ha empleado la metodología en "V" con el propósito de desarrollar un sistema para la creación de métricas que sean más apropiadas para evaluar las diferentes formas de inteligencia artificial previstas para operar en el proyecto, además de definir las variables de evaluación. El enfoque de esta investigación es de tipo mixto, con un diseño preexperimental y un alcance descriptivo. A lo largo de esta investigación, se ha destacado la utilidad potencial de los chatbots en el ámbito de la ingeniería clínica para facilitar la toma de decisiones; pues su arquitectura y entrenamiento permiten un acceso más eficiente a la tecnología. Además, se ha examinado la infraestructura necesaria para implementar los chatbots evaluados, resaltando su facilidad de acceso al requerir únicamente de un dispositivo móvil y conexión a internet. Asimismo, destaca su capacidad para desarrollarse de manera orgánica y sin complicaciones al mantener una conversación. La implementación de estos chatbots, incluyendo aspectos como el aprendizaje, la mejora en la búsqueda de información y la selección de bases de datos, ha sido destacada como un resultado significativo durante la experimentación de este proyecto. Aunque la implementación de estas formas de inteligencia artificial se espera que sea gradual y pueda llevar tiempo, es una tendencia que podría materializarse en un futuro cercano. Estos sistemas podrían ser un apoyo para los ingenieros que trabajan en la ingeniería clínica, especialmente en el mantenimiento de equipos y sistemas médicos; al facilitar el acceso a información y ser capaz de crear herramientas complementarias para automatizar partes repetitivas del trabajo realizado.

Palabras claves: aplicabilidad, biomédica, ingeniería clínica, inteligencia artificial, chatbot.

ABSTRACT

Healthcare chatbots represent artificial intelligence (AI) applications designed to provide medical guidance, emotional support, health-related information and answers to wellness and healthcare queries through online dialogues. These chatbot systems find applicability in various contexts within the field of health and medical care, with the ability to confer several advantages, yet at the same time face challenges and critical considerations. In this context, the "V" methodology has been employed for the purpose of developing a system for the creation of metrics that are most appropriate for evaluating the different forms of artificial intelligence envisioned to operate in the project, in addition to defining the evaluation variables. The approach of this research is of a mixed type, with a pre-experimental design and a descriptive scope. Throughout this research, the potential usefulness of chatbots in the field of clinical engineering to facilitate decision making has been highlighted, as their architecture and training allow more efficient access to the technology. In addition, the infrastructure necessary to implement the evaluated chatbots has been examined, highlighting their ease of access by requiring only a mobile device and internet connection. It also highlights their ability to develop organically and without complications when maintaining a conversation. The implementation of these chatbots, including aspects such as learning, improved information search and database selection, has been highlighted as a significant result during the experimentation of this project. Although the implementation of these forms of artificial intelligence is expected to be gradual and may take time, it is a trend that could materialize soon. These systems could be a support for engineers working in clinical engineering, especially in the maintenance of medical equipment and systems; by facilitating access to information and being able to create complementary tools to automate repetitive parts of the work performed.

Keywords: applicability, biomedical, clinical engineering, artificial intelligence, chatbot.

ÍNDICE DE CONTENIDO

I. INTRODUCCIÓN.....	3
II. PLANTEAMIENTO DEL PROBLEMA	4
2.1 PRECEDENTES DEL PROBLEMA.....	4
2.2 DEFINICIÓN DEL PROBLEMA.....	5
2.3 JUSTIFICACIÓN	5
2.4 PREGUNTAS DE INVESTIGACIÓN.....	6
2.4.1 PREGUNTA GENERAL	6
2.4.2 PREGUNTAS ESPECÍFICAS	6
2.4.3 OBJETIVO GENERAL	7
2.4.4 OBJETIVOS ESPECÍFICOS.....	7
III. ESTADO DEL ARTE	8
3.1 ANÁLISIS DE LA SITUACIÓN ACTUAL.....	8
3.1.1 MACROENTORNO.....	8
3.1.2 MICROENTORNO	13
3.2 CONCEPTUALIZACIÓN	19
3.2.1 CHATGPT	19
3.2.2 GOOGLE BARD	20
3.2.3 CHATSONIC	21
3.2.4 PERPLEXITY	21
3.2.5 INGENIERÍA CLÍNICA.....	22
3.3 TEORÍAS DE SUSTENTO.....	24
3.3.1 BASES TEÓRICAS	24
3.3.2 METODOLOGÍAS DESARROLLADAS POR OTROS INVESTIGADORES.....	29
3.3.3 INSTRUMENTOS UTILIZADOS.....	30
3.4 MARCO LEGAL	31
3.4.1 LEGISLACIÓN INTERNACIONAL.....	31
3.4.2 LEGISLACIÓN REGIONAL.....	32
IV. METODOLOGÍA	33
4.1 ENFOQUE	33
4.2 VARIABLES DE INVESTIGACIÓN.....	33
4.2.1 VARIABLE DEPENDIENTE.....	33

4.2.2	VARIABLES INDEPENDIENTES.....	34
4.3	TÉCNICAS E INSTRUMENTOS UTILIZADOS.....	35
4.3.1	TÉCNICAS.....	35
4.3.2	ELECCIÓN DE INSTRUMENTOS	37
4.3.3	INSTRUMENTOS	38
4.4	MUESTRA.....	45
4.5	METODOLOGÍA DE ESTUDIO.....	45
4.5.1	DEFINIR PAUTAS EN EL PROCESO DE EVALUACIÓN DE LOS CHATBOTS.....	46
4.5.2	EVALUACIÓN Y CREACIÓN DEL CUESTIONARIO JUNTO CON LA EVALUACIÓN DE USABILIDAD BASADA EN HEURÍSTICAS.....	48
4.5.3	ESTABLECER CRITERIOS PARA LAS PRUEBAS	50
4.5.4	EXPLORACIÓN DEL EMPLEO DE CHATBOTS EN LAS ACTIVIDADES DEL INGENIERO CLÍNICO	52
4.6	METODOLOGÍA DE VALIDACIÓN.....	52
4.7	CRONOGRAMA DE ACTIVIDADES	53
4.8	MATRIZ METODOLÓGICA.....	54
V.	RESULTADOS Y ANÁLISIS	56
5.1	EVALUACIÓN DE METODOLOGÍA EN V	56
5.1.1	DEFINICIÓN DE PAUTAS PARA EL PROCESO DE EVALUACIÓN DE CHATBOTS.....	56
5.1.2	ELABORACIÓN Y CREACIÓN DE CUESTIONARIOS JUNTO CON LA EVALUACIÓN DE USABILIDAD	57
5.1.3	ESTABLECER CRITERIOS DE EVALUACIÓN PARA PRUEBAS.....	57
5.2	ANÁLISIS ESTADÍSTICO	59
5.2.1	ENCUESTA DE DIAGNÓSTICO.....	59
5.1.2	HEURÍSTICA DE USABILIDAD	64
5.1.3	ENCUESTA DE ACEPTACIÓN	86
5.1.4	COSTOS OPERACIONALES	96
VI.	CONCLUSIONES.....	99
VII.	RECOMENDACIONES.....	100
VIII.	IMPLEMENTACIÓN.....	101
IX.	EVOLUCIÓN DEL TRABAJO ACTUAL	102
BIBLIOGRAFÍA	103
ANEXOS	109

ÍNDICE DE ILUSTRACIONES

Ilustración 1 - Diagrama de la relación entre IA, ML y DL.	9
Ilustración 2 – Conceptos de la IA y preguntas que plantean.	12
Ilustración 3 - Modalidades de datos y oportunidades para la IA biomédica multimodal	15
Ilustración 4 – Ejemplo de una Red Neuronal.	25
Ilustración 5 - Diagrama de variables de investigación.....	35
Ilustración 6 - Logo de ChatGPT.	39
Ilustración 7 - Flujograma del funcionamiento de ChatGPT.....	40
Ilustración 8 - Logo de Google Bard.....	41
Ilustración 9 - Flujograma del funcionamiento de Google Bard.....	41
Ilustración 10 - Logo de Chatsonic.	42
Ilustración 11 - Flujograma del funcionamiento de Chatsonic.....	42
Ilustración 12 - Logo de Perplexity.....	43
Ilustración 13 - Flujograma del funcionamiento de Perplexity.....	44
Ilustración 14 - Metodología en V.	52
Ilustración 15 - Línea de tiempo del proceso de desarrollo y formulación de evaluación de encuestas.	56
Ilustración 16 - Habilidades de los chatbots útiles en el ámbito de la ingeniería clínica.	58
Ilustración 17 – Calificación de Likert de la encuesta de diagnóstico.....	64
Ilustración 18 - Visibilidad del estado del sistema.....	67
Ilustración 19 - Correspondencia entre el sistema y el mundo real.....	69
Ilustración 20 - Control y libertad del usuario.	71
Ilustración 21 - Consistencia y normas.....	73
Ilustración 22- Prevención de errores.	75
Ilustración 23 - Ayuda y orientación.	77
Ilustración 24 - Flexibilidad y eficacia de uso.....	78
Ilustración 25 - Diseño estético, minimalista y atractivo.	80
Ilustración 26 - Ayudar a los usuarios a reconocer, diagnosticar y recuperarse de los errores.	82

Ilustración 27 - Conservación del contexto.	84
Ilustración 28 - Fiabilidad.	86
Ilustración 29 - Datos cuantitativos sobre la aceptación.	91
Ilustración 30 – Comparativa de la media de las características evaluadas en la heurística.....	91
Ilustración 31 - Calificación general de cada chatbot de acuerdo con la media.	92
Ilustración 32 - Características más valiosas de los chatbots según la opinión.....	93
Ilustración 33 - Características esenciales para que un chatbot sea efectivo en el ámbito de la ingeniería clínica.....	94
Ilustración 34 - Chatbot más recomendado según la muestra.....	95
Ilustración 35 - Dificultad al utilizar chatbots de acuerdo con la opinión de la muestra.	96

ÍNDICE DE TABLAS

Tabla 1 - Comparativa entre las ventajas y desventajas ofrecidas por ChatGPT.....	20
Tabla 2 - Comparativa entre las ventajas y desventajas ofrecidas por Google Bard.	20
Tabla 3 - Comparativa entre las ventajas y desventajas ofrecidas por Chatsonic.....	21
Tabla 4 - Comparativa entre las ventajas y desventajas ofrecidas por Perplexity.	22
Tabla 5 - El proceso de diseño de cuatro fases.....	29
Tabla 6 - Resumen de las metodologías utilizadas por otros investigadores.....	30
Tabla 7 - Metodología de la investigación.	33
Tabla 8 - Características por la que destaca los instrumentos elegidos.	38
Tabla 9 - Cronograma de actividades.....	53
Tabla 10 - Matriz metodológica.....	54
Tabla 11 - IB con acceso a una computadora.....	59
Tabla 12 - IB con acceso a Wifi/Internet.....	60
Tabla 13 - Uso previo de chatbots.....	60
Tabla 14 - Uso previo de otras IA.....	60
Tabla 15 - Nivel de conocimiento de las TICs de los IB.....	61
Tabla 16 - Infraestructura tecnológica.....	61
Tabla 17 - Factores que influyen el uso de chatbots.....	62
Tabla 18 - Implementación de chatbots y facilitación del trabajo.....	62
Tabla 19 - Análisis estadístico comparativo de datos cualitativos.....	63
Tabla 20 - Análisis estadístico comparativo de la visibilidad del sistema.....	66
Tabla 21 - Análisis estadístico comparativo de la correspondencia entre el sistema y el mundo real	68
Tabla 22 - Análisis estadístico comparativo del control y la libertad del usuario.....	70
Tabla 23 - Análisis estadístico comparativo de la consistencia y normas.....	72
Tabla 24 - Análisis estadístico comparativo de la prevención de errores.....	74
Tabla 25 - Análisis estadístico comparativo sobre la ayuda y orientación.....	76
Tabla 26 - Análisis estadístico comparativo sobre la flexibilidad y eficacia de uso.....	78
Tabla 27 - Análisis estadístico comparativo sobre el diseño.....	79
Tabla 28 - Análisis estadístico comparativo sobre la ayuda a los usuarios para lidiar con errores.....	81

Tabla 29 - Análisis estadístico comparativo sobre la conservación del contexto.	83
Tabla 30 - Análisis estadístico comparativo sobre la fiabilidad.	85
Tabla 31 - Empleo de chatbots como herramienta en sus actividades cotidianas.	86
Tabla 32 - Implementarlos facilitará su trabajo.	87
Tabla 33 - Frecuencia con el que los implementaría en el trabajo.	87
Tabla 34 - Recomendación para su uso en unidad de biomédica.	88
Tabla 35 - Nivel de confianza en chatbots.	88
Tabla 36 - Preocupación por la privacidad.	88
Tabla 37 - Análisis estadístico de variables cuantitativas sobre la aceptación.	90
Tabla 38 - Costos operacionales ChatGPT.	97
Tabla 39 - Costos operacionales de Google Bard.	97
Tabla 40 - Costos operacionales de Perplexity.	98
Tabla 41 - Costos operacionales de ChatSonic.	98

ÍNDICE DE ANEXOS

Anexo 1 - Tres categorías de problemas éticos de la inteligencia artificial.....	109
Anexo 2 - La heurística de Nielsen comparada con la heurística final creada para el agente conversacional por Langevin et. al.....	110
Anexo 3 - Encuesta de diagnóstico aplicada.	113
Anexo 4 – Cuarta prueba con el chatbot ChatGPT.....	116
Anexo 5 -Décima prueba con el chatbot ChatGPT.....	116
Anexo 6 - Segunda prueba con el chatbot Google Bard.....	117
Anexo 7 - Quinta prueba con el chatbot Google Bard.	117
Anexo 8 - Historial de búsquedas con el chatbot Perplexity AI.....	118
Anexo 9 - Sexta prueba con el chatbot Perplexity AI.....	118
Anexo 10 - Segunda prueba con el chatbot ChatSonic.....	119
Anexo 11 - Novena prueba con el chatbot ChatSonic.....	119
Anexo 12 - Heurística aplicada.....	120
Anexo 13 - Encuesta de aceptación aplicada al finalizar las pruebas.	121
Anexo 14 - Matriz bibliográfica	126

LISTA DE SIGLAS

AGI	Artificial General Intelligence
DL	Deep Learning
GPT	Generative Pre-trained Transformer
IA	Inteligencia Artificial
IB	Ingeniero Biomédico
LLM	Large Language Model
NLP	Natural Language Processing
ML	Machine Learning
RLHF	Reinforcement Learning with Human Feedback
TIC	Tecnologías de la Información y las Comunicaciones

GLOSARIO

- Automatización: proceso en el que se llevan a cabo tareas o actividades de manera automática, no requiere de intervención directa de los humanos (Corvalán, 2019).
- Inteligencia artificial fuerte (AGI): inteligencia artificial que posee una capacidad cognitiva igual o superior a la inteligencia humana (Comisión Europea, 2019).
- Chatbot o agente conversacional: programa de inteligencia artificial creado para interactuar con los usuarios a través de una conversación (Zeroual & Lakhouaja, 2018).
- Chunks: del inglés "trozo" o "pedazo", es un fragmento de información, contenido en muchos formatos multimedia (Rodríguez-Blanco et al., 2016).
- Tokens: palabras reservadas de un lenguaje, secuencia de caracteres que representa una unidad de información en el programa fuente (Jiang et al., 2020).

I. INTRODUCCIÓN

Se espera que la inteligencia artificial (IA) evolucione el futuro. Pues esta, está siendo cada vez más utilizada en diversos ámbitos, incluyendo la salud y la ingeniería biomédica. La biomédica, al ser un sector de estudio en constante evolución, requiere de adaptación a las nuevas tecnologías y tendencias. Es por esto, que la aplicabilidad de chatbots podría beneficiar de gran manera a los profesionales en el ámbito.

A pesar de las notables cualidades y ventajas que los chatbots podrían brindar, como la reducción de costos económicos y la mejora de la eficiencia y la efectividad, se presentan riesgos vinculados a su implementación. Entre estos desafíos, se pueden señalar la confiabilidad y la seguridad de los datos. Además, se ven implicados varios factores éticos relacionados con su utilización.

La investigación busca explorar el uso de agentes conversacionales en el ámbito sanitario, específicamente, en las labores diarias de los ingenieros biomédicos dedicados a la ingeniería clínica. Esto se hará con diversas pruebas y evaluaciones cualitativas y cuantitativas, a las que se someterá una muestra de profesionales seleccionados.

En el siguiente documento, se presentarán los fundamentos teóricos y su conexión con la evaluación de los chatbots. Se establecerá la información pertinente sobre la metodología empleada siguiendo el modelo en "V". Asimismo, se destacarán los resultados de los análisis estadísticos realizados. Para finalizar, se expondrán las conclusiones y se ofrecerán recomendaciones para proyectos futuros.

II. PLANTEAMIENTO DEL PROBLEMA

En este capítulo, se exponen todos los detalles relacionados al planteamiento del problema de la investigación. Se introducen precedentes sobre el tema y se demuestra la justificación para el abordaje de este. Posteriormente, se plantean las preguntas de investigación, que finalmente se transforman en los objetivos de esta.

2.1 PRECEDENTES DEL PROBLEMA

Los chatbots representan aplicaciones de inteligencia artificial (IA) con la capacidad de sostener conversaciones naturales en lenguaje humano con los usuarios. En el ámbito de la atención médica, los chatbots han experimentado un desarrollo continuo y se han empleado para múltiples propósitos. Su función principal radica en proporcionar información relevante y oportuna a los pacientes, lo cual puede resultar esencial en situaciones críticas, como el acceso a recursos relacionados con la salud mental. Desde la concepción inicial del primer chatbot a finales de la década de 1960, se han dedicado considerables esfuerzos a la creación de chatbots especializados en distintos aspectos de la atención médica, empleando diversas metodologías y enfoques de desarrollo (Safi et al., 2020).

Desde una perspectiva más amplia, la inteligencia artificial se visualiza como una tecnología horizontal con aplicaciones de relevancia en diversos sectores y ramas industriales. Cuando se analiza en conjunto con el análisis de macrodatos (big data) y la ciberseguridad, estas tres tecnologías fundamentales conforman lo que se ha denominado el "ABC de la tecnología". Se estima que estas tecnologías se encuentran en un punto de inflexión y se espera que experimenten una adopción más rápida en los próximos años, a medida que empresas y gobiernos enfocan sus recursos y esfuerzos en estas.

La creación de chatbots de dominio abierto se presenta como un desafío complejo en el ámbito de la investigación en aprendizaje automático. A pesar de la tendencia hacia modelos neuronales con un aumento en el número de parámetros y el tamaño de los conjuntos de datos utilizados para su entrenamiento, se reconoce la importancia de otros componentes esenciales. Una conversación efectiva no se reduce únicamente a la tecnología subyacente, sino que requiere una combinación de habilidades, como la capacidad para generar temas de conversación atractivos, transmitir conocimientos de manera efectiva, expresar empatía y personalidad, y mantener una coherencia global en la interacción (Roller et al., 2021).

La incorporación de chatbots en las responsabilidades del ingeniero clínico en un entorno hospitalario posee un potencial significativo para mejorar la eficiencia en la gestión de equipos

médicos y recursos técnicos. Estos sistemas automatizados pueden proporcionar respuestas rápidas a preguntas frecuentes y pueden apoyar en la resolución de problemas técnicos. Para garantizar la maximización de los beneficios de esta tecnología, es imperativo proporcionar una capacitación adecuada a los ingenieros clínicos, y asegurar soporte técnico en caso de consultas más complejas o problemas en la implementación (Safi et al., 2020).

2.2 DEFINICIÓN DEL PROBLEMA

La ingeniería biomédica es una profesión con diversos enfoques, siendo uno de ellos la ingeniería clínica. Esta rama puede presentar diversos retos como ser la carencia de instrucciones debido al extravío de manuales de servicio o la limitación de soluciones como resultado directo del conocimiento y experiencia del ingeniero o técnico encargado. La inteligencia artificial puede ser integrada a las labores operativas de una manera complementaria positiva en muchos escenarios.

2.3 JUSTIFICACIÓN

Actualmente existe un creciente interés por el uso de la inteligencia artificial a nivel mundial, específicamente aplicando el uso de chatbots. Statista (2023) informó que un total de 902 empresas y organizaciones de todo el mundo, ya utilizan productos de OpenAI, por sector industrial. OpenAI es una empresa estadounidense dedicada a la investigación de inteligencia artificial, a la cual se le atribuye la creación de ChatGPT. Este mismo informe, reportó que 24 de las empresas mencionadas, se dedican a la salud, 98 de ellas a los servicios y 251 a la tecnología. Esto resalta la importancia de investigar la aplicación de agentes conversacionales en la atención médica.

Según estimó Insider Intelligence (2022), hasta el 73% de las tareas administrativas en el ámbito de la atención médica podrían ser automatizadas por IA. Asimismo, la implementación de chatbots podría generar ahorros anuales de hasta 11 mil millones de dólares americanos para los sectores de atención médica, banca y venta minorista al finalizar el 2023. Esto implica una oportunidad para mejorar la eficiencia en el sector de la salud al contar con un indicativo financiero como motivación.

Adicionalmente, el Banco de Desarrollo de América Latina y el Caribe (2021) reportó que la IA está revolucionando el ámbito sanitario a nivel mundial. Debido a esto, se aproximó que las inversiones en IA en este sector, a nivel global, alcanzarán los USD 36 mil millones para el año 2025. Esto representaría un crecimiento del 50% en comparación al 2018. Este crecimiento

proyectado demuestra la idea de cómo los chatbots, una aplicación específica de la IA, pueden contribuir al avance en el campo de la ingeniería clínica y la atención médica en general.

Sin embargo, la OMS (2023) explicó que es necesario analizar minuciosamente los peligros asociados con la utilización de estos modelos de lenguaje en la ampliación del acceso a la información de salud, ya sea como una herramienta para respaldar la toma de decisiones o para mejorar la capacidad de diagnóstico en entornos con recursos limitados, con el objetivo de salvaguardar la salud de las personas y disminuir las desigualdades. Esta información resalta la necesidad de llevar a cabo una investigación ética y responsable, garantizando que la implementación de chatbots no genere desigualdades o ponga en peligro la salud de las personas.

Los chatbots basados en inteligencia artificial, poseen el potencial para revolucionar la atención médica. Sin embargo, también es necesario abordar los desafíos éticos y de seguridad asociados. La presente investigación busca explorar los beneficios y riesgos en la implementación de chatbots en el sector de la salud, específicamente, para uso del ingeniero clínico en un entorno hospitalario.

2.4 PREGUNTAS DE INVESTIGACIÓN

2.4.1 PREGUNTA GENERAL

¿Cuál es el impacto específico y medible en el desempeño de los ingenieros biomédicos en la zona noroccidental de Honduras en 2023, después de ser expuestos a cuatro herramientas de inteligencia artificial utilizadas para las consultas relacionadas con actividades laborales técnicas?

2.4.2 PREGUNTAS ESPECÍFICAS

- a) ¿Cómo afecta la implementación de cuatro herramientas de inteligencia artificial en el trabajo de los ingenieros biomédicos en la zona noroccidental de Honduras en 2023, en términos de eficiencia en la toma de decisiones en sus actividades laborales técnicas?
- b) ¿Cuál es la percepción de los ingenieros biomédicos en la zona noroccidental de Honduras en 2023 sobre la utilidad de las herramientas de inteligencia artificial en sus actividades laborales técnicas y cómo esto se refleja en su satisfacción en el trabajo?

- c) ¿Cuál es el chatbot más eficaz para ingenieros biomédicos en la zona noroccidental de Honduras en 2023, identificado a través de una heurística y tomando en consideración las necesidades específicas de esta profesión?

2.4.3 OBJETIVO GENERAL

Analizar de manera específica y medible en la zona noroccidental de Honduras, en 2023, el impacto en el desempeño de los ingenieros biomédicos tras su exposición a cuatro herramientas de inteligencia artificial utilizadas para consultas asociadas con actividades laborales técnicas.

2.4.4 OBJETIVOS ESPECÍFICOS

- a) Evaluar el impacto de la implementación de cuatro herramientas de inteligencia artificial en el trabajo de los ingenieros biomédicos en la zona noroccidental de Honduras en 2023, centrándose en la mejora de la eficiencia en la toma de decisiones y la reducción de errores en sus actividades laborales técnicas.
- b) Examinar la percepción de los ingenieros biomédicos en la zona noroccidental de Honduras en 2023 sobre la utilidad de las herramientas de inteligencia artificial en sus actividades laborales técnicas.
- c) Identificar el chatbot más eficaz para los ingenieros biomédicos en la zona noroccidental de Honduras en 2023, mediante la aplicación de una heurística que considere las necesidades específicas de esta profesión y que optimice la asistencia técnica proporcionada por los chatbots.

III. ESTADO DEL ARTE

En el presente capítulo se expone una recopilación de la fundamentación teórica referente a la investigación. Esto incluye conceptos, estudios y publicaciones realizadas en los últimos años. Este capítulo inicia con un análisis de la situación actual, seguido por la conceptualización y teorías de sustento, finalizando con el actual marco legal enfocado en la legislación internacional y regional. Este capítulo, por tanto, sienta las bases necesarias para la comprensión y el análisis en profundidad de la investigación que se llevó a cabo en las siguientes secciones.

3.1 ANÁLISIS DE LA SITUACIÓN ACTUAL

Se exhiben y se analizan temas trascendentes que fomentan la comprensión de la situación actual como puntos de referencia para uso y aplicación de la IA, desde sus beneficios hasta sus limitaciones.

3.1.1 MACROENTORNO

En el macroentorno se expresan los aspectos a nivel global que se encuentran relacionados al problema de investigación.

3.1.1.1 Contexto tecnológico

Rouhiainen (2018) describió a la IA como la habilidad de los ordenadores de realizar actividades que requieren del intelecto humano. Esto se puede detallar como la capacidad de los equipos para hacer uso de algoritmos, aprender datos y hacer uso de este aprendizaje para tomar decisiones; de la misma manera como lo haría un ser humano. La diferencia y ventaja que posee la IA yace en que esta es capaz de analizar grandes volúmenes de información y reducir el porcentaje de error (Abid Haleem et al., 2022).

La IA se subdivide en diversos enfoques (Ilustración 1), uno de ellos siendo el aprendizaje automático (ML). Este consiste en el uso de algoritmos para que las máquinas aprendan de los patrones de datos, es decir, brindan una experiencia personalizada basada en la manera en que se utilizan. Un subcampo del aprendizaje automático es el aprendizaje profundo (DL). El mencionado anteriormente, se produce por medio del uso de redes neuronales, las cuales se organizan en múltiples capas para poder reconocer relaciones y patrones complejos en los datos (Rouhiainen, 2018). Para ser aplicado, se requiere contar con enormes conjuntos de información y una alta capacidad de procesamiento.

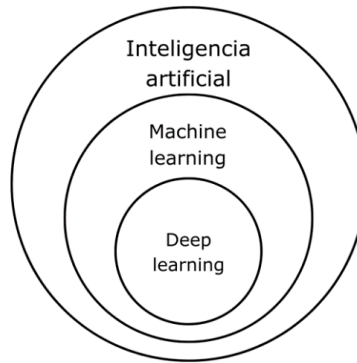


Ilustración 1 - Diagrama de la relación entre IA, ML y DL.

Fuente: Elaboración propia.

Más recientemente, los avances en el DL incluyen los modelos de lenguaje de gran tamaño (LLM). Los LLM son entrenados con inmensas cantidades de datos de texto, de esta manera se vuelven capaces de generar textos similares a los del lenguaje natural humano y así, contestar preguntas y realizar tareas solicitadas (Kasneci et al., 2023). La IA ha revolucionado múltiples sectores desde su introducción. Debido a esto, los chatbots que hacen uso de LLM, tienen la habilidad de cambiar trascendentalmente, la manera en que se llevan a cabo diversas profesiones (Abid Haleem et al., 2022).

Los principios mencionados en conjunto con otras técnicas de aprendizaje son métodos utilizados para el desarrollo de inteligencia artificial que está siendo aplicada en múltiples ámbitos. Hasta la fecha, se ha aplicado la inteligencia artificial en la educación, salud, industria y manufactura, publicidad, transporte, finanzas, servicio de atención al cliente, agricultura y otros. Sin embargo, el potencial de la IA abarca todas las áreas de la vida humana.

3.1.1.2 Impacto y tendencias en el ámbito laboral

Se encontró que la presencia de la inteligencia artificial se hace más evidente cada día. Como consecuencia de las constantes innovaciones tecnológicas, se producen cambios en los entornos laborales que compelen a los colaboradores a adaptarse a estas transformaciones. Se descubrió que es común observar la coexistencia de máquinas y personas en un mismo entorno laboral. Incluso, se visualiza como algunos empleados han sido reemplazados por equipos especializados debido a que realizaban actividades mecánicas y rutinarias, las cuales podrían ser automatizadas.

Corvalán (2019) describió que los recursos humanos no deben temer ser sustituidos por máquinas inteligentes si se encuentran lo suficientemente preparados con habilidades necesarias para satisfacer el evolutivo campo laboral. Es por esto por lo que una de las

tendencias internacionales durante el proceso de transición de automatización e implementación de la IA, radica en la inversión que las diferentes entidades deben realizar para capacitar y reentrenar a su personal; con el objetivo de aprovechar estas nuevas habilidades y desarrollar instituciones más competitivas.

3.1.1.3 *Cobotización*

El término emergente *cobotización*, sugiere el cotrabajo entre sistemas de inteligencia artificial y colaboradores humanos, para crear un personal laboral con inteligencia aumentada. Se cree que la inteligencia aumentada está basada en una cobotización inclusiva dirigida hacia unir esfuerzos para crear cuatro herramientas que optimicen el trabajo. Estas son: la asistencia inteligente, automatización, diagnóstico o detección inteligente y predicción (Corvalán, 2019).

3.1.1.4 *Ética*

Se discutieron los principios y dilemas éticos que giran en torno al uso de la inteligencia artificial.

3.1.1.4.1 *Principios éticos*

La Comisión Europea (2019) creó directrices para la creación de IA digna de confianza, dentro de esto se encuentran cuatro principios éticos: respeto de la autonomía humana, prevención del daño, equidad y explicabilidad. Tomando en cuenta estos principios, se seleccionaron siete requisitos clave para la creación de IA confiable, estos fueron:

- 1) Supervisión humana
- 2) Solidez técnica y seguridad
- 3) Privacidad y gestión de datos
- 4) Transparencia
- 5) Diversidad, no discriminación y equidad
- 6) Bienestar social y medioambiental
- 7) Responsabilidad.

El propósito de crear estos lineamientos fue construir un enfoque basado en principios consensuados para que así, los desarrolladores de IA tuvieran una lista de métodos de evaluación y recomendaciones políticas. Asimismo, estos requisitos, se fundaron en la ética biomédica o bioética. La bioética posee cuatro principios base propuestos por Beauchamp and Childress: autonomía, justicia, beneficencia y no maleficencia (Stahl, 2021).

3.1.1.4.2 *Beneficios y problemas éticos*

La IA además de brindar habilidades técnicas para facilitar tareas en diversos ámbitos, aporta beneficios éticos. La Comisión Europea (2019) describió que la IA no solo cumple con su objetivo final, sino que además se convierte en un valioso recurso para fomentar el desarrollo humano y de esta manera mejorar el bienestar individual y social. Adicionalmente, aporta progreso e innovación a la sociedad.

Por el otro lado, se encontró que la presencia de la IA viene acompañado de problemáticas o dilemas éticos (Ilustración 2). Esto se clasificó en tres grandes grupos (Anexo 1):

- Problemas éticos que surgen del ML

La mayoría de los éxitos actuales de la inteligencia artificial se lograron a través de técnicas de aprendizaje automático, las cuales están basadas en redes neuronales artificiales. Sin embargo, algunas características de estos enfoques generan preocupaciones éticas. La opacidad y la imprevisibilidad de estos sistemas, junto con su dependencia de grandes conjuntos de datos para su entrenamiento, son aspectos que plantean inquietudes. Debido a que ni el desarrollador, ni el implantador y tampoco el usuario del sistema, pueden predecir cómo este reaccionará ante un conjunto específico de entradas. Además, debido a que el sistema es adaptativo y dinámico al aprender, sus comportamientos pasados no garantizan un comportamiento futuro idéntico en situaciones similares.

- Problemas generales relacionados a vivir en un mundo digital

Como se ha discutido anteriormente, la inteligencia artificial tendrá un fuerte impacto en la vida de los seres humanos. Un ejemplo de esto es la alta probabilidad que la IA tenga un impacto significativo en la economía, afectando tanto el empleo como la concentración del poder económico. Asimismo, es plausible que tenga repercusiones en la justicia, la equidad y el acceso a los servicios. Además, existe el riesgo de que genere desequilibrios de poder que pongan en peligro los principios democráticos. De hecho, el uso de la IA en diversos campos de aplicación, como las armas autónomas, plantea una serie de cuestiones éticas que deben abordarse.

- Problemas metafísicos

Este tipo de problemática se vincula con la AGI, la naturaleza de la conciencia y la dinámica de la interacción entre humanos y máquinas. La idea principal se basa en que, en un futuro, la AGI podría demostrar habilidades de razonamiento humano, por lo cual serían capaces de ayudar a mejorar la raza humana o destruirla. Aunque esto aún no sea una realidad, es esencial

tenerlas en cuenta durante el progreso continuo de la IA. A partir de esto surgen interrogantes como la posibilidad de atribuir responsabilidad moral a la IA, o programarla con una moral artificial para tomar decisiones éticas.

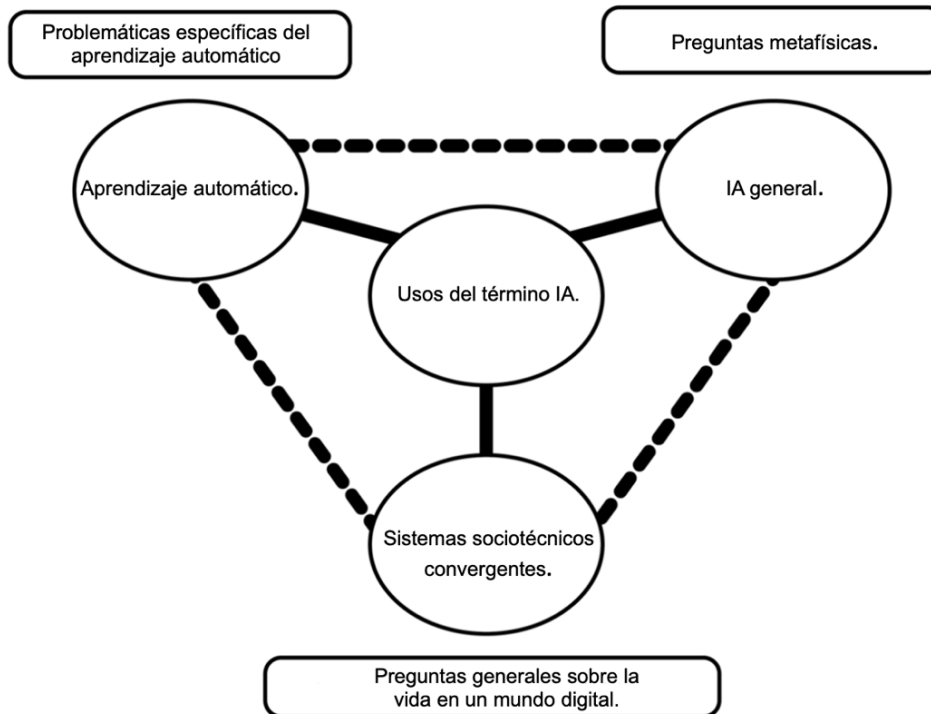


Ilustración 2 – Conceptos de la IA y preguntas que plantean.

Fuente: (Stahl, 2021)

Claro está que los problemas presentados son complejos y polifacéticos, por lo cual no existe una respuesta concreta a estas interrogativas. Sin embargo, es imprescindible comenzar la conversación para poder preparar a la humanidad sobre los posibles retos que se avecinan. Pues a medida que la IA se integra más en nuestras vidas, surgen más interrogantes. Es imperativo reflexionar sobre estas cuestiones y garantizar que la IA se desarrolle y se utilice de manera ética y responsable.

En resumen, para comprender la ética es fundamental entender la situación humana. Debido a que la ética se desarrolla y se aplica porque somos seres capaces de empatizar con otros que comparten temores y esperanzas similares. Por otro lado, la IA no comparte la misma categoría que los seres humanos, ya que carece de la misma capacidad empática y comprensión.

3.1.2 MICROENTORNO

En el microentorno se expone un análisis de los elementos relacionados directamente con el rol de la inteligencia artificial en el ámbito sanitario, su relación con la ingeniería en biomédica y la aplicación en la región latinoamericana.

3.1.2.1 *Inteligencia artificial aplicada a la salud*

La tecnología ha ido evolucionando a lo largo de la historia humana, impulsando el desarrollo en cada una de sus etapas. En los últimos años la IA se ha popularizado y ha sido aplicada a diversos ámbitos, la salud no ha sido la excepción. Se han evidenciado diversos avances que contribuyen a la optimización de los procesos sanitarios, lo que ha influido en el mejoramiento de la calidad de vida de las personas. Martínez-García et al. (2019) mencionaron que entre las aplicaciones más populares de la IA se encuentran aquellas relacionadas a la salud debido a su utilidad. A continuación, se presentan ejemplos de la aplicación de la IA en los procesos sanitarios.

3.1.2.1.1 *IA en la sala de operaciones*

Las intervenciones quirúrgicas representan un método de tratamiento para variadas afecciones, sin embargo, su éxito depende de diversos factores. Entre los cuales Wall & Krummel (2020) describieron que la experiencia del cirujano, la preparación previa a la operación y la toma de decisiones en tiempo real son aspectos importantes a considerar. Teniendo como objetivo primordial brindar un tratamiento óptimo, se considera que se requiere de un "modelo integral" que vaya más allá del juicio del cirujano; aquí es donde entran los modelos predictivos que poseen la capacidad para identificar factores no considerados.

Se han aplicado diversos modelos predictivos en el ámbito de la cirugía. En Maubert et al. (2019) haciendo uso de un modelo basado en aprendizaje profundo lograron predecir la resecabilidad de tumores causados por carcinomatosis peritoneal con una precisión elevada de 97.82%. Por otro lado, haciendo uso de redes neuronales, De Ramón Fernández et al. (2019) lograron determinar el tratamiento quirúrgico en un 96% de embarazos ectópicos. Cabe destacar que el uso de la IA se extiende a más allá de procesos quirúrgicos, Lundberg et al. (2018) lograron el desarrollo de un sistema de deep learning capaz de predecir la hipoxemia intraoperatoria inclusive minutos antes de que la misma suceda.

3.1.2.1.2 *IA en el cuidado de los pacientes*

El personal de enfermería cumple una labor fundamental en el tratamiento y cuidado de la salud de los pacientes, desempeñan funciones de brindar seguridad y educación a los

mismos. En este ámbito también ha incursionado la IA con la “enfermería virtual”. Esta consiste en un asistente virtual impulsado por IA ofrece experiencias personalizadas a los pacientes, ayudándoles a identificar su enfermedad en función de los síntomas, controlar su estado de salud, programar citas con el médico y entre otras.

De igual forma, Martínez-García et al. (2019) mencionan otra forma de enfermería virtual, a través de una aplicación para dispositivos móviles, que permite realizar un seguimiento diario de los valores y mediciones que el paciente puede tomar por sí mismo desde su hogar, todo ello con la asistencia de un profesional virtual.

3.1.2.1.3 IA en el diagnóstico de enfermedades

Cada vez se vuelve más popular la aplicación de redes neuronales y otros tipos de IA para el diagnóstico de enfermedades. El principal objetivo de la IA en el diagnóstico es reducir el tiempo de evaluación del caso y la obtención de un diagnóstico certero que permita un tratamiento adecuado y oportuno. Se han abordado diversas afecciones desde las que se consideran poco complejas, hasta las de alta complejidad. A continuación, se presentan ejemplos precisos.

Kavya et al. (2021) propusieron un marco asistido por computadora para el diagnóstico de alergias. Los autores aplicaron algoritmos de ML de servidor y luego seleccionaron el algoritmo de mejor rendimiento utilizando la validación cruzada de k-fold. Además, los autores implementaron el marco asistido por computadora en la aplicación móvil, que puede ayudar a los médicos junior a confirmar las predicciones de diagnóstico.

Peng et al (2021) propusieron un marco XAI que puede ayudar a los médicos con el pronóstico de los pacientes con hepatitis. Los autores compararon los métodos XAI intrínsecos como la regresión logística (LR), el árbol de decisión (DT) y kNN con los modelos complejos.

H. Wu et al. (2021) propusieron un gráfico multigranularidad contrafactual que apoya la extracción de hechos (CMCE) para el diagnóstico de linfedema. CMGE es una red neuronal basada en gráficos que puede extraer datos de registros médicos electrónicos. Además, podría obtener una relación causal entre características.

Sabol et al. (2020) propusieron una aplicación XAI denominada criterio acumulativo de pertenencia de clase difusa (CFCMC) que podría ayudar a los patólogos y usarse para el diagnóstico del cáncer colorrectal. El sistema propuesto proporciona un diagnóstico a través de la explicación semántica de los resultados de la clasificación del tipo de tejido.

3.1.2.2 Utilización en el área biomédica

Debido a que se caracteriza por ser polifacética, la ingeniería biomédica puede ser definida de diversas formas, sin embargo, en su mayoría esas definiciones pueden concordar en que el biomédico busca innovar a través de la tecnología. La IA en el ámbito biomédico se encuentra presente en el diagnóstico por imágenes, desarrollo de dispositivos, precisión de diagnósticos de equipos de laboratorio, aplicaciones terapéuticas, entre otros. Cabe destacar que luego de realizar una revisión bibliográfica Acosta et al. determinaron que la IA en el ámbito biomédico posee diversas oportunidades de desarrollo (Ilustración 3).

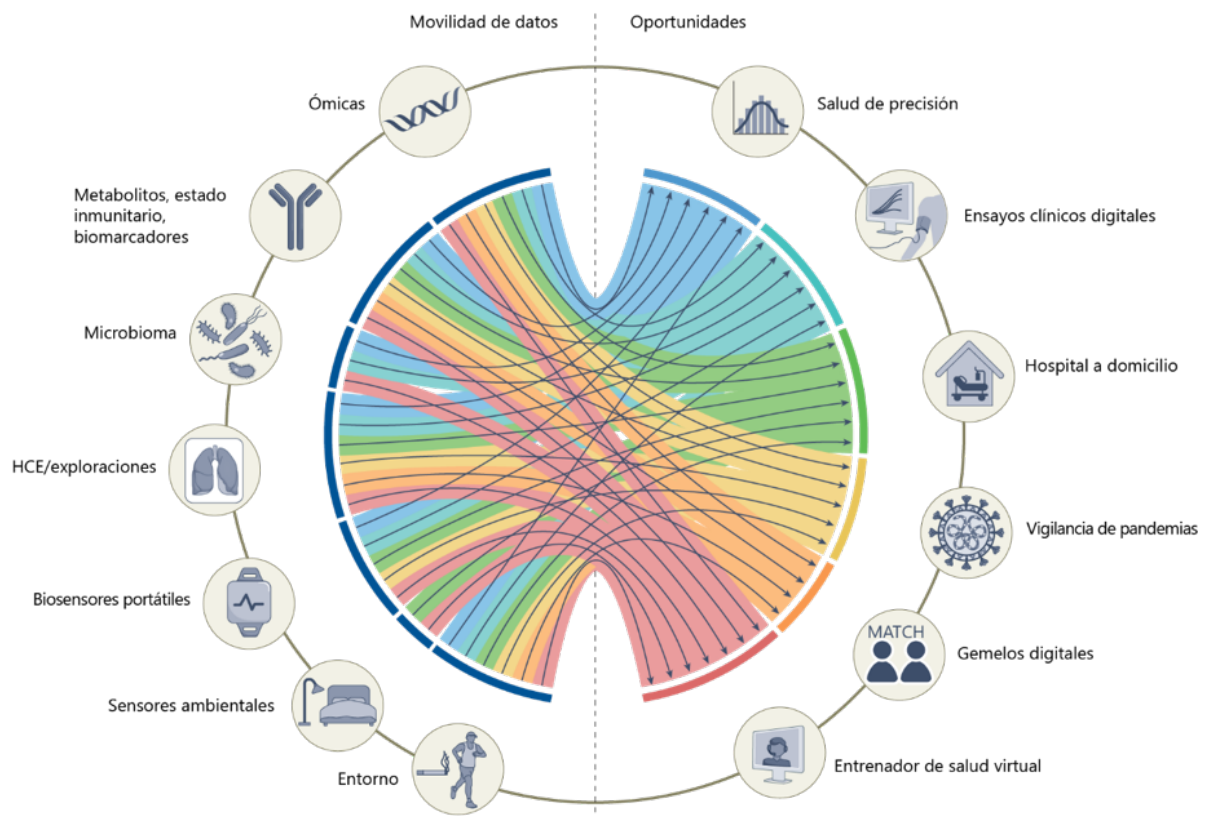


Ilustración 3 - Modalidades de datos y oportunidades para la IA biomédica multimodal

Fuente: (Acosta et al. 2022)

Haciendo un énfasis en las oportunidades, es evidente que las posibilidades de implementación de IA para la mejora de la atención sanitaria son variadas.

3.1.2.2.1 Precisión en la salud

Con el progreso en la secuenciación durante las últimas dos décadas, se experimenta una revolución en la cantidad de datos biológicos detallados que se pueden obtener utilizando nuevos desarrollos técnicos. "Estos se denominan colectivamente "omas" e incluyen el genoma, el proteoma, el transcriptoma, el inmunoma, el epigenoma, el metaboloma y el

microbioma" (Karczewski & Snyder, 2018). Estos análisis son posibles gracias a herramientas como Redes neuronales gráficas y modelos PHATE, que pueden aprender representaciones abstractas de datos biológicos y clínicos en diferentes niveles de granularidad, para determinar un diagnóstico. Estas nuevas técnicas son de suma utilidad ya que permiten el análisis a nivel celular, permitiendo la detección de patologías como el cáncer de forma precisa.

3.1.2.2.2 *Ensayos clínicos digitales*

Los ensayos clínicos poseen diversas limitantes, entre ellas altos costos, limitantes geográficas y socioculturales, además de tomar un periodo de tiempo sumamente largo. Estas barreras se pueden eliminar mediante la digitalización de ensayos clínicos, al reducir las barreras para la inscripción y retención de participantes, promover el compromiso y optimizar las mediciones e intervenciones de los ensayos. Steinhubl et al. (2019) menciona que el uso de tecnologías digitales puede mejorar la granularidad de la información obtenida de los participantes en un ensayo clínico, aumentando así el valor de estos estudios.

3.1.2.2.3 *Monitoreo remoto: el "hospital en casa"*

Noah et al. (2018) destacaron que, gracias a los avances recientes en biosensores, monitoreo continuo y análisis, se han abierto posibilidades de crear un entorno similar al hospitalario, desde el hogar del paciente. Esto supone una promesa de reducción significativa de costos y una menor necesidad de personal de atención médica. Adicionalmente, se evitan de infecciones nosocomiales y errores médicos que suelen ocurrir en instalaciones médicas. A parte de todo eso, brinda mayor comodidad, conveniencia y apoyo emocional para el paciente, al estar rodeado de miembros de su familia.

El desarrollo de biosensores y otros sistemas de telesalud han sido impulsados gracias a los avances de la IA en las últimas décadas, por ello esta es un área de la biomédica a la que se le puede apostar.

3.1.2.2.4 *Vigilancia de pandemias y detección de brotes*

La pandemia del COVID-19 evidenció la importancia de la vigilancia epidemiológica. Algunos países, lograron integrar con éxito datos multimodales de mapas de migración, utilización de teléfonos móviles y datos de prestación de servicios de salud para pronosticar la propagación del brote e identificar casos potenciales. Ejemplos de esto son: J. T. Wu et al., (2020) donde en su artículo "Pronóstico inmediato y pronóstico de la posible propagación nacional e internacional del brote de nCoV 2019 que se originó en Wuhan, China: un estudio de modelado", detallan la aplicación de estos servicios en China.

3.1.2.2.5 *“Gemelos digitales”*

Acosta et al. (2022) explicó que, en la actualidad, se confía en los ensayos clínicos como la mejor evidencia para identificar intervenciones exitosas. Sin embargo, estas intervenciones pueden beneficiar únicamente a una parte de individuos. Un enfoque complementario llamado "gemelos digitales" puede suplir estas lagunas de conocimiento al utilizar grandes cantidades de datos para modelar y predecir con alta precisión cómo una intervención terapéutica específica beneficiaría o afectaría a un paciente particular. De esta manera, se busca personalizar y mejorar la eficacia de las intervenciones médicas para cada paciente en base a su perfil y características personales.

El término hace referencia a la búsqueda de crear un "gemelo" digitalmente al que se le puedan aplicar pruebas siempre digitalmente para determinar las intervenciones adecuadas a realizarle.

3.1.2.2.6 *Asistente virtual digital*

Este es otro término para enfermería digital, aunque ya son bastante populares, se busca incursionar en hacerlos más especializados para brindar una atención personalizada al paciente. Esta atención personalizada siendo en base a datos específicos como la secuenciación del genoma, el monitoreo continuo de biomarcadores y otros datos biomédicos relevantes, para responder preguntas relacionadas con la salud, clasificar síntomas o comunicarse con proveedores de atención médica cuando corresponda.

Acosta et al. (2022) menciona ciertas oportunidades para incursionar, sin embargo, considerando la multidisciplinaria que es la biomédica, existen aún más oportunidades. Una de esta es la gestión de las tecnologías médicas en las diversas etapas de su ciclo de vida en el ambiente hospitalario, es decir IA aplicada a la ingeniería clínica o a un departamento de biomédica. Los beneficios que esta brindaría serían infinitos; facilitarían procesos, toma de decisiones, gestión de mantenimientos correctivos e incluso presenta la posibilidad de implementar mantenimiento predictivo. Además de representar grandes beneficios económicos en la reducción de costos de mantenimiento, entre otros.

3.1.2.3 *Actualidad de la IA en salud en Latinoamérica*

Aunque ha sido de forma lenta, se ha logrado la introducción de la IA en la atención en salud latinoamericana. Su implementación se ha realizado fundamentalmente en el diagnóstico de enfermedades, haciendo uso de diversas técnicas. Cabe destacar, que los

sistemas sanitarios en Latinoamérica varían en gran manera de región a región, por lo que establecer un protocolo de IA estándar resulta todo un reto.

Recientemente, debido al riesgo que supone, se ha hecho hincapié en la integración de la IA y oncología de precisión en Latinoamérica. A pesar de los avances médico-tecnológicos, el cáncer sigue representando un desafío en el diagnóstico, tratamiento y pronosis, por ello se busca el desarrollo e implementación de herramientas que faciliten dicho trabajo. Sussman et al. (2022) expresó que gracias al avance de la biología molecular y al desarrollo de métodos de diagnóstico más precisos, así como herramientas terapéuticas mejoradas; la IA tiene el potencial de mejorar diversas situaciones clínicas. Entre ellas, destacan la capacidad de predecir parámetros clínicamente relevantes, realizar diagnósticos más precisos y agilizar el progreso de la medicina personalizada.

Sin embargo, aunque la IA es muy prometedora, su implementación presenta algunos problemas intrínsecos. Estos problemas a considerar son la normalización, la comprensibilidad, la verificación y el valor de los datos, la incorporación de estas tecnologías en la práctica médica y temas relacionados con la regulación y la legislación, entre otros (Chua et al., 2021). De igual forma, es importante reconocer las limitantes económicas que sufren la mayoría de los países latinoamericanos. Werutsky et al. (2021) menciona que algunas disparidades adicionales en la investigación y ejecución prevalecen y son más evidentes en los países de ingresos bajos, debido a los desequilibrios socioeconómicos que pueden afectar el financiamiento de áreas críticas para el desarrollo y la integración de estos modelos de IA.

En el caso de Honduras, aún no se encuentra ningún proyecto de IA aplicada al ámbito sanitario. Porque, aunque la IA como herramienta es sumamente innovadora para los sistemas sanitarios latinoamericanos, se deben tomar en consideración diversos factores relevantes a los mismos para que su implementación sea exitosa. Realizar un análisis previo a la implementación de un sistema de IA es de alta importancia en los países de Latinoamérica, ya que estos suponen un alto costo que muchos sistemas sanitarios, como el nacional, no pueden pagar.

3.2 CONCEPTUALIZACIÓN

La adquisición de conocimientos y comprensión acerca de las IA se han convertido en un imperativo ineludible en la contemporaneidad. La creciente influencia y penetración de estas en diversas esferas de la sociedad demanda un acercamiento fundamentado y sistemático para abordar sus aplicaciones en distintas disciplinas.

En el contexto actual del desarrollo tecnológico, la comprensión de las IA es base para estar al tanto de los avances y aplicaciones que emergen de manera acelerada en una amplia gama de industrias. La comprensión de las capacidades y limitaciones inherentes en las IA permite una ponderación adecuada a sus implementaciones, evitando la adopción irreflexiva y resguardando la ética en la incorporación de tales entidades en distintas facetas.

En el ámbito del procesamiento del lenguaje natural (NLP), la perplejidad constituye una métrica habitual empleada para cuantificar el desempeño de un modelo de lenguaje. En vista de la aparición de los modelos de lenguaje de inteligencia artificial más avanzados, como Megatron de Microsoft y GPT-3 de OpenAI, resulta imperativo comprender cómo evaluar adecuadamente su rendimiento.

De igual manera, es relevante adquirir un conocimiento introductorio acerca de una de las disciplinas que fundamenta la investigación propuesta: la ingeniería clínica. Esta rama posee un alto interés en el respaldo de la atención hospitalaria. La aparición de la necesidad de contar con ingenieros especializados en la preservación y mantenimiento de los dispositivos médicos en entornos hospitalarios se remonta a la última etapa de la década de los setenta. Sus inicios coinciden con el progreso de la tecnología electrónica en los equipos médicos, así como con el desarrollo de computadoras, automatización y telecomunicaciones. Previamente, predominaban dispositivos electromecánicos, y la experiencia y conocimiento de los médicos y su equipo eran las principales fuentes de apoyo en esta área (Saldarriaga, 2008).

3.2.1 CHATGPT

GPT son las siglas de "Generative Pre-trained Transformer", cuyo significado se traduce al español como "Transformador pre-entrenado generativo". Es una herramienta de inteligencia artificial desarrollado por la empresa OpenAI que se basa en el aprendizaje automático y la generación de lenguaje natural para interactuar con los usuarios (OpenAI, 2022).

En la Tabla 1, se presenta una descripción más detallada de los aspectos positivos y negativos relacionados con el uso de este agente conversacional. Esto con el fin de brindar una comprensión más completa de su funcionamiento y sus implicaciones.

Tabla 1 - Comparativa entre las ventajas y desventajas ofrecidas por ChatGPT.

Ventajas Chat GPT	Desventajas Chat GPT
Entiende perfectamente español, puede hacerle preguntas y pedir que se desarrolle textos en este idioma.	Todo lo que se escribe en el ChatGPT queda registrado.
Es capaz de escribir guiones originales.	Suele ser inestable debido a la alta demanda.
Produce escritos de alta calidad.	En diferentes temas el contenido que genera puede ser preciso o correcto.
Comprende instrucciones complejas.	Puede generar información errónea,

Fuente: Elaboración propia.

3.2.2 GOOGLE BARD

Es basada en el modelo de lenguaje laMDA en el que Google lleva trabajando años, este tiene como objetivo ofrecer una herramienta de IA conversacional al usuario. Según Google Bard es capaz de tomar la información de Internet para ofrecer respuestas actuales y de alta calidad, en donde la idea es que el usuario pueda obtener una respuesta completa y general de cualquier concepto, impulsando las búsquedas en Internet a través del lenguaje natural.

Su funcionamiento en el estado actual es ser una herramienta sin límite, obteniendo sus respuestas a partir de diversas fuentes de internet en segundos. De misma manera su ecosistema trata de ser lo más amigable con el usuario, de manera que su lenguaje es casual "como si estuviera hablando con un amigo", se pretende ser multilingüe, comprende la situación de forma que puede descifrar lo que se está hablando y responder de manera que tengo sentido lógico, al ser un producto de Google se relaciona con otros para dar una asistencia más fluida (Google, 2023). Las ventajas y desventajas se exponen en la Tabla 2.

Tabla 2 - Comparativa entre las ventajas y desventajas ofrecidas por Google Bard.

Ventajas	Desventajas
Construido en base al modelo PaLM.	Sesgos, imprecisiones y ambigüedades.
Conversaciones similares a las humanas.	Limitaciones creativas notables.
Capacidades generativas específicas.	No cita ni vincula a as fuentes.
Acceso a tiempo real a sitios web.	Inconsistencias notables en las respuestas.
Extensión a través de complementos.	Capacidades generativas limitadas.

Fuente: Elaboración propia.

3.2.3 CHATSONIC

De la empresa Writesonic, es una herramienta de inteligencia artificial que se basa en GPT-3.5 y GPT-4. Tiene funciones tales como creación de texto e imágenes, comandos de voz y sección de preguntas y respuestas. Sus características más destacadas se encontraron, al ser una herramienta de chat pretende generar texto en base a recopilación de datos, crea imágenes, responde a preguntas, traducción y comando de voz.

Esta IA trabaja bajo con un modelo de negocios que permite características básicas del servicio de manera gratuita y otras características avanzadas, por un costo adicional. Debido a esto tiene limitantes como puede ser la restricción de palabras, siendo de 2500 palabras gratis mensuales. También de manera gratuita, posee una calidad de investigación baja, problemas con las imágenes y falta de respuesta ante las interrogantes presentadas. Esto debido a que las bases de la IA se formaron con información hasta el 2021 y al tener la opción de incluir datos de Google, si no se activan no se encontrará la información requerida (Writesonic, 2023). La Tabla 3, expone una comparativa entre las ventajas y desventajas de uso ofrecidas por esta tecnología.

Tabla 3 - Comparativa entre las ventajas y desventajas ofrecidas por Chatsonic.

Ventajas	Desventajas
Fácil de usar.	Errores de hecho.
Asequible.	Se requiere mucha intervención humana.
Creativa.	Límite de palabras.
Prueba gratuita.	Repetición.

Fuente: Elaboración propia.

3.2.4 PERPLEXITY

Perplexity AI es una innovadora plataforma de búsqueda impulsada por inteligencia artificial que se propone brindar respuestas precisas a interrogantes complejas mediante la utilización de modelos lingüísticos de gran envergadura. La fundación de esta empresa tuvo lugar en el año 2022, siendo los cofundadores Aravind Srinivas, Denis Yarats y Andy Konwinski. Srinivas previamente había destacado en el campo de los modelos generativos de lenguaje y la difusión en OpenAI, mientras que Yarats había ejercido como científico de investigación en IA en Facebook AI Research, y Konwinski ostenta el cargo de cofundador en DataBricks (Perplexity AI, 2023).

Sus características se basan en responder preguntas complejas que los motores de búsqueda tradicionales no pueden en el cual se le proporcionara al usuario información

precisa y fiable, posee un ecosistema amigable. Se presenta una comparación de beneficios y desventajas de su uso en la Tabla 4.

Tabla 4 - Comparativa entre las ventajas y desventajas ofrecidas por Perplexity.

Ventajas	Desventajas
Ofrece respuestas rápidas y completas a preguntas consideradas complejas.	Su ecosistema se encuentra en inglés, pero entiende y responde en otros idiomas.
Ayuda al usuario a aprender cosas nuevas y explorar perspectivas diferentes.	No es capaz de manejar las preguntas subjetivas o ambiguas.
Mejora el pensamiento crítico y la capacidad de investigación, mostrando al usuario las fuentes de la información	Se basa en modelos matemáticos, por ende, no es capaz de captar los matices y sutilezas del lenguaje.

Fuente: Elaboración propia.

3.2.5 INGENIERÍA CLÍNICA

La ingeniería clínica, denominada como tal, constituye una subdisciplina de la ingeniería biomédica orientada hacia la gestión, implementación y progreso de los instrumentos o aparatos que establecen interacciones directas con el paciente, ya sea con propósitos terapéuticos o diagnósticos. Este campo, de amplio alcance en la ingeniería, abarca de manera integral todas sus ramas y especializaciones, destacando especialmente la electrónica, las comunicaciones, los sistemas, la mecánica, la electricidad y los materiales, entre otras áreas de importancia (Saldarriaga, 2008).

Con el objetivo de esclarecer de manera más precisa la función actual del ingeniero clínico en el entorno hospitalario, se presentaron a continuación las responsabilidades particulares que han sido definidas por varias instituciones y universidades de renombre internacional, haciendo especial énfasis en el American College of Clinical Engineering (ACCE). Según Ibáñez & Soto (2006) las actividades que más destacan dentro de este ámbito se dividen en: administración, adquisición, desarrollo de instalaciones físicas especiales, seguridad tecnológica y capacitación.

Administración. Se trata de un procedimiento orientado a garantizar la disposición adecuada de la tecnología biomédica en el entorno hospitalario, con el propósito de proporcionar los niveles necesarios de seguridad y calidad para satisfacer las demandas tecnológicas que requieren los servicios de atención médica. Esto implica llevar a cabo diversas actividades, tales como:

- La elaboración de inventarios para supervisar el control tecnológico.
- La gestión y adquisición de componentes y suministros tecnológicos.

- La administración de aspectos relacionados con el mantenimiento, la gestión de riesgos y la calidad de la tecnología.
- La evaluación de la tecnología biomédica.
- La evaluación y supervisión de la contratación de servicios externos.
- La creación de recursos técnicos especializados, como bibliotecas técnicas.

Adquisición. Este procedimiento tiene como objetivo la integración de tecnología biomédica en el entorno hospitalario, en consonancia con las necesidades específicas. Entre las actividades que engloba se incluyen:

- Planificar la adquisición de tecnología médica.
- Llevar a cabo una evaluación técnica, normativa y de costo-beneficio de la tecnología biomédica antes de su adquisición.
- Coordinar la instalación y calibración de la tecnología biomédica recién adquirida.
- Realizar una evaluación de la tecnología después de su implementación.
- Diseñar o modificar tecnología biomédica e instrumental médico.

Desarrollo de instalaciones físicas especiales. Se trata del procedimiento destinado a preparar las infraestructuras físicas de un hospital de manera que la tecnología biomédica alojada funcione de manera eficiente. Su función consiste en:

- Brindar asesoramiento en la planificación, diseño y evaluación de proyectos destinados a la construcción o renovación de instalaciones clínicas particulares.

Seguridad tecnológica. Este proceso tiene como finalidad garantizar un funcionamiento adecuado de la tecnología biomédica, reduciendo al mínimo el riesgo de accidentes o daños que puedan poner en peligro la salud de los pacientes, el personal hospitalario y la propia tecnología. Sus actividades principales incluyen:

- Coordinar la seguridad en aspectos eléctricos.
- Coordinar la seguridad en radiología.
- Coordinar la seguridad en cuestiones electromagnéticas.
- Coordinar la seguridad en sistemas hidráulicos.
- Coordinar la bioseguridad y la gestión de residuos hospitalarios.
- Supervisar e investigar incidentes relacionados con la tecnología.

Capacitación. Este proceso tiene como objetivo garantizar que el personal médico y paramédico del hospital cuente con el conocimiento necesario para operar y respaldar eficazmente la tecnología biomédica. Sus actividades típicas comprenden:

- Proporcionar capacitación en el uso de la tecnología a profesionales médicos y de atención médica.
- Formular políticas para el entrenamiento y la formación en tecnología dirigidos al personal médico y paramédico.

3.3 TEORÍAS DE SUSTENTO

Para comprender las teorías fundamentales de los IA, se expondrán las diferentes bases para estudiar y comprender las IAs, el deep learning y el uso de las TICs.

3.3.1 BASES TEÓRICAS

3.3.1.1 Programación

Programación en informática, la expresión hace referencia al proceso de concepción y creación de programas o aplicaciones mediante el desarrollo de un código fuente, el cual consiste en un conjunto de instrucciones que el ordenador sigue para llevar a cabo la ejecución de un programa (Trejos, 2022). La programación es el primer paso para crear un software, que luego lo ejecutará el hardware de la computadora para obtener así los resultados esperados del código realizado.

La programación especifica la estructura y el comportamiento de un programa rectificando si este está funcionando adecuadamente bajo el propósito planteado. También la programación incluye la especificación del algoritmo definido como la secuencia de pasos y operaciones que el programa debe realizar para poder resolver un problema; para que dicho algoritmo se desempeñe correctamente debe estar ejecutado en un lenguaje compatible además de correcto.

3.3.1.2 Aprendizaje automático (*machine learning*)

Se define como un tipo de inteligencia artificial en el que las computadoras utilizan grandes cantidades de datos para aprender a realizar tareas en lugar de ser programadas para realizarlas, lo que permite que el aprendizaje automatizado haga posible que los sistemas informáticos se vuelvan más inteligentes a medida que se encuentran datos adicionales (Zhou, 2021).

En la recopilación realizada por Janiesch et al. (2021) se expuso su objetivo primordial como la automatización de la construcción de modelos analíticos para facilitar tareas cognitivas, tales como la detección de objetos o la traducción en lenguaje natural. Esta meta se logra alcanzar mediante el uso de algoritmos que, de manera repetitiva, aprenden de los datos de específicos del problema a investigar, permitiendo a las computadoras descubrir información oculta y patrones más complejos.

3.3.1.3 Redes neuronales profundas

Se trata de una metodología de aprendizaje automático que habilita a una computadora, a través de su entrenamiento, para ejecutar tareas que presentarían una dificultad al emplear enfoques de programación tradicionales. Los algoritmos de redes neuronales están inspirados en la estructura y funciones del cerebro humano: que al igual que la mente, están diseñados para operar no solo siguiendo un conjunto predeterminado de reglas, sino también para predecir posibles soluciones y deducir conclusiones basadas en iteraciones y experiencias previas.

Estas redes neuronales se componen por varias capas de nodos que reciben entradas de otras capas y producen una salida hasta llegar al resultado final. Las redes neuronales pueden componerse de cualquier número de capas ocultas, por ello cuantas más capas de nodos tenga la red, mayor será la complejidad de esta (Ilustración 4).

La capacidad de adaptación al aprendizaje es una de las características más atractivas de las redes neuronales. Estas redes adquieren habilidades para realizar ciertas tareas mediante un proceso de entrenamiento basado en ejemplos ilustrativos. Dado que las redes pueden identificar patrones a través de ejemplos y entrenamiento, más no es necesario construir modelos previos ni especificar funciones de distribución de probabilidad.

Las redes neuronales se logran comportar como sistemas dinámicos auto adaptativos, esta capacidad proviene de la autorregulación de sus elementos procesales que conforman dicho sistema (Sarmiento-Ramos, 2020).

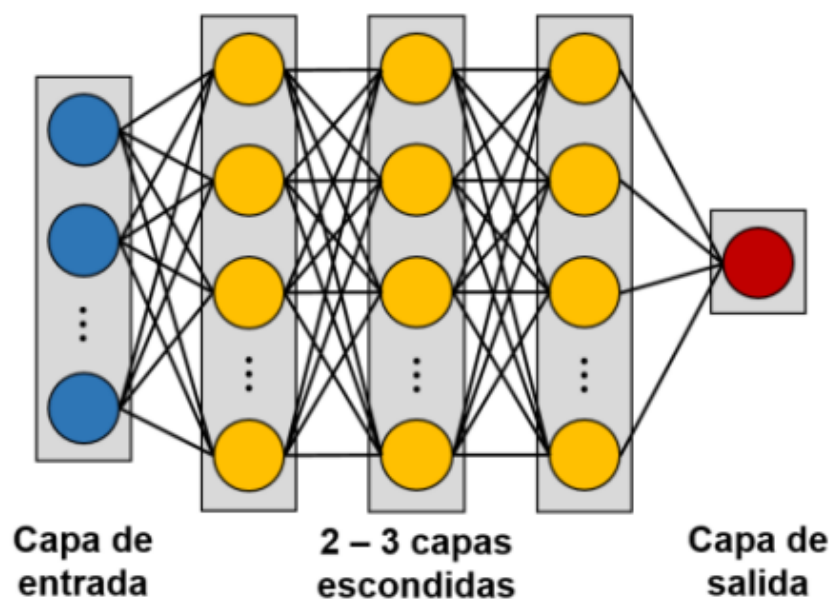


Ilustración 4 – Ejemplo de una Red Neuronal.

Fuente: (Sarmiento-Ramos, 2020).

3.3.1.4 *Aprendizaje profundo (deep learning)*

De acuerdo a Kelleher (2019), el aprendizaje profundo es una subdivisión de la inteligencia artificial que se enfoca en crear grandes modelos de redes neuronales, capaces de tomar decisiones basadas en datos. Este trata de imitar la forma en que los humanos obtienen ciertos tipos de conocimiento.

De acuerdo a Kelleher (2019), el aprendizaje profundo es una subdivisión de la inteligencia artificial que se enfoca en crear grandes modelos de redes neuronales, capaces de tomar decisiones basadas en datos. Este trata de imitar la forma en que los humanos obtienen ciertos tipos de conocimiento.

El aprendizaje profundo constituye un componente necesario en el ámbito de la ciencia de datos, el cual engloba el uso de estadísticas y modelos productivos. Esto es especialmente ventajoso para los científicos de datos cuya responsabilidad recae en la recolección, análisis e interpretación de volúmenes masivos de información. Mediante el aprendizaje profundo, todo este proceso se agiliza y simplifica de manera significativa.

Este tipo de IA es aplicado en muchas áreas, pero en específico; en el ámbito sanitario, el deep learning es utilizado para el procesamiento de imágenes médicas en rayos X, resonancias magnéticas y tomografías computarizadas. Así como también es empleado en el diagnóstico de enfermedades.

3.3.1.5 *Modelo grande de lenguaje*

Según Kojima et al. (2022) un LLM es un tipo del aprendizaje profundo capaz de identificar, resumir, traducir, anticipar y producir texto y otros tipos de contenido utilizando el conocimiento adquirido a partir de grandes conjuntos de datos. Recientemente, la utilización de conjuntos de datos más extensos, han permitido que los modelos grandes de lenguaje preentrenados sean altamente competentes en diversas tareas posteriores de procesamiento del lenguaje natural. Algunas de las tareas de procesamiento de lenguaje natural incluyen la generación y clasificación de texto, respuesta a preguntas en formato conversacional y traducción de textos.

Los modelos grandes de lenguaje generalmente adoptan una arquitectura fundamentada en transformadores. Esta clase de arquitectura de IA emplea mecanismos de autoatención para calcular una suma ponderada sobre una secuencia de entrada, permitiendo determinar de manera dinámica qué tokens en la secuencia tienen una mayor relevancia entre ellos (Jiang et al., 2020).

Un LLM emplea redes neuronales profundas para producir resultados fundamentados en patrones aprendidos de los datos de entrenamiento. Generalmente, las arquitecturas de transformadores permiten que un modelo de aprendizaje automatizado identifique las relaciones entre las palabras en una oración, sin importar su posición en la secuencia de texto, gracias al uso de mecanismos de autoatención (Jiang et al., 2020).

3.3.1.6 *Procesamiento de lenguaje natural*

Conocido también por sus siglas PLN, es una disciplina de la inteligencia artificial que ha adquirido relevancia en el ámbito digital. Gracias a estas técnicas, en la actualidad es posible llevar a cabo traducciones automáticas de texto, revisar la ortografía y realizar conteos de palabras de manera efectiva.

Gracias a la recopilación de Zeroual & Lakhouaja (2018) donde llego a definir que este proceso posee como objetivo "aprender, comprender, reconocer y producir contenido de lenguaje humano". Lo cual en sí mismo trata de mejorar una comunicación eficiente entre el humano y la maquina por medio del uso de lenguas naturales.

Así mismo dentro de la diversidad de aplicaciones del PLN, se destaca en especial el desarrollo de sistemas conversacionales, también conocidos como "chatbots". Los cuales pretenden obtener como objetivo lograr una comunicación más cercana y natural con el usuario, del cual se brinca un canal de comunicación interactivo, así como eficiente Zeroual & Lakhouaja (2018).

Rodríguez-Blanco et al. (2016) en su texto "Extracción de datos enlazados desde información no estructurada aplicando técnicas PLN y ontologías" realizaron una mención sobre las técnicas PLN, del cual se destacaron 3 fases:

- Procesamiento de datos: fase en el cual el contenido de los archivos de entrada es categorizado mediante el proceso de extracción de texto plano, haciendo uso de bibliotecas especializadas para dicho fin.
- Análisis sintáctico superficial: en este enfoque que tiene como objetivo inicial la identificación de conceptos, lo cual consiste en agrupar los tokens de la oración en chunks que representan estructuras gramaticales.
- Análisis de dependencia: de este enfoque, se establecen las relaciones de dependencias concurrentes entre las estructuras gramáticas, que son representadas en un árbol de dependencias.

3.3.1.7 *Ingeniería rápida*

Ingeniería rápida o prompt engineering, según Meyer et al. (2023), es la habilidad para seleccionar la indicación de texto precisa y lograr obtener de manera efectiva, los resultados deseados.

3.3.1.8 *Aprendizaje por refuerzo (RLHF)*

Liu (2023) conceptualizó que el RLHF, es una técnica de ML utilizada para producir agentes que aprenden a trabajar para lograr algún objetivo, facilitado por las interacciones con el entorno y la retroalimentación dada en términos de una señal de recompensa. Sirve de puente importante al incorporar la retroalimentación humana en el proceso de aprendizaje, permitiendo a los servidores aprender mediante la experiencia humana.

Cabe destacar las principales formas en que RLHF aporta positivamente a la experiencia humana con las IA, siendo

- Lucha contra la desinformación: siendo como técnica de alineación eficaz, tiene una gran potencia, para ayudar a mitigar la generación de contenido nocivos.
- Reforzar la alineación de valores: como objetivo central de la investigación en IA es producir sistemas que comporten de forma coherente con los valores y las intenciones humanas.
- Mitigación de sesgos: puede reducir el sesgo en varios niveles de la cadena de producción de la IA. En términos generales, la IA se ve afectada por sesgo de representación que afecta a los estudios de muestreo y población.
- Mejorar el acceso equitativo y la privacidad en la IA: puede abrir la puerta a la democratización de las tecnologías de IA a todos los niveles de la sociedad, independientemente de su nivel de desarrollo.
- Tendiendo puentes entre culturas: tiene el potencial de transformar la forma en que conciliamos las perspectivas interculturales y acordamos el dialogo.
- Impulsar la industria: pueden llegar a facilitar el desarrollo de sistemas de IA más estables.
- Transformación del trabajo: influirán en el grado en que los distintos trabajos son susceptibles de automatización.

3.3.2 METODOLOGÍAS DESARROLLADAS POR OTROS INVESTIGADORES

Los chatbots tomaron popularidad en los últimos meses, junto con esto vinieron interrogantes sobre su utilidad y aplicabilidad. Con el objetivo de otorgar respuestas a estas preguntas, se han llevado a cabo diversos estudios con metodologías distintas (Tabla 6).

Langevin et al. (2021) en su estudio titulado "Evaluación heurística de agentes conversacionales", desarrollaron un comprensivo conjunto de heurísticas para analizar las interacciones con los agentes conversacionales o chatbots. El estudio mencionado utiliza el enfoque heurístico de Nielsen, publicado en 1990, como fundamento base. Los autores de este estudio crearon un sistema de diseño de cuatro fases (Tabla 5). Se evaluaron dos interfaces, uno con asistente de voz y otro en línea. Se comparó la heurística de Langevin et al contra la de Nielsen. Esto sirvió como punto clave para determinar que la mejor fue la desarrollada por Langevin y sus coautores.

Tabla 5 - El proceso de diseño de cuatro fases.

Fase 1	Generación heurística
Fase 2	Revisión de expertos
Fase 3	Validación a través de la evaluación heurística
Fase 4	Validación de las heurísticas revisadas

Fuente: (Langevin et al. 2021)

Ye et al. (2023) realizaron un experimento enfocado en estudiar la mejora de la confianza entre humanos y robots. Este consistió en que los participantes realizaran una tarea de montaje. El objetivo era ensamblar una placa en una pieza de trabajo y fijarla utilizando un taladro. Los participantes realizaron la tarea por sí solos y también hicieron la misma actividad, pero en esta ocasión, con ayuda de un robot. Se usó RoboGPT, habilitado por ChatGPT para su funcionamiento. Al concluir cada paso, debían completar una escala de confianza y un cuestionario. Finalmente, se les hizo una entrevista a los participantes para conocer más sobre la colaboración entre humanos y robots. Los resultados implicaron un impacto positivo de la utilización de LLM emergentes para mejorar la confianza en la cobotización.

Rivas et al. (2018) con el objetivo de conocer como las personas juzgan a los bots basado en aprendizaje automático, encuestaron a una muestra de población estadounidense. Esto se realizó presentando un escenario en que una persona y un chatbot interactuaban con ellas de tal forma que condujera a una situación incómoda en la que el ser humano o el bot les faltaba el respeto. El propósito de la encuesta era analizar la percepción de responsabilidad moral, culpa y confianza de los participantes acerca de los chatbots. Los resultados obtenidos

sugirieron que las personas juzgan a los chatbots con los mismos estándares de moralidad que a los humanos, aunque con una menor escala de culpabilidad.

Tabla 6 - Resumen de las metodologías utilizadas por otros investigadores.

Autores del estudio	Agente conversacional utilizado	Modalidad de entrada	Instrumentos utilizados para la evaluación	Muestra y población	
Langevin et al. (2021)	Amazon chatbot revelado.	Alexa, no	Voz y texto.	Entrevistas, encuestas, heurística de Nielsen y heurística mejorada.	Nueve expertos en heurística y 8 profesionales en diseño de interfaz de usuario.
Ye et al. (2023)	RoboGPT, ChatGPT.		Voz.	Experimentación y encuesta.	15 participantes mayores de edad.
Rivas et al. (2018)	Chatbot revelado.	no	Texto.	Encuestas.	180 participantes estadounidenses mayores de edad.

Fuente: Elaboración propia.

3.3.3 INSTRUMENTOS UTILIZADOS

La práctica de la utilización de heurísticas es bastante común para guiar el diseño y hacer la evaluación de interfaces. Adicionalmente, se han utilizado entrevistas y encuestas para determinar la fiabilidad de los agentes conversacionales.

3.3.3.1 *Entrevista*

Juaréz (2018) describió las ventajas de la entrevista al detallar que estas logran un mayor peso e importancia si se siguen las reglas básicas para aplicarlas. Un ejemplo de esto es registrar las entrevistas adecuadamente, poniendo la suficiente atención, objetividad y sutileza requerida en el lenguaje no verbal. Pues los gestos y actitudes que la acompañan, lograr recolectar más información cercana a los pensamientos del entrevistado. De esta manera, se obtendrán respuestas con un mayor nivel de confiabilidad y validez.

3.3.3.2 *Encuesta*

Avila et al. (2020) describió a la encuesta como un método de empírica que hace uso de un formulario, con el propósito de conocer las respuestas sobre una problemática de estudio. Una característica de la encuesta es que los participantes la llenan por sí mismos. No se debe confundir la encuesta con otros métodos de recolección de datos que igualmente hacen uso de un cuestionario. Una de las mayores ventajas que posee la encuesta es que permite obtener una mejor manera de obtener opiniones en un grupo grande de personas, que pueden o no, estar dispersas geográficamente.

3.3.3.3 *Heurística*

La heurística es un método que hace uso de un conjunto de técnicas para encontrar soluciones. El beneficio de hacer uso de la heurística yace en la habilidad que posee este procedimiento para aportar soluciones óptimas o cercanas a este nivel, para problemas de pequeña escala; bajo un proceso realizado en poco tiempo.

Generalmente, la heurística de usabilidad se enfoca en brindar una experiencia clara y eficiente. Sin embargo, el diseño de interfaces de agentes conversacionales puede necesitar considerar aspectos más allá de la usabilidad. Garantizar una experiencia de usuario satisfactoria puede implicar evaluar tanto la calidad de la conversación como las interacciones del usuario. Es por eso por lo que la heurística creada por Langevin et al en 2021, incorporó las heurísticas de preservación del contexto y fiabilidad (Anexo 2).

3.4 MARCO LEGAL

En virtud de la investigación y en consideración de las regulaciones tanto a nivel nacional como en regiones de Latinoamérica y Europa, se pretende profundizar en el conocimiento de las legislaciones y normativas destinadas a salvaguardar los derechos de los individuos en lo que concierne a la aplicación de la inteligencia artificial, con un enfoque primordial en el ámbito de la salud. Este enfoque se alinea con el área de especialización de la investigación, la ingeniería biomédica.

3.4.1 LEGISLACIÓN INTERNACIONAL

En abril de 2021, se presentó ante la Comisión de la Unión Europea un proyecto de regulación conocido como la Ley de Inteligencia Artificial, que busca establecer normativas para el ámbito de las inteligencias artificiales y regular sus diversos usos, incluso prohibiendo ciertos usos específicos. Dentro de este proyecto de ley se contempla la creación de un comité a nivel europeo denominado Comité Europeo de Inteligencia Artificial. Este comité será responsable de brindar asesoramiento y asistencia tanto a los organismos reguladores de los Estados miembros en relación con las inteligencias artificiales, como a la propia Comisión Europea, abarcando temas como el impacto en el mercado y la aplicación de esta ley una vez que entre en vigor (Cuesta Ríos, 2021).

Uno de los desafíos actuales en el campo de la inteligencia artificial es garantizar el control y evitar la presencia de cualquier tipo de sesgo en sus sistemas. La presencia de sesgos representa una preocupación significativa en el ámbito de la inteligencia artificial, no solo desde un punto de vista ético, sino también de acuerdo con el marco ético establecido para

la inteligencia artificial, la robótica y las tecnologías relacionadas en sistemas de alto riesgo, donde se requerirá el cumplimiento de medidas especializadas para prevenir su aparición. Aunque los sesgos están prohibidos en los sistemas de alto riesgo, es posible que las inteligencias artificiales apliquen tratamientos diferenciados a diferentes grupos de personas cuando este tratamiento sea objetivo, razonable y legítimo (Cuesta Ríos, 2021).

3.4.2 LEGISLACIÓN REGIONAL

La Red Iberoamérica de Protección de Datos que la Inteligencia Artificial engloba un conjunto diverso de técnicas computacionales y procesos enfocados en mejorar la capacidad de las máquinas para llevar a cabo diversas actividades. Esto abarca desde modelos algorítmicos hasta sistemas de aprendizaje automático (machine learning) y técnicas de aprendizaje profundo (deep learning).

Así mismo en la "Declaración relativa a la ética y protección de datos en la Inteligencia Artificial" se reconoce el vínculo entre la recolección, uso y revelación de información personal y el desarrollo de ciertas áreas de la IA.

En América Latina, en los últimos han mostrado avances en la creación de estrategias nacionales de IA. Siendo como ejemplo, Argentina creó el "Plan Nacional de Inteligencia Artificial" llamado *ARGENIA*; Brasil cuenta con su "Estrategia Brasileña de Inteligencia Artificial" *EBIA*; Chile inició sus trabajos con un "Plan de Trabajo de Inteligencia Artificial"; Colombia decretó la "Política Nacional para la Transformación Digital e Inteligencia Artificial"; México es considerado el primer país de la región en iniciar sus trabajos para la creación de la "Estrategia de Inteligencia Artificial", y en Uruguay construyó la "Estrategia de Inteligencia Artificial para el Gobierno Digital" (García Benítez & Ruvalcaba-Gómez, 2021).

Se desarrolló una investigación mediante la cual no se logró identificar artículos, documentos o publicaciones donde se hiciera referencia directa a un ente para la protección de datos en la inteligencia artificial, por ello como alternativa, se tiene planificado utilizar la información recopilada por los entes reguladores de otros países latinoamericanos para abordar esta situación. Posterior a esto no se logró identificar artículos o documentos publicados referentes al marco legal.

IV. METODOLOGÍA

En este capítulo se presenta la metodología aplicada en el desarrollo de la investigación para abordar y resolver el problema definido. Para esto se detallan temas como el enfoque, variables dependientes e independientes, técnicas e instrumentos utilizados, y la población y muestra estudiada. Continuando, se detalla la metodología de estudio basada en el modelo en V y para finalizar se describe la metodología de validación utilizada.

4.1 ENFOQUE

En la presente investigación se utilizó el enfoque mixto. Hernández-Sampieri & Mendoza, (2018) definen este método como una serie de enfoques sistemáticos, empíricos y críticos de investigación que involucran la recopilación y el análisis tanto de datos cuantitativos como cualitativos. Estos datos se integran y discuten en conjunto con el objetivo final de obtener una comprensión más profunda del fenómeno que está siendo investigado.

En la presente investigación, el tipo de estudio seleccionado fue el transversal debido a que los datos se recolectaron en un periodo de tiempo. El tipo de diseño fue el preexperimental con un alcance descriptivo, ya que se desea conocer la trascendencia de la implementación de esta tecnología, previo a ejecutarlo como tal. Finalmente, se definió una muestra no probabilística ya que esta se eligió tomando en cuenta los criterios de la investigación (Tabla 7).

Tabla 7 - Metodología de la investigación.

Tipo de enfoque	Mixto
Tipo de alcance	Descriptivo
Tipo de estudio	Transversal
Tipo de diseño	Preexperimental
Tipo de muestra	No probabilística

Fuente: Elaboración propia.

4.2 VARIABLES DE INVESTIGACIÓN

En este apartado se detallan las variables de investigación. Para la ejecución del estudio, se plantó una variable dependiente y múltiples variables independientes que se relacionaron entre sí (Ilustración 5).

4.2.1 VARIABLE DEPENDIENTE

La variable dependiente es la que se ve influenciada bajo manipulación de las variables independientes. Se estableció un estudio mono céntrico con relación a una sola variable

dependiente. Esta se definió como la aplicabilidad y usabilidad de chatbots en las actividades del ingeniero clínico.

4.2.2 VARIABLES INDEPENDIENTES

Las variables independientes son aquellas que se pueden modificar con el objetivo final de obtener el resultado deseado, estas no se ven alteradas por las demás. Estas fueron elegidas tomando como referencias los puntos específicos a evaluar para permitir alcanzar las respuestas de las preguntas de investigación planteadas. Se consideraron las siguientes variables:

- **Arquitectura y entrenamiento.** La arquitectura del chatbot establece su funcionamiento y la manera en que lleva a cabo el procesamiento lingüístico. Por otro lado, el proceso de entrenamiento conlleva la instrucción del modelo en la tarea de comprender y generar respuestas pertinentes, basándose en ejemplos de entrada y salida provistos a lo largo del proceso de calibración. Se evaluó haciendo uso de una heurística de usabilidad.
- **Costos operacionales.** Hace referencia al precio de adquisición del chatbot, ya sea mensual o anual. Se realizó un análisis comparativo de los costos de suscripción de los agentes conversacionales.
- **Acceso tecnológico.** Es imprescindible disponer de acceso a sistemas informáticos y conectividad a la red de Internet para la implementación y utilización efectiva de chatbots. Asimismo, se refiere al nivel de comprensión y pericia que el personal posee con relación a las tecnologías de la información y la comunicación. Se estudió mediante una encuesta de diagnóstico para conocer la situación actual de los expertos, y los instrumentos a su disponibilidad.
- **Aceptación.** Alude al grado de aprobación y confianza que el personal otorga a los chatbots. En ausencia de un nivel sustancial de aceptación, se compromete la cohesión del proceso de integración y la eficacia de los agentes conversacionales se ve disminuida. Se aplicó una encuesta de aceptación para conocer el nivel de percepción de la muestra con relación al uso de chatbots.

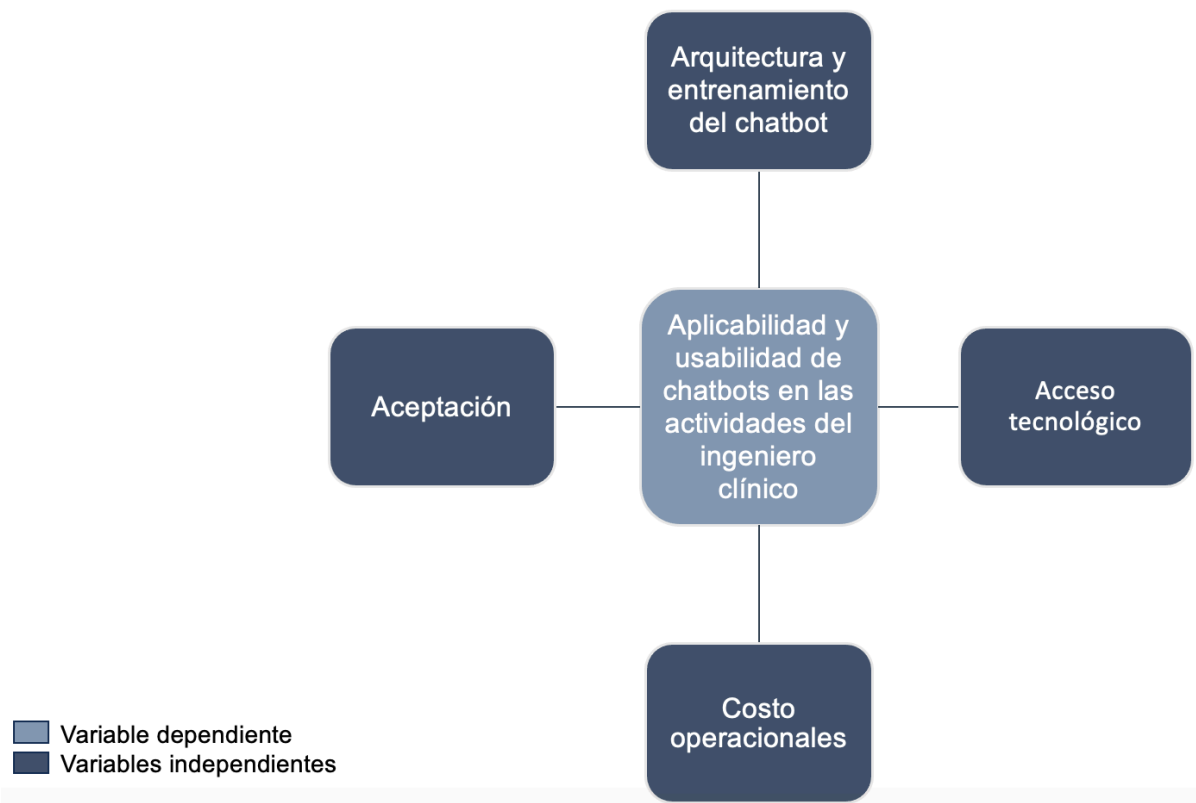


Ilustración 5 - Diagrama de variables de investigación.

Fuente: Elaboración propia.

4.3 TÉCNICAS E INSTRUMENTOS UTILIZADOS

En esta subdivisión se presentan las técnicas e instrumentos utilizados, los cuales fueron esenciales en el proceso del desarrollo del estudio.

4.3.1 TÉCNICAS

4.3.1.1 Encuesta

Según Hernández-Sampieri & Mendoza (2018), la encuesta es de los más utilizados para la recolección de datos; consiste en un conjunto de preguntas con respecto a una o más variables a medir. Una vez que se ha establecido el protocolo de la encuesta, se procede a someterlo a un proceso de validación con el propósito de identificar sus limitaciones y mejorar los elementos antes de su implementación en la prueba piloto. Esta última fase se somete a un análisis que establece su confiabilidad, lo cual es esencial para su uso como evaluación final.

Tras la creación del protocolo, es imperativo que se someta a un proceso de validación, en este caso, llevado a cabo por un panel de cinco expertos en el campo relevante. El propósito de este proceso es evaluar la validez de contenido del protocolo, asegurando así su capacidad para generar resultados fidedignos y auténticos.

La validez se refiere a la capacidad intrínseca de un enfoque de investigación para abordar y resolver de manera efectiva las preguntas planteadas. La fiabilidad, por otro lado, denota la consistencia en la obtención de resultados equivalentes a través de múltiples circunstancias y contextos diversos.

4.3.1.2 *Heurística de usabilidad*

La usabilidad se configura como un atributo intrínseco de calidad que aborda la evaluación de la facilidad de empleo en las interfaces de usuario. El término "usabilidad" también encapsula los procedimientos concebidos para mejorar la simplicidad de utilización durante el curso del proceso de diseño.

La evaluación de la usabilidad se erige como un paso cardinal en el itinerario del diseño orientado al usuario. Diversos métodos se encuentran disponibles para determinar el grado de usabilidad en sistemas interactivos de software. Entre estos, la evaluación heurística se destaca como el enfoque preponderante. En esta metodología, los evaluadores examinan la interfaz de un sistema apoyados en un conjunto de preceptos de diseño heurísticos, diseñados específicamente para maximizar la usabilidad.

Múltiples compilaciones de heurísticas de usabilidad coexisten, abarcando desde formulaciones generales hasta aquellas que atienden peculiaridades contextuales. La literatura confirma que muchas de estas heurísticas han derivado de las heurísticas de Nielsen, ampliándose o refinándose en función de requerimientos particulares. En el panorama actual, ha surgido una propuesta metodológica para la definición de heurísticas de usabilidad. Las evaluaciones de usabilidad posibilitan la detección y subsanación de inconvenientes particulares de usabilidad, con el fin de mejorar la interacción entre los usuarios y el sistema.

La evaluación heurística se erige como uno de los métodos prominentes en el ámbito de la inspección de la usabilidad. Este enfoque involucra a un conjunto de peritos que proceden de manera sistemática a analizar la interfaz de un sistema, considerando un conjunto de principios heurísticos bien conocidos en el ámbito del diseño de la usabilidad. Nielsen propuso una serie de diez heurísticas de usabilidad ampliamente utilizadas en diversos tipos de evaluaciones. A pesar de la simplicidad de aplicación de este método, ocasionalmente la identificación de problemas de usabilidad no resulta tan evidente, ya que estos inconvenientes pueden ser demasiado generales para evaluar sistemas con características altamente específicas (Jiménez et al., 2012).

Nielsen (1994) en su libro "Métodos de inspección de la usabilidad", planteó una serie de reglas que todo sistema debería cumplir y comunes a las distintas interfaces que nos podemos encontrar, suelen ser las bases para realizar una evaluación de usabilidad:

- 1) Visibilidad y estado del sistema.
- 2) Coincidencia entre mundo real y sistema.
- 3) Control y libertad del usuario.
- 4) Estándares y consistencia.
- 5) Prevención de errores.
- 6) Reconocimiento para evitar el recuerdo.
- 7) Flexibilidad y eficiencia de uso.
- 8) Diseño estético y minimalista.
- 9) Ayuda a los usuarios para reconocer, diagnósticas y recuperarse de los errores.
- 10) Ayuda y documentación.

4.3.2 ELECCIÓN DE INSTRUMENTOS

ChatGPT y ChatSonic se destacan como los chatbots de mayor popularidad, contando con una base de usuarios de 10 millones cada uno. Estas soluciones son ampliamente reconocidas por su competencia en tareas relacionadas con la creatividad y la conversación, tales como la generación de texto, la traducción de idiomas y la creación de diversos tipos de contenido creativo.

Por otro lado, Google Bard, un chatbot experimental desarrollado por Google AI, ha conseguido rápidamente notoriedad debido a su capacidad para proporcionar información precisa y esclarecedora. Google Bard cuenta con una base de usuarios de 5 millones y es especialmente popular para llevar a cabo tareas informativas e investigativas.

En cuanto a Perplexity, se presenta como un motor de búsqueda conversacional que también puede generar texto, traducir idiomas y producir distintos tipos de contenido creativo. Perplexity goza de una base de usuarios de 5 millones y es apreciado en escenarios que demandan acceso a información proveniente de diversas fuentes.

En resumen, ChatGPT, ChatSonic, Google Bard y Perplexity se configuran como chatbots poderosos, cada uno con sus propias características (Tabla 8). La elección más apropiada dependerá en última instancia de las necesidades particulares de cada usuario.

Tabla 8 - Características por la que destaca los instrumentos elegidos.

Características	ChatGPT	Bard	Perplexity	Chatsonic
Popularidad	Alta	Media	Media	Alta
Tiempo de existencia	Desde 2022	Desde 2023	Desde 2023	Desde 2022
Base de datos	Texto y código de internet	Texto y código de Google	Texto y código de internet	Texto y código de internet
Interfaz amigable	Sí	Sí	Sí	Sí
Accesibilidad	Gratuita	Gratuita	Gratuita	Gratuita
Número de usuarios	10 millones	5 millones	5 millones	10 millones
Uso principal	Creatividad y conversación	Información e investigación	Información e investigación	Creatividad y conversación

Fuente: Elaboración propia.

4.3.3 INSTRUMENTOS

Un chatbot o robot de chat, en pocas palabras es un programa informático que se comunica con un humano a través de mensajes de texto o voz, siendo esto en tiempo real y de manera que se personaliza a las necesidades del usuario. Estos se han desarrollado mediante inteligencia artificial con la unión de la tecnología de procesamiento del lenguaje natural; en el cual su diseño se fundamenta en la simulación de un entorno de conversación entre un programa y un ser humano de forma muy similar a una conversación real de seres humanos (Chandel et al., 2019).

En el presente momento, se observa una notable presencia de chatbots en áreas predominantemente representadas por los centros de atención al cliente, el comercio electrónico, el ámbito de la salud y las plataformas de mensajería. Estos sistemas asisten de manera efectiva a los usuarios, quienes únicamente requieren emitir mensajes breves para adquirir respuestas que suplen sus necesidades y que, de otra forma, no lograrían obtener al explorar sitios web afines o al establecer comunicación telefónica.

4.3.3.1 ChatGPT

El funcionamiento del modelo de GPT (Generative Pre-trained Transformer) y, por ende, de ChatGPT (Ilustración 6), se basa en una arquitectura de red neuronal denominada "Transformer" que ha demostrado su eficacia en tareas de procesamiento de lenguaje natural. (OpenAI, 2022). De manera simplificada su principio de funcionamiento está basado en (Ilustración 7):

- 1) Preentrenamiento. Fase en la cual el modelo expone cantidades enormes de texto de diversos tipos y temáticas. Mediante este proceso denominado como "aprendizaje no

supervisado”, en el cual el modelo trata de aprender patrones, estructuras gramaticales, significados y relaciones entre palabras y frases.

- 2) Atención y contexto. El componente clave de la arquitectura Transformer es la atención. El modelo descompone las oraciones en tokens lo que luego, calcula la atención de cada uno de los tokens. Permitiendo capturar el contexto a largo plazo y comprender mejor las dependencias en el texto.
- 3) Generación de texto. El modelo ha aprendido a asociar palabras y generar texto coherente. En la fase de ajuste fino (fine-tuning), se adapta el modelo para tareas específicas, como responder preguntas o generar conversaciones.
- 4) Interacción. Cuando se realiza una interacción con ChatGPT, se ingresa el texto en forma de oración o pregunta. Lo cual el modelo analiza las palabras y su contexto utilizando la atención y el conocimiento que adquirió previamente, por último, genera respuestas basadas en esa comprensión.
- 5) Contexto y longitud. Su memoria no es perfecta y se basa en una ventana limitada de tokens anteriores. Esto puede llevar a respuestas coherentes en el contexto, pero también a olvidar información a medida que aumenta la longitud de la conversación.
- 6) Salida generada. La respuesta generada por el modelo es una predicción basada en las probabilidades aprendidas durante el entrenamiento. Puede variar en calidad y precisión según la complejidad de la solicitud y la comprensión contextual.

En resumen, la funcionalidad de ChatGPT se basa en la capacidad del modelo para aprender patrones lingüísticos y generar textos coherentes en función del contexto proporcionado en una conversación. Su capacidad de generar respuestas relevantes se deriva de su entrenamiento en una gran cantidad de datos de lenguaje natural.



Ilustración 6 - Logo de ChatGPT.

Fuente: (Vecteezy, 2023)

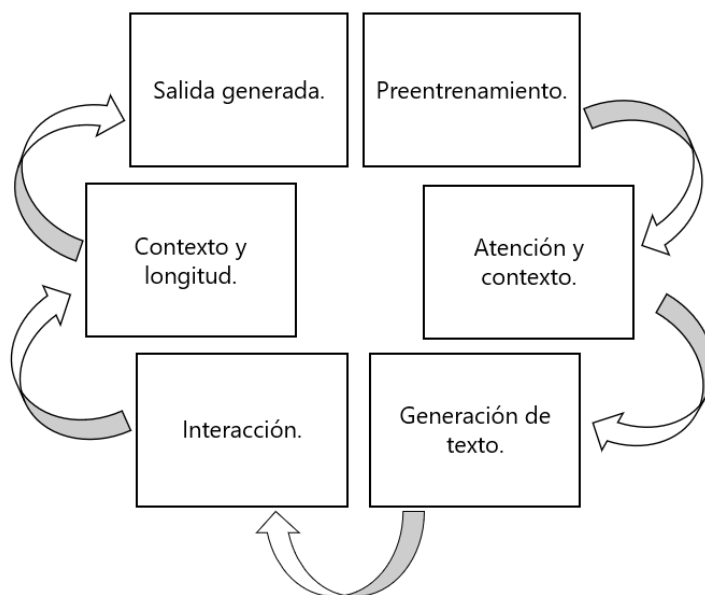


Ilustración 7 - Flujo del funcionamiento de ChatGPT.

Fuente: Elaboración propia.

4.3.3.2 Google Bard

Google Bard (Ilustración 8) es un modelo de lenguaje factual de Google AI, entrenado en un conjunto de datos masivo de texto y código. Puede generar texto, traducir idiomas, escribir diferentes tipos de contenido creativo y responder a sus preguntas de manera informativa (Google, 2023).

Bard se basa en la arquitectura Transformer, que fue desarrollada por Google AI en 2017. Transformer es un modelo de redes neuronales que es particularmente bueno para tareas de procesamiento de lenguaje natural (PLN). Bard está entrenado en un conjunto de datos masivo de texto y código, que incluye libros, artículos, código y otros tipos de texto. Este conjunto de datos es lo que le permite a Bard generar texto, traducir idiomas, escribir diferentes tipos de contenido creativo y responder a sus preguntas de manera informativa. (Google, 2023).

Bard todavía está en desarrollo, pero ha aprendido a realizar muchos tipos de tareas, incluyendo (Ilustración 9):

- Seguir instrucciones y completar solicitudes cuidadosamente.
- Usar su conocimiento para responder a sus preguntas de manera integral e informativa, incluso si son abiertas, desafiantes o extrañas.
- Generar diferentes formatos de texto creativo de contenido de texto, como poemas, código, guiones, piezas musicales, correo electrónico, cartas, etc.

Bard está disponible a través de la consola de Google AI y la biblioteca de búsqueda de Google.



Ilustración 8 - Logo de Google Bard.

Fuente: (Branditecture, 2023)

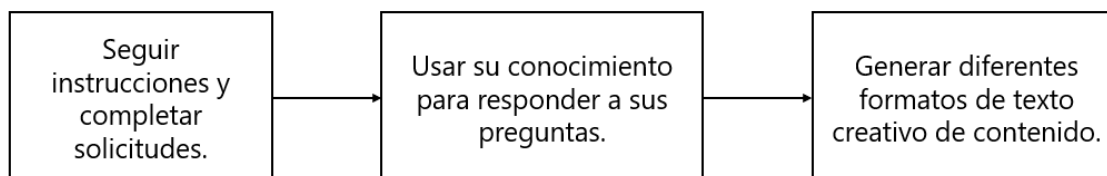


Ilustración 9 - Flujograma del funcionamiento de Google Bard.

Fuente: Elaboración propia.

4.3.3.3 Chatsonic

Funciona basándose en una combinación de algoritmos de procesamiento del lenguaje natural (PLN), aprendizaje automático (AM) y aprendizaje profundo (AD). NLP permite a ChatSonic (Ilustración 10) entender y procesar el lenguaje humano, mientras que los algoritmos ML y DL le permiten aprender de los datos de entrenamiento y generar texto similar al humano.

El principio de funcionamiento implica varios pasos clave (Ilustración 11):

- 1) Preprocesamiento: El texto de entrada se transforma en tokens y se convierten en una representación numérica que el modelo pueda entender.
- 2) Codificación: El texto preprocesado se hace pasar por una arquitectura codificador-decodificador, como una red neuronal recurrente (RNN) o un modelo transformador. Este paso captura la información contextual de la entrada.
- 3) Entrenamiento: El modelo se entrena en un gran conjunto de datos de texto, donde aprende a predecir la siguiente palabra más probable dadas las palabras anteriores. Para ello se utilizan técnicas como el aprendizaje supervisado o el aprendizaje por refuerzo.
- 4) Inferencia: Cuando un usuario introduce una pregunta o indicación, el modelo entrenado genera una respuesta basada en los patrones y estructuras que ha aprendido durante el entrenamiento.

- 5) Aprendizaje iterativo: ChatSonic aprende y mejora continuamente a lo largo del tiempo incorporando los comentarios de los usuarios y volviendo a entrenarse con conjuntos de datos actualizados. Este proceso le permite adaptarse a los cambios en los patrones lingüísticos y las preferencias de los usuarios.



Ilustración 10 - Logo de Chatsonic.

Fuente: (Hitakshi, 2023)

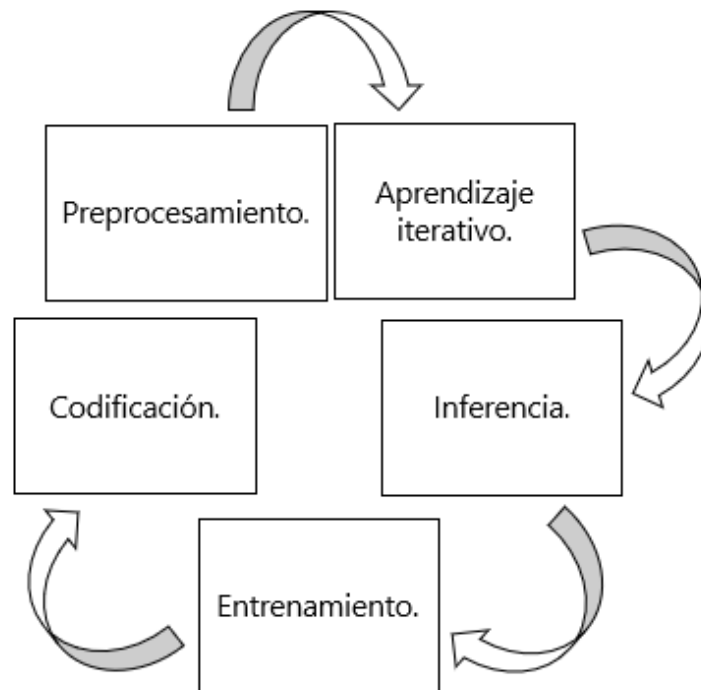


Ilustración 11 - Flujograma del funcionamiento de Chatsonic.

Fuente: Elaboración propia.

4.3.3.4 *Perplexity*

El principio de funcionamiento de Perplexity (Ilustración 12) puede explicarse en el contexto de la teoría de la información y los modelos lingüísticos. La perplejidad es una medida de lo bien que una distribución de probabilidad o un modelo de probabilidad predice una muestra. A menudo se utiliza como métrica de evaluación de los modelos lingüísticos en el Procesamiento del Lenguaje Natural (PLN). (Perplexity AI, 2023).

He aquí una explicación científica del principio de funcionamiento de la perplejidad (Ilustración 13):

- 1) **Distribución de probabilidades:** La perplejidad se basa en el concepto de distribución de probabilidad. En el contexto de los modelos lingüísticos, una distribución de probabilidad asigna probabilidades a diferentes palabras o secuencias de palabras en función de su probabilidad en un conjunto de datos dado.
- 2) **Precisión de la predicción:** La perplejidad mide lo bien que una distribución de probabilidad predice una muestra. Una perplejidad baja indica que la distribución de probabilidad es buena para predecir la muestra, mientras que una perplejidad alta sugiere que la distribución es menos precisa en sus predicciones.
- 3) **Métrica de evaluación:** La perplejidad se utiliza como métrica de evaluación de los modelos lingüísticos. Ayuda a comparar distintos modelos y a evaluar su rendimiento a la hora de generar textos coherentes y con sentido.
- 4) **Probabilidad inversa:** La perplejidad puede definirse como la probabilidad inversa del conjunto de pruebas, normalizada por el número de palabras. Representa el número medio de palabras que pueden codificarse con un número determinado de bits.
- 5) **Entropía cruzada:** Otra forma de definir la perplejidad es mediante la entropía cruzada. La entropía cruzada mide el número medio de bits necesarios para codificar una palabra, y la perplejidad es el número de palabras que pueden codificarse con esos bits.

En resumen, la perplejidad es una métrica utilizada para evaluar el rendimiento de los modelos lingüísticos. Mide lo bien que una distribución de probabilidad o un modelo predicen una muestra, y una perplejidad menor indica una mayor precisión en la predicción. Puede definirse en términos de probabilidad inversa o entropía cruzada.



Ilustración 12 - Logo de Perplexity.

Fuente: (Diction, 2023)

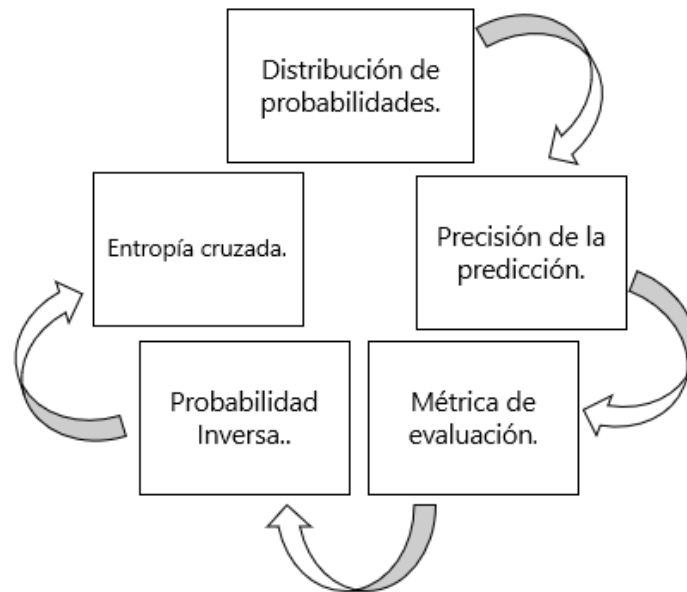


Ilustración 13 - Flujograma del funcionamiento de Perplexity.

Fuente: Elaboración propia.

4.3.3.5 Microsoft forms

Microsoft Forms es una herramienta basada en la nube que permite la creación y administración de encuestas y formularios en línea. Su funcionamiento se fundamenta en la recopilación y análisis de datos estructurados, permitiendo a los usuarios diseñar cuestionarios personalizados y obtener información significativa de manera eficiente.

La dinámica de Microsoft Forms involucra la creación de formularios digitales con preguntas y opciones de respuesta diversas, tales como casillas de verificación, preguntas de opción múltiple, escalas de calificación y campos de texto libre. Una vez configurado el formulario, se genera un enlace que puede ser compartido con los participantes a través de diversos canales, como correo electrónico o redes sociales.

Los datos recopilados a través de Microsoft Forms se almacenan en una hoja de cálculo de Microsoft Excel, lo que facilita su organización y análisis. La herramienta ofrece capacidades de análisis básicas, como la generación de gráficos y tablas para visualizar los resultados. Además, la integración con otras herramientas de Microsoft, como Power BI, permite un análisis más avanzado y una comprensión profunda de los datos recopilados.

En resumen, Microsoft Forms opera como una plataforma versátil para la recopilación y análisis de datos estructurados a través de encuestas y formularios en línea. Facilita la generación y distribución de cuestionarios, así como el procesamiento y visualización efectiva de los resultados (Microsoft, 2023a).

4.3.3.6 *Microsoft Excel*

Microsoft Excel, es una aplicación de software desarrollada por Microsoft Corporation que se emplea principalmente para el trabajo contable y financieras gracias a sus funciones, diseñadas específicamente para crear y trabajar con hojas de cálculo.

Este software es ampliamente utilizado en entornos empresariales, académicos y personales para diversidad de propósitos, que trascienden desde la contabilidad y administración financiera hasta la monitorización de inventarios, la planificación presupuestaria, la confección de informes y el escrutinio de datos estadísticos (Microsoft, 2023b).

4.3.3.7 *Minitab*

Minitab es un software estadístico dirigidos a todo tipo de organizaciones que permite analizar datos complejos y resolver problemas en el proceso de producción. Este software exhibe una amplia gama de alternativas altamente precisas y sumamente intuitivas, lo cual posibilita a numerosas empresas a nivel global supervisar y optimizar los resultados de sus operaciones productivas. Esto se traduce en la capacidad de representar visualmente los resultados obtenidos mediante diversos métodos gráficos y estadísticos, tales como gráficos, tablas, diagramas, histogramas y pruebas de hipótesis. En esencia, Minitab habilita la apropiada manipulación y utilización de la información, lo que motiva su empleo frecuente en el ámbito del control de calidad (Minitab, 2023).

4.4 MUESTRA

El tipo de muestreo utilizado es no probabilístico, esto está fundamentado en la consideración de estudiar a ingenieros biomédicos con experiencia en el área de la ingeniería clínica. Este requisito es indispensable, siendo indistinto si laboran en hospitales, clínicas, fundaciones o empresas, ya sean de carácter público o privado. La única consideración es su experiencia directa o indirecta con el ámbito hospitalario y conocimiento sobre las necesidades que se presentan en las actividades diarias de la naturaleza de este ámbito.

4.5 METODOLOGÍA DE ESTUDIO

Se utilizó el Modelo en V (Ilustración 14) propuesto en VDI 2206 – Una nueva guía para el diseño de sistemas mecatrónicos (Gausemeier & Moehring, 2002). Cabe destacar que este modelo fue modificado y enfocado en el desarrollo de software; para el cual cuenta con dos ramas, una de verificación y otra de validación. Se dividió en niveles específicos, cada uno con

su respectivo proceso de evaluación, esto para alcanzar el objetivo final que es la implementación de un sistema de IA para un departamento de biomédica.

4.5.1 DEFINIR PAUTAS EN EL PROCESO DE EVALUACIÓN DE LOS CHATBOTS

La definición de pautas para el proceso de evaluación de chatbots implicó establecer un conjunto de criterios y procedimientos sistemáticos para medir y analizar el rendimiento y la efectividad de los chatbots. Estas pautas fueron esenciales para asegurarse de que los chatbots cumplan con los objetivos previstos y proporcionen una experiencia de usuario coherente y satisfactoria. Para esto se tomaron en consideración los siguientes aspectos:

- 1) Definición de objetivos. Establecer claramente los objetivos y propósitos del chatbot. ¿Qué se espera que logre el chatbot? ¿Cuál es su función principal? Esto proporciona una base para la evaluación.
- 2) Diseño de escenarios de evaluación. Creación de situaciones o escenarios realistas en los que se pondrá a prueba el chatbot. Estos escenarios cubrieron una variedad de posibles interacciones que los usuarios podrían tener con el chatbot.
- 3) Selección de métricas de evaluación. Elección de métricas cuantificables y cualitativas para medir el rendimiento del chatbot. Se incluyó tasas de resolución de consultas, tiempo de respuesta, tasa de satisfacción del usuario y comprensión de preguntas complejas.
- 4) Conjunto de datos de prueba. Preparación de un conjunto de datos diverso que refleje situaciones del mundo real. Estos datos se utilizaron para evaluar cómo el chatbot maneja diversas interacciones y desafíos.
- 5) Pruebas de funcionalidad. Evaluar si el chatbot cumple con las funciones básicas para las que fue creado. ¿Puede comprender y responder adecuadamente a las preguntas de los usuarios? ¿Puede llevar a cabo tareas específicas, como proporcionar información o realizar acciones?
- 6) Pruebas de contexto y continuidad. Evaluación de la capacidad del chatbot para mantener un contexto coherente a lo largo de una conversación y recordar interacciones anteriores.
- 7) Evaluación de la experiencia del usuario. Recopilación de opiniones de los usuarios para medir su satisfacción general. Esto se hizo a través de encuestas.

- 8) Comparación con otros chatbots. Comparar el rendimiento del chatbot con otros similares para determinar su competitividad.
- 9) Seguridad y privacidad. Revisar que el chatbot cumpla con las normativas de seguridad y privacidad, especialmente si maneja datos sensibles de los usuarios.

4.5.1.1 *Análisis de la literatura científica fundamentado en el enfoque de ingeniería clínica*

En el enfoque de ingeniería clínica aplicado al análisis de literatura científica se vio involucrada una metodología que busca comprender y evaluar de manera sistemática aspectos técnicos y clínicos de investigación de campo de la salud y la medicina.

Siendo que la ingeniería clínica se basa en la aplicación de principios y técnicas de ingeniería para mejorar la eficiencia, seguridad y efectividad de los diferentes dispositivos médicos, sistemas de atención médica y procesos clínicos. Al aplicar este enfoque al análisis de la literatura científica, se buscó identificar avances tecnológicos, mejoras en la atención médica y oportunidades para la innovación en este ámbito.

El análisis de esta literatura científica se basó de acuerdo con una estructura sugerida para realizar esta tarea, siendo:

- 1) Selección de artículos. Localización de las publicaciones pertinentes vinculadas a la disciplina de la ingeniería clínica y la implementación de tecnologías en el dominio de la salud.
- 2) Revisión de antecedentes. Adquisición de una comprensión profunda del contexto histórico y tecnológico en el cual se enmarca a la investigación. Análisis de la previa aplicación de los principios de la ingeniería en problemas clínicos similares.
- 3) Objetivos y métodos. Evaluación de los objetivos perseguidos por el estudio y los métodos para emplear su consecución. Identificación de la adopción de un enfoque de ingeniería para abordar específicamente una cuestión clínica.
- 4) Resultados técnicos. Se examinaron los resultados técnicos presentados en el artículo. Se evaluó si los enfoques de ingeniería propuestos han logrado los objetivos.
- 5) Discusión e interpretación. Análisis de la discusión del autor sobre los resultados obtenidos y como fueron relacionados con los objetivos iniciales.

- 6) Conclusiones y futuras direcciones. Se evaluaron las conclusiones del estudio en relación con los avances en ingeniería clínica.
- 7) Comparación con trabajos previos. Se llevó a cabo una comparación de los resultados y enfoques presentados en el artículo con otros estudios similares.
- 8) Impacto y aplicabilidad. Consideración del posible impacto de los resultados en la esfera de la práctica clínica y la industria de la salud.
- 9) Valoración crítica. Se efectuó una evaluación integral y crítica del artículo desde una perspectiva de ingeniería clínica.

4.5.2 EVALUACIÓN Y CREACIÓN DEL CUESTIONARIO JUNTO CON LA EVALUACIÓN DE USABILIDAD BASADA EN HEURÍSTICAS

La elaboración y creación del cuestionario, junto con la heurística de usabilidad, fue un proceso esencial para evaluar la experiencia del usuario y la facilidad de uso del sistema.

4.5.2.1 *Elaboración y creación del cuestionario*

- 1) Identificación de objetivos. Definición clara de los objetivos de la evaluación. ¿Qué aspectos específicos de la usabilidad se desean evaluar? Incluyó aspectos como la navegación, la claridad del contenido y la facilidad de encontrar información.
- 2) Selección de participantes. Elección de la muestra de los usuarios involucrados en la evaluación. Estos usuarios fueron representativos del público objetivo del sistema.
- 3) Diseño de preguntas. Creación de preguntas que abordaron los aspectos que se deseaban evaluar. Las preguntas fueron ser claras y específicas para recopilar información relevante.
- 4) Tipo de preguntas. Combinación de tipos de preguntas, abiertas y cerradas. Las primeras permitieron respuestas detalladas, mientras que las últimas ofrecieron opciones predefinidas y escalas de calificación que permitieron medir la satisfacción o acuerdo.
- 5) Secuencia lógica. Organización de las preguntas de manera lógica y coherente. Se comenzó con preguntas generales para establecer el contexto y luego se profundizó en aspectos específicos.
- 6) Verificación del cuestionario. Lectura del cuestionario para identificar posibles problemas, ambigüedades o preguntas confusas.

- 7) Ajustes y finalización. Basado en lo detectado en la verificación, se realizaron los ajustes necesarios en el cuestionario para garantizar claridad y comprensión. Luego, se finalizó el cuestionario.

4.5.2.2 *Heurística de usabilidad*

La evaluación de usabilidad implicó utilizar un conjunto predefinido de principios de diseño conocidos como "heurísticas" para identificar problemas de usabilidad en una interfaz. Los pasos en este proceso fueron:

- 1) Selección de heurísticas. Elección un conjunto de heurísticas de usabilidad. Estas son pautas generales que se han demostrado eficaces en la mejora de la usabilidad. Para este estudio, se utiliza la heurística de Nielsen, con las modificaciones realizadas por Langevin et al.
- 2) Inspección de la interfaz. Examinación de la interfaz del sistema en busca de problemas que violen las heurísticas seleccionadas.
- 3) Revisión. Basado en la información recopilada en los pasos previos, se mantuvo la heurística seleccionada.

4.5.2.3 *Evaluación minuciosa de registro de evaluaciones siguiendo directrices predefinidas*

La evaluación minuciosa de los registros de evaluaciones, en concordancia con directrices predefinidas, constituyó un proceso necesario en la búsqueda de la calidad y la integridad de la información recopilada, mediante esta rigurosa metodología se aseguró que las evaluaciones se llevaran a cabo de manera coherente y objetiva, evitando de esta manera sesgos y promoviendo la fiabilidad de los resultados obtenidos.

En un mundo donde la desinformación y la manipulación de la información son desafíos cada vez más presentes, confirmar la autenticidad de las respuestas se vuelve imperativo. Esto se aplica a ámbitos diversos, como el periodismo, la investigación académica, la toma de decisiones empresariales y gubernamentales, entre otros. La confiabilidad de la información es un pilar fundamental para la construcción de conocimiento y la adopción de decisiones basadas en evidencia.

La evaluación minuciosa de registros según directrices predefinidas contó con múltiples beneficios, entre ellos:

- 1) Objetividad. Al basarse en directrices establecidas de antemano, se redujo la influencia de sesgos personales en el proceso de evaluación.

- 2) Consistencia. La aplicación consistente de las directrices garantizó resultados uniformes a lo largo del tiempo.
- 3) Confianza en los resultados. Los resultados de la evaluación fueron más confiables y pueden ser utilizados de manera efectiva para tomar decisiones informadas.
- 4) Reproducibilidad. Al seguir pautas predefinidas, es más probable que otros investigadores puedan reproducir los mismos resultados al utilizar los mismos registros y directrices.
- 5) Transparencia. La aplicación de directrices predefinidas aumentó la transparencia y permite una revisión y verificación más efectiva por parte de terceros.

4.5.3 ESTABLECER CRITERIOS PARA LAS PRUEBAS

Los criterios actuaron como estándares que ayudaron a medir y juzgar el éxito o el cumplimiento de ciertos aspectos clave. Los pasos para establecer estos criterios fueron:

- 1) Definición de objetivos claros. Se tuvo una comprensión clara de los objetivos de las pruebas. ¿Qué aspecto del producto o sistema estás evaluando? ¿Qué resultados deseas obtener de las pruebas?
- 2) Identificación de aspectos clave. Desglose de los aspectos clave del chatbot, como ser: funcionalidad, usabilidad, rendimiento y seguridad.
- 3) Desarrollo de criterios específicos. Para cada aspecto identificado, se crearon criterios específicos que definan lo que se considera como cumplimiento. Estos criterios fueron claros, medibles y alcanzables.
- 4) Uso de métricas cuantificables. Cuando fue posible, se incluyó métricas cuantificables en los criterios.
- 5) Inclusión de criterios de calidad y rendimiento. Criterios relacionados con la calidad del producto y su rendimiento. Estos abordaron la ausencia de errores críticos, el cumplimiento de estándares de diseño y la capacidad de respuesta.
- 6) Se fue realista y alcanzable. Los criterios fueron realistas y alcanzables dentro del contexto de las pruebas.
- 7) Consideración de diferentes escenarios. Se garantizó que los criterios fuesen apropiados para cada uno de los escenarios en los que se llevaron a cabo las pruebas.

- 8) Aseguramiento de consistencia y objetividad. Los criterios fueron consistentes y objetivos. Se evitaron criterios ambiguos que pudiesen ser interpretados de diferentes maneras.
- 9) Documentación de los criterios. Se documentaron todos los criterios de manera clara y accesible para uso de todos los involucrados en las pruebas. Esto aseguró que todos tuviesen una comprensión común de lo que se esperaba.
- 10) Revisión y validación. Antes de comenzar las pruebas, se revisaron y validaron los criterios para garantizar que fuesen apropiados y justos.
- 11) Actualización y adaptación. Si a medida que avanzaron las pruebas, surgieron nuevas consideraciones, problemas o descubrimientos, se actualizaron o adaptaron los criterios para reflejar estos cambios.

4.5.3.1 *Confirmar la autenticidad de las respuestas*

La confirmación de la autenticidad de las respuestas emergentes fungió como un paso crítico en la búsqueda de información precisa y veraz en una variedad de contextos. Este proceso se enfocó en verificar la credibilidad y origen legítimo de las respuestas de la información y la confianza en los datos presentados.

Algunos elementos clave en el proceso de confirmación de autenticidad incluyen:

- 1) Evaluación de fuentes. Verificación de la credibilidad y reputación de las fuentes de las respuestas. Fuentes confiables y reconocidas.
- 2) Cotejo de datos. Contrastación de la información con datos provenientes de diferentes fuentes para identificar alteraciones digitales.
- 3) Verificación tecnológica. Utilización de herramientas tecnológicas y algoritmos para detectar manipulación de contenido.
- 4) Contextualización. Análisis del contexto en el que se presentan las respuestas para determinar si se alinean con la lógica y a la realidad.
- 5) Evaluación crítica. Examinación de las respuestas desde una perspectiva crítica, considerando posibles sesgos o intereses ocultos.
- 6) Colaboración y validación. En situaciones complejas, realización de búsqueda de validación y opiniones adicionales de expertos en el campo.

La confirmación de la autenticidad de las respuestas fue esencial para preservar la calidad y veracidad de la información; este proceso contribuyó a establecer bases sólidas para la toma de decisiones y el avance del conocimiento.

4.5.4 EXPLORACIÓN DEL EMPLEO DE CHATBOTS EN LAS ACTIVIDADES DEL INGENIERO CLÍNICO

Se propuso explorar y analizar el potencial empleo de un sistema de inteligencia artificial en las actividades diarias de los ingenieros clínicos; con el objetivo de mejorar la gestión, operación y eficiencia de los equipos médicos y dispositivos. El enfoque se centró en la evaluación de como la implementación de una solución basada en IA que puede incidir en la optimización de procesos, la toma de decisiones clínicas y calidad de la atención médica.

La estrategia de esta exploración implicó una revisión exhaustiva de la literatura científica actualizada en el campo de la ingeniería biomédica y la aplicación de IA en el ámbito clínico. La selección de fuentes incluye artículos científicos, publicaciones en conferencias especializadas y documentos técnicos que aborden la convergencia de estas disciplinas.

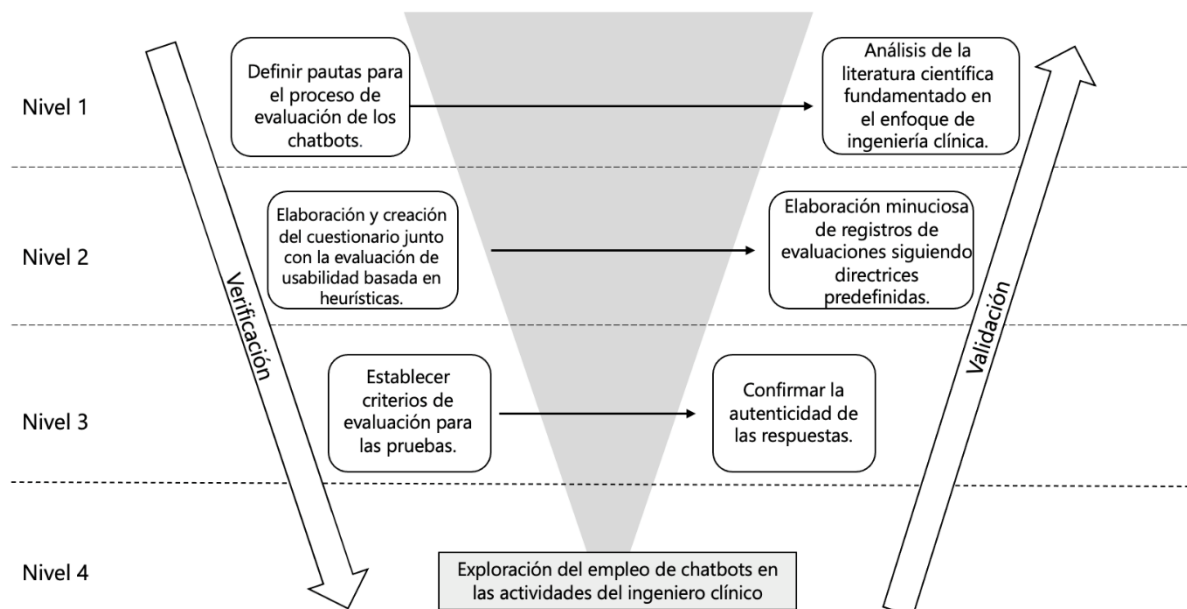


Ilustración 14 - Metodología en V.

Fuente: Elaboración propia.

4.6 METODOLOGÍA DE VALIDACIÓN

Con el propósito de asegurar la calidad y la confiabilidad de la información recabada durante el desarrollo de esta investigación, se procedió a evaluar y examinar dichos datos mediante enfoques matemático-estadísticos. La aplicación de técnicas estadísticas posibilitó la interpretación y detección precisa de patrones y tendencias presentes en los datos obtenidos. De esta manera, se obtuvieron resultados exactos y de sencilla interpretación.

4.7 CRONOGRAMA DE ACTIVIDADES

En esta sección se presenta el cronograma de actividades (Tabla 9) realizadas para el desarrollo del proyecto de investigación. Dichas actividades se llevaron a cabo en un periodo de tiempo de aproximadamente 10 semanas, entre los meses de julio y septiembre.

Tabla 9 - Cronograma de actividades

Actividades desarrolladas para la investigación	Semana 1	Semana 2	Semana 3	Semana 4	Semana 5	Semana 6	Semana 7	Semana 8	Semana 9	Semana 10
Investigación y selección de los chatbots a utilizar	■									
Creación de cuentas en los diferentes chatbots	■									
Búsqueda de metodologías aplicables		■	■							
Suscripción a los agentes conversacionales de pago				■						
Primer acercamiento a la muestra				■						
Creación de encuestas y heurística					■					
Pruebas y recolección de datos						■	■			
Análisis de resultados							■	■		
Conclusiones obtenidas								■		
Finalización del proyecto									■	

Fuente: Elaboración propia.

4.8 MATRIZ METODOLÓGICA

La metodología aplicada se sintetiza en la matriz metodológica (Tabla 10).

Tabla 10 - Matriz metodológica

Titulo	Problema de Investigación	Preguntas de investigación	Objetivos	Variables	Metodología y herramientas
Explorando el empleo de chatbots en las actividades de una unidad de biomédica en un hospital	La inteligencia artificial puede ser integrada a las labores operativas de una manera complementaria positiva en muchos escenarios. Por tanto, se desea explorar la incorporación de cuatro herramientas de inteligencia artificial al entorno de trabajo de ingeniería biomédica	General ¿Cuál es el análisis del resultado de la exposición de los ingenieros biomédicos al empleo de cuatro herramientas de inteligencia artificial para consultas asociadas con actividades laborales técnicas?	General Analizar el resultado de la exposición de los ingenieros biomédicos al empleo de cuatro herramientas de inteligencia artificial para consultas asociadas con actividades laborales técnicas.	Dependiente Empleo de chatbots en una unidad de biomédica.	Microsoft forms. Metodología de modelo en V. ChatGPT.
		Específicas ¿Qué literatura se encuentra disponible en bases de datos de publicaciones científicas con respecto a la implementación de herramientas de inteligencia artificial en la práctica profesional contemporánea a nivel nacional e internacional?	Específicos Evaluar la literatura existente en bases de datos de publicaciones científicas con respecto a la implementación de herramientas de inteligencia artificial en la práctica profesional moderna bajo el contexto nacional e internacional.	Independientes Arquitectura y entrenamiento del chatbot.	Google Bard. Perplexity AI.

Título	Problema de Investigación	Preguntas de investigación	Objetivos	Variables	Metodología y herramientas
		Específicas	Específicos	Independientes	
Explorando el empleo de chatbots en las actividades de una unidad de biomédica en un hospital	La inteligencia artificial puede ser integrada a las labores operativas de una manera complementaria positiva en muchos escenarios. Por tanto, se desea explorar la incorporación de cuatro herramientas de inteligencia artificial al entorno de trabajo de ingeniería biomédica	¿Cuál podría ser una configuración metodológica apropiada para exponer a los profesionales en ingeniería biomédica al uso de las herramientas específicas de inteligencia artificial?	Diseñar una configuración metodológica adecuada para la exposición de profesionales de la ingeniería biomédica al uso de las herramientas de inteligencia artificial elegidas para la investigación.	Acceso tecnológico.	ChastSonic.
		¿Cuál es el método adecuado para implementar la aplicación de la investigación a los profesionales seleccionados de manera discrecional?	Implementar los métodos escogidos de modelo de aplicación de la investigación a los profesionales seleccionados como muestra discrecional.	Costos operacionales.	Microsfot Excel.
		¿Cuál es el análisis resultante de la aplicación metodológica en todos los participantes, llevado a cabo a través del conjunto de técnicas prácticas y teóricas seleccionadas para este estudio?	Analizar los resultados de la aplicación metodológica a todos los participantes por medio del paquete de técnicas prácticas y teóricas escogidas para el estudio.	Aceptación.	Minitab.

Fuente: Elaboración propia.

V. RESULTADOS Y ANÁLISIS

En este capítulo, se presentaron las respuestas a las interrogantes investigativas formuladas. Asimismo, se estableció una correlación con las variables de investigación empleadas y debatieron los descubrimientos en relación con el cuerpo literario existente. Este apartado posibilitó el análisis e interpretación de los datos adquiridos mediante encuestas y heurísticas, con el propósito de derivar conclusiones entorno al empleo de estas herramientas.

5.1 EVALUACIÓN DE METODOLOGÍA EN V

Se detallan los pasos de la metodología empleados para la obtención de resultados (Ilustración 15).



Ilustración 15 - Línea de tiempo del proceso de desarrollo y formulación de evaluación de encuestas.

Elaboración propia.

5.1.1 DEFINICIÓN DE PAUTAS PARA EL PROCESO DE EVALUACIÓN DE CHATBOTS

Con el propósito de establecer directrices claras, se llevó a cabo una fase preliminar que involucró un análisis exhaustivo de la utilización de diversos sistemas de chat. Este análisis riguroso permitió determinar su idoneidad para su inclusión en el estudio. Una vez validada su pertinencia, se procedió a diseñar un marco que identificara las variables relevantes a evaluar, tomando como base las variables de investigación. Además, se delinearon los aspectos a considerar para optimizar la explotación de las capacidades de cada una de estas entidades de inteligencia artificial.

Siguiendo una metodología análoga, se llevaron a cabo pruebas preliminares de los sistemas seleccionados. Estas pruebas fueron sometidas a una revisión minuciosa para

asegurar su coherencia con los objetivos predefinidos. Estas etapas iniciales se sustentaron en la revisión de documentación académica, artículos científicos y otros recursos de investigación pertinentes, recopilados a lo largo del proceso investigativo relacionado con cada una de las plataformas de inteligencia artificial consideradas.

5.1.2 ELABORACIÓN Y CREACIÓN DE CUESTIONARIOS JUNTO CON LA EVALUACIÓN DE USABILIDAD

Al abordar la evaluación de las entidades de inteligencia artificial desarrolladas, se planteó la cuestión de qué enfoque resultaría más viable y esclarecedor. Con el propósito de adentrarnos en la totalidad de la estructura y los componentes inherentes al proceso, se idearon tres pruebas fundamentales.

La primera de estas pruebas consistió en un examen de diagnóstico o conocimientos previos, destinado a establecer una línea de base para la evaluación. La segunda fase implicó una evaluación de usabilidad que abordó directamente las capacidades de cada entidad de inteligencia artificial previamente empleada. La tercera y última prueba implicó la realización de encuestas de aceptación, considerando la posibilidad de que los encuestados, en un futuro cercano, puedan emplear estas mismas entidades de inteligencia artificial en contextos laborales subsiguientes.

Es importante resaltar que estas pruebas proporcionaron una fuente continua de información adicional, permitiendo una ampliación constante de la base de conocimiento que sirvió de apoyo para la investigación.

5.1.3 ESTABLECER CRITERIOS DE EVALUACIÓN PARA PRUEBAS

Habiendo completado la fase de elaboración de los cuestionarios, se delinearon los aspectos de relevancia que buscábamos establecer. Estos aspectos se construyeron en consonancia con los objetivos previamente definidos, asegurando que las respuestas que se obtuvieran llenaran los vacíos informativos deseados. Además, se aspiraba a profundizar en la perspectiva de utilidad a largo plazo de las inteligencias artificiales, no solo para el presente proyecto, sino también para futuros contextos. Se procuró captar las opiniones de los ingenieros participantes en esta investigación, con el propósito de evaluar cada entidad de inteligencia artificial en términos de su potencial para mejorar el trabajo en campo.

Uno de los objetivos centrales era lograr una interacción fluida y natural, de tal manera que el uso de estas inteligencias artificiales se asemejara a una experiencia auténtica y sin artificios. Asimismo, se pretendía examinar la precisión, pertinencia y coherencia de las

respuestas generadas por cada entidad, y con base en estos criterios, establecer la robustez y consistencia necesaria para fundamentar las conclusiones que se persiguen extraer de este estudio.

Además, se aplicaron directrices rigurosas y veraces para la selección de las pruebas, las cuales se diseñaron con el fin de obtener resultados sólidos. Para lograr esto, se incluyeron aspectos como el entendimiento básico de procesos de mantenimiento preventivo y la búsqueda de información detallada sobre equipos específicos, a elección de los entrevistados o ingenieros participantes. En consonancia con la aspiración de establecer una interacción lo más fluida y natural posible, se incorporaron preguntas adicionales de interés para los entrevistados. Con el propósito de mantener la coherencia informativa, se mantuvo una continuidad en la exploración, en línea con las características intrínsecas de cada entidad de inteligencia artificial considerada (Ilustración 16).



Ilustración 16 - Habilidades de los chatbots útiles en el ámbito de la ingeniería clínica.

Fuente: Elaboración propia.

En el transcurso de la evaluación, mientras se atravesaba el procedimiento de implementación del estudio y en la fase de adaptación de los agentes conversacionales, los ingenieros se sometieron a cada uno de ellos a diferentes pruebas minuciosas. Durante este proceso, pudieron discernir distintivas características y determinar las variables subyacentes que generaban variaciones entre los agentes conversacionales mencionados. Consecuentemente, los participantes en el estudio adquirieron un conocimiento más

profundo acerca de las particularidades que distinguen a ChatGPT (Anexo 4 y Anexo 5), Bard (Anexo 6 y Anexo 7), Perplexity (Anexo 8 y Anexo 9) y ChatSonic (Anexo 10 y Anexo 11). Siendo estas características lo que hacen que cada uno de estas más atractivas a los usuarios.

5.2 ANÁLISIS ESTADÍSTICO

En esta sección se expone el análisis estadístico descriptivo llevado a cabo utilizando los datos adquiridos en la investigación. Los datos recopilados fueron interpretados posibilitando una evaluación objetiva. Para esto, se utilizaron enfoques estadísticos con la asistencia de herramientas de visualización de datos como Minitab y Excel. Estas herramientas permitieron describir y analizar los datos de las encuestas y heurísticas.

5.2.1 ENCUESTA DE DIAGNÓSTICO

Dentro del marco de la encuesta efectuada a la muestra de ingenieros clínicos (Anexo 3), se recabaron y evaluaron diversos datos. Esto permitió que se identificaran tendencias y patrones en las respuestas brindadas por los participantes, mediante la fusión de enfoques cuantitativos y cualitativos. De esta manera, se logró obtener una comprensión más amplia de las experiencias y situación actual de los profesionales en relación con el tema central del proyecto.

5.2.1.1 Análisis de datos cualitativos

Con relación a las herramientas tecnológicas destinadas al empleo de chatbots en entornos de atención médica, se observó que el total de ingenieros, representado por el 100% (n=10), disponen de computadoras para su utilización, como se observa en la Tabla 11. Estos hallazgos ponen de manifiesto que, a pesar de que todos los profesionales encuestados afirmaron tener acceso a una computadora, esto no implica necesariamente que estas se encuentren en un estado óptimo.

Tabla 11 - IB con acceso a una computadora.

Acceso a computadora					
	Conteo	Porcentaje	Conteo Acumulado	Porcentaje Acumulado	
Si	10	100.00	10	100.00	
N=	10				

Fuente: Elaboración propia.

Simultáneamente, se procedió a consultar a los encuestados si, además de estar en posesión de ordenadores, disponían de conectividad a Internet (Tabla 12). Los resultados concernientes a la disponibilidad de acceso a Internet o conexión Wifi en su entorno laboral

arrojaron los siguientes datos: el 90.00% (n=9) de los ingenieros afirmaron contar con dicho acceso, mientras que el 10.00% (n=1) restante mencionó lo contrario.

Tabla 12 - IB con acceso a Wifi/Internet.

Acceso a Wifi/Internet	Conteo	Porcentaje	Conteo Acumulado	Porcentaje Acumulado
No	1	10.00	1	10.00
Si	9	90.00	10	100.00
N=	10			

Fuente: Elaboración propia.

La mayor parte de los encuestados, equivalente al 60.00% (n=6), señaló haber empleado previamente esta herramienta tecnológica tanto en el ámbito laboral como en su rutina diaria. En contraposición, el 40.00% restante (n=4) de los encuestados mencionó no haber utilizado chatbots en ninguna instancia hasta la fecha de la encuesta (Tabla 13).

Tabla 13 - Uso previo de chatbots.

Uso previo de chatbots	Conteo	Porcentaje	Conteo Acumulado	Porcentaje Acumulado
No	4	40.00	4	40.00
Si	6	60.00	10	100.00
N=	10			

Fuente: Elaboración propia.

Adicionalmente, se evidenció una notable receptividad y una valoración positiva hacia la adopción de otro género de inteligencias artificiales, como se ejemplifica con Alexa, Siri y Bixby. Una mayoría de los expertos, equivalente al 90.00% (n=9) de los participantes comunicó haber utilizado alguna de las mencionadas inteligencias artificiales en algún momento. En contraste, un 10.00% (n=1) de los encuestados respondió que no había tenido experiencia previa con ninguna de estas tecnologías (Tabla 14).

Tabla 14 - Uso previo de otras IA.

Uso de otras IAs	Conteo	Porcentaje	Conteo Acumulado	Porcentaje Acumulado
No	1	10.00	1	10.00
Si	9	90.00	10	100.00
N=	10			

Fuente: Elaboración propia.

La mayoría de los participantes, equivalente al 50.00% (n=5), manifestó situarse en un nivel intermedio de competencia en cuanto a las tecnologías de la información y la comunicación (TICs). Por otro lado, un 40.00% (n=4) de los encuestados señaló que, desde su perspectiva, poseen un nivel avanzado en este dominio, mientras que únicamente un 10.00% (n=1) de los encuestados afirmó que sus conocimientos en las TICs son particularmente sólidos, considerándose un experto en esta área (Tabla 15).

Tabla 15 - Nivel de conocimiento de las TICs de los IB.

Nivel de Conocimiento de las IA	Conteo	Porcentaje	Conteo Acumulado	Promedio Acumulado
Avanzado	4	40.00	4	40.00
Experto	1	10.00	5	50.00
Intermedio	5	50.00	10	100.00
N=	10			

Fuente: Elaboración propia.

La mayor parte de los encuestados (70.00%, n=7) sostiene la opinión de que la infraestructura tecnológica en su entorno laboral es adecuada para llevar a cabo labores con chatbots. No obstante, un 30.00% (n=3) de los participantes expresó que existe margen para mejorar en este aspecto dentro de su lugar de trabajo (Tabla 16).

Tabla 16 - Infraestructura tecnológica.

Infraestructura tecnológica	Conteo	Porcentaje	Conteo acumulado	Porcentaje acumulado
Necesita mejorar	3	30.00	3	30.00
Si	7	70.00	10	100.00
N=	10			

Fuente: Elaboración propia.

Conforme a lo previamente expuesto, los expertos encuestados mencionaron los elementos que ejercieron influencia en la potencial aceptación para emplear diversas categorías de chatbots (Tabla 17). Se constató que el 40.00% (n=4) de los profesionales asignan un alto valor a la fiabilidad y precisión de las respuestas del sistema, siendo este uno de los factores de mayor relevancia en la consideración para la eventual adopción de chatbots. De igual manera, el 40.00% (n=4) de la muestra destacó la conveniencia y facilidad de uso de los chatbots como un factor influyente en su decisión de uso. Esto subraya que los profesionales encuestados otorgan el mismo grado significativo de importancia a la

autenticidad y exactitud de las respuestas brindadas; así como a la comodidad y simplicidad de utilización de este tipo de herramientas.

Por otro lado, el 20.00% (n=2) de los participantes identificó la disponibilidad y accesibilidad del sistema como el segundo factor más influyente.

Tabla 17 - Factores que influyen el uso de chatbots.

Factores que influyen su uso	Conteo Porcentaje		Conteo acumulado	Porcentaje acumulado
Conveniencia y facilidad de uso	4	40.00	4	40.00
Disponibilidad y acceso a la herramienta	2	20.00	6	60.00
Fiabilidad y precisión de las respuestas	4	40.00	10	100.00
N=	10			

Fuente: Elaboración propia.

La incorporación de sistemas de chatbots en el contexto de la ingeniería clínica podría resultar en una notable optimización de la eficiencia y la eficacia de las funciones llevadas a cabo por estos expertos (Tabla 18). Por este motivo, exactamente el 30.00% (n=3), indican que, de acuerdo a su perspectiva, no están seguros si la introducción de estas herramientas podría facilitar su trabajo. En contraposición, una proporción significativamente mayor, es decir, el 70.00% (n=7) restante; afirma que existe la posibilidad de que los chatbots, en efecto, contribuyan a la simplificación de las tareas si se integrasen a sus labores diarias.

Tabla 18 - Implementación de chatbots y facilitación del trabajo.

La implementación de chatbots facilitaría el trabajo	Conteo	Porcentaje	Conteo acumulado	Porcentaje acumulado
Si	7	70.00	7	30.00
Tal vez	3	30.00	10	100.00
N=	10			

Fuente: Elaboración propia.

5.2.1.2 Análisis de datos cuantitativos

Se procedió a llevar a cabo un análisis estadístico de cuatro variables específicas: disponibilidad de trabajar con chatbots, confianza en este tipo de inteligencia artificial, facilidad para mantener un diálogo y ser una herramienta de solución, como se ilustra en la Ilustración 17. Para cada una de estas variables, se calcularon diversas medidas relativas a la centralidad y dispersión de los datos, como se detalla en la Tabla 19. En términos de medidas de centralidad, se constató que la variable disponibilidad de trabajo presenta una media de 4.500, mientras que confianza registra una media ligeramente inferior de 3.400. Por su parte,

tanto facilidad de diálogo como herramienta de solución exhiben medias de 3.900 y 4.400, respectivamente. Los errores estándar asociados con estas medias resultaron relativamente reducidos, oscilando entre 0.167 y 0.267.

En lo que concierne a la dispersión de los datos, se observa que las variables confianza y facilidad de diálogo presentan desviaciones estándar de 0.843 y 0.738 respectivamente; siendo estos los valores más elevados. Van acompañados por varianzas correspondientes de 0.711 y 544. En contraste, las variables disponibilidad de trabajo y herramienta de solución revelan desviaciones estándar de 0.527 y 0.699, junto a varianzas de 0.278 y 0.489, respectivamente. Los coeficientes de variación, que señalan la proporción de la desviación estándar en relación con la media, fluctúan en el rango del 11.71% al 24.80%, indicando diferentes niveles de variabilidad en las distribuciones de los datos.

En cuanto a la forma de las distribuciones, se identificó que las medianas de las variables disponibilidad de trabajo, confianza, facilidad de diálogo y herramienta de solución son de 4.500, 3.000, 4.000 y 4.500, respectivamente. Los valores de rango, que denotan la diferencia entre los valores mínimo y máximo, oscilan entre 1.000 y 3.000. El rango intercuartílico (IQR), que refleja la dispersión de los valores centrales, varía entre 1.000 y 1.250. Además, se ha establecido el valor modal, o moda, para cada variable, siendo 4 para facilidad de diálogo, 3 para confianza, 4 y 5 para disponibilidad de trabajo y 5 para herramienta de solución.

La asimetría, una métrica de la distribución de los datos en relación con la media, oscila entre -0.78 y 0.000. Valores negativos sugieren una distribución sesgada hacia la izquierda, mientras que valores ceros indican la falta de asimetría. Por su parte, la curtosis, que evalúa la forma de los picos en la distribución, exhibe valores en el rango de -2.57 a 0.37. Mientras que valores negativos indican una distribución achatada, valores positivos sugieren una distribución más puntiaguda (Ilustración 17).

Tabla 19 - Análisis estadístico comparativo de datos cualitativos.

Variable	Conteo total	N	N*	Número acumulado	Porcentaje	Porcentaje acumulado	Media	Error estándar de la media
Disponibilidad de trabajo	10	10	0	10	100	100	4.500	0.167
Confianza	10	10	0	10	100	100	3.400	0.267
Facilidad de diálogo	10	10	0	10	100	100	3.900	0.233
Herramienta de solución	10	10	0	10	100	100	4.400	0.221

Variable	Desviación estándar	Varianza	Coefficiente de variación	Mínimo	Q1	Mediana	Q3	Máximo
Disponibilidad de trabajo	0.527	0.278	11.71	4.000	4.000	4.500	5.000	5.000
Confianza	0.843	0.711	24.80	2.000	3.000	3.000	4.000	5.000
Facilidad de diálogo	0.738	0.544	18.92	3.000	3.000	4.000	4.250	5.000
Herramienta de solución	0.699	0.489	15.89	3.000	4.000	4.500	5.000	5.000

Variable	Rango	IQR	Modo	Número para moda	Asimetría	Curtosis
Disponibilidad de trabajo	1.000	1.000	4; 5	5	0.00	-2.57
Confianza	3.000	1.000	3	5	0.39	0.37
Facilidad de diálogo	2.000	1.250	4	5	0.17	-0.73
Herramienta de solución	2.000	1.000	5	5	-0.78	-0.15

Fuente: Elaboración propia.

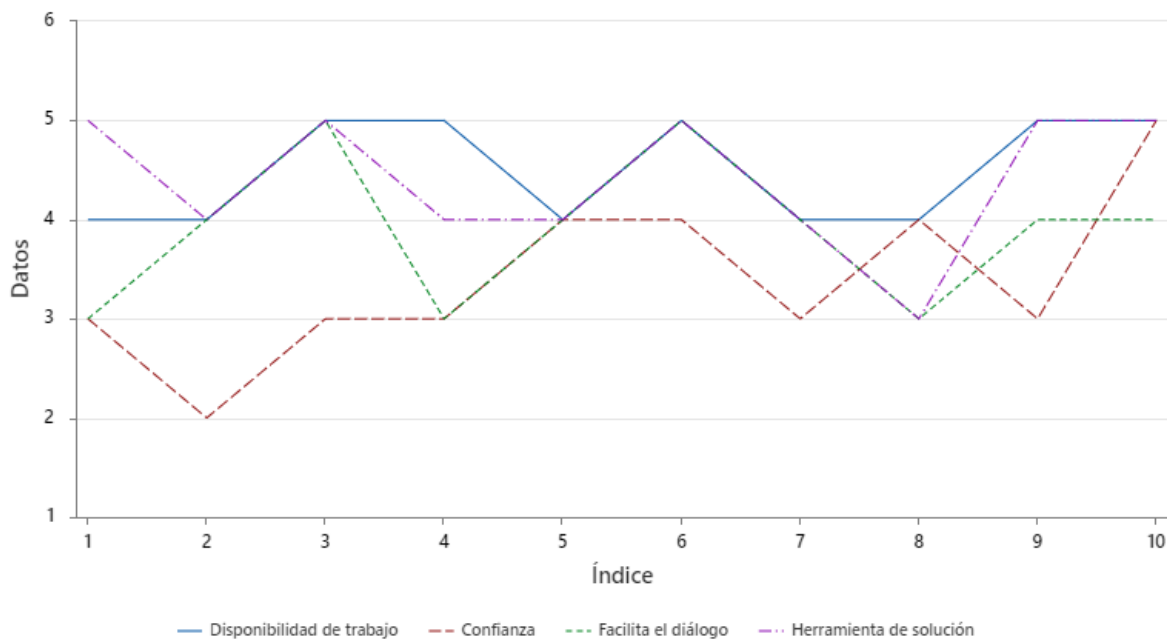


Ilustración 17 – Calificación de Likert de la encuesta de diagnóstico.

Fuente: Elaboración propia.

5.1.2 HEURÍSTICA DE USABILIDAD

Se llevó a cabo el análisis de las variables de investigación directamente relacionadas con las plataformas de los chatbots a través de una heurística de evaluación. Las variables de interés examinadas comprendieron: la visibilidad del estado del sistema, la correspondencia entre el sistema y el mundo real, el control y la libertad del usuario, la consistencia y normas, la prevención de errores, la ayuda y orientación, la flexibilidad y eficacia de uso, el diseño

estético, minimalista y atractivo, la ayuda a los usuarios a reconocer, diagnosticar y recuperarse de los errores, la conservación del contexto y la fiabilidad.

La evaluación heurística (Anexo 12) de cada una de estas variables permitió una comprensión profunda de la utilidad de cada uno de los chatbots para las tareas y actividades diarias del ingeniero clínico, basada en la experiencia de los profesionales.

5.1.2.1 Análisis comparativo de los agentes conversacionales

Con el objetivo de conocer el agente conversacional de mayor utilidad en las actividades diarias de un ingeniero clínico, se realizó un análisis comparativo de las respuestas otorgadas por la muestra de profesionales. Para esto, se estudió cada una de las variables de manera individual.

Para todas las variables se utilizó el mismo sistema métrico. Esto fue mediante una calificación del 1 al 5, cada número tenía un significado correspondiente. El 1 hacía referencia a "totalmente insatisfecho", el 2 significaba "insatisfecho", el 3 era "ni satisfecho ni insatisfecho", mientras que el 4 hacía mención de "satisfecho" y el 5 a "totalmente satisfecho".

Se mostraron los resultados para cada una de las once variables evaluadas en la heurística. Estos resultados incluyeron tanto el número total de respuestas válidas (N) como el porcentaje que representaban con respecto al total de la muestra. Además, se presentaron medidas estadísticas como la media, el error estándar de la media y la desviación estándar. También se proporcionaron valores como la varianza, el coeficiente de variación, la mediana, el valor mínimo y máximo, así como el primer y tercer cuartil para cada una de las variables. Junto a esto, se compartieron otros detalles como el rango, el rango intercuartil, el valor más frecuente o modo, el número de casos que coincidían con el modo, la asimetría y la curtosis de cada variable.

5.1.2.2 Visibilidad del estado del sistema

En primer lugar, se observan las medias de las evaluaciones no varían mucho entre los agentes, siendo 4.500 el valor para ChatGPT, Google Bard y ChatSonic. Únicamente Perplexity obtuvo un valor más alto de 4.800. Estos valores representan las calificaciones promedio asignadas por los participantes a la visibilidad del estado del sistema.

La precisión de estas medias es indicada por el error estándar de la media, que resulta más bajo para Perplexity (0.133) y 0.167 para el resto de chatbots. Al analizar las desviaciones estándar, se observa que son comparativamente bajas en todos los agentes, sugiriendo consistencia en las respuestas de los participantes.

En cuanto a la variabilidad y distribución, las varianzas son relativamente bajas en todos los casos, indicando que las calificaciones no están ampliamente dispersas alrededor de las medias. Esto se refuerza con el coeficiente de variación, que es similar en todos los agentes, lo que implica niveles comparables de dispersión relativa en relación con las medias. Al observar los valores mínimos, Q1, mediana, Q3 y máximo, se puede deducir que las calificaciones tienden a concentrarse entre los valores 4.000 y 5.000. La mediana de 4.500 para todos los chatbots a excepción de Perplexity AI (5.000) sugiere una inclinación hacia calificaciones cercanas al valor medio.

En el contexto de la moda y la asimetría-curtosis, se evidencia que Perplexity AI tiene un modo de 5, mientras que ChatGPT, Google Bard y ChatSonic muestran dos modos, 4 y 5, lo que representa que las dos son las calificaciones más frecuentes. En términos de asimetría, todos son valores de cero lo que indica una distribución simétrica. Excluyendo a Perplexity cuya asimetría es de -1.78, esto indica que hay una inclinación hacia calificaciones más altas en estas distribuciones. Por otro lado, las curtosis reflejan la forma de la distribución. Tres de los agentes conversacionales muestran una curtosis negativa (-2.57), señalando una distribución aplanada y una dispersión más amplia. En comparación al valor positivo (1.41) que indica que las colas de la distribución son más pesadas que las de una distribución normal, lo que significa que hay más valores extremos en la distribución.

En resumen, el análisis estadístico (Tabla 20) destaca a Perplexity como el agente con la media más alta en las evaluaciones de la visibilidad del estado del sistema, insinuando que los participantes perciben una mejor claridad en su funcionamiento. Este análisis cuantitativo proporciona una visión más profunda de cómo los participantes perciben la visibilidad de los diferentes agentes conversacionales (Ilustración 18).

Tabla 20 - Análisis estadístico comparativo de la visibilidad del sistema.

Variable	Cuento		Número		Porcentaje		Media	Error estándar de la media	Desviación estándar
	total	N	N*	acumulado	Porcentaje	acumulado			
ChatGPT	10	10	0	10	100	100	4.500	0.167	0.527
Bard	10	10	0	10	100	100	4.500	0.167	0.527
Perplexity	10	10	0	10	100	100	4.800	0.133	0.422
Chatsonic	10	10	0	10	100	100	4.500	0.167	0.527

Variable	Coeficiente									
	Varianza	de variación	Mínimo	Q1	Mediana	Q3	Máximo	Rango	IQR	Modo
ChatGPT	0.278	11.71	4.000	4.000	4.500	5.000	5.000	1.000	1.000	4; 5
Bard	0.278	11.71	4.000	4.000	4.500	5.000	5.000	1.000	1.000	4; 5
Perplexity	0.178	8.78	4.000	4.750	5.000	5.000	5.000	1.000	0.250	5

Coeficiente										
Variable	Varianza	de variación	Mínimo	Q1	Mediana	Q3	Máximo	Rango	IQR	Modo
Chatsonic	0.278	11.71	4.000	4.000	4.500	5.000	5.000	1.000	1.000	4; 5
Número										
Variable	para moda	Asimetría	Curtosis							
ChatGPT	5	0.00	-2.57							
Bard	5	0.00	-2.57							
Perplexity	8	-1.78	1.41							
Chatsonic	5	0.00	-2.57							

Fuente: Elaboración propia.

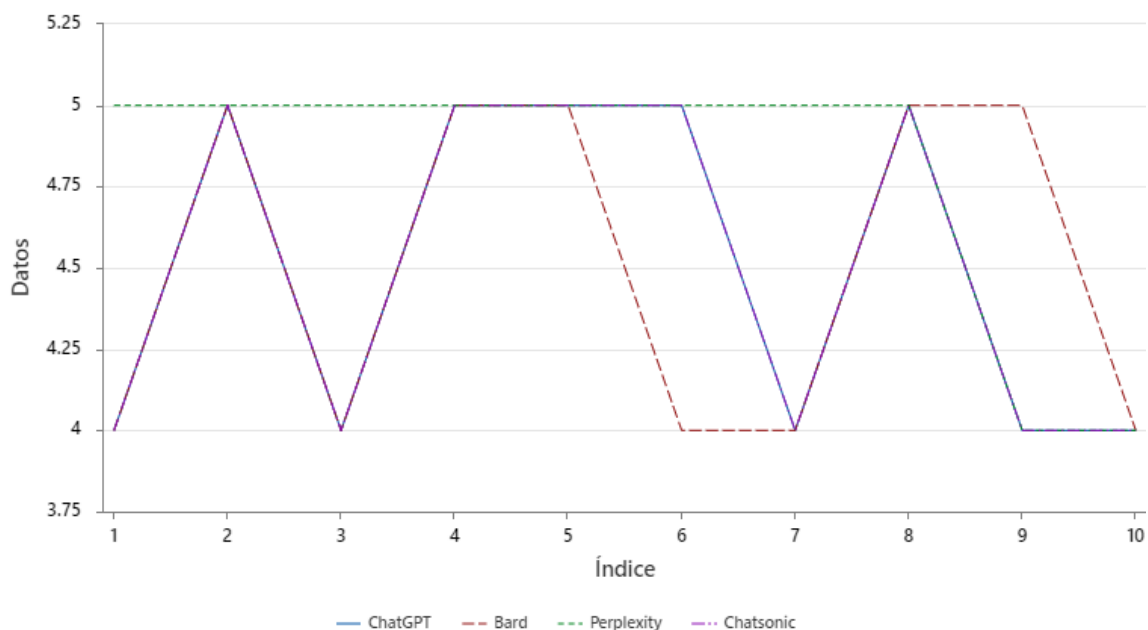


Ilustración 18 - Visibilidad del estado del sistema.

Fuente: Elaboración propia.

5.1.2.3 Correspondencia entre el sistema y el mundo real

En el contexto del análisis estadístico de las variables que exploran la correspondencia entre el sistema y el mundo real, se pueden extraer varias conclusiones. Cada una de las cuatro variables en estudio (ChatGPT, Bard, Perplexity y ChatSonic) se basa en un conjunto de diez observaciones, lo que abarca la totalidad del conjunto analizado (Ilustración 19).

La media, que representa el valor promedio de las puntuaciones, varía entre las variables. Google Bard y ChatSonic tienen una media de 3.800 y 3.700 respectivamente, mientras que ChatGPT se sitúa en 4.100 y Perplexity en 4.700. La desviación estándar, que cuantifica la dispersión de los datos respecto a la media, presenta diferencias notables, oscilando entre 0.483 y 0.949.

Las medidas de variabilidad también ofrecen perspectivas interesantes. La varianza, que refleja la extensión de los datos, tiende a ser similar entre ChatGPT, Bard y ChatSonic, aproximadamente en 0.767, mientras que para Perplexity es cercano a cero, indicando uniformidad en sus valores. El coeficiente de variación, que expresa la dispersión relativa en relación con la media, varía entre 10.28% y 25.64%. En cuanto al rango, que cuantifica la diferencia entre los valores máximo y mínimo, presenta un cambio de 2.000 para ChatGPT y Bard, de 3.000 para ChatSonic y Perplexity muestra un rango de 1.000.

Los valores de mediana son consistentes a lo largo de las variables (4.000), exceptuando a Perplexity (5.000). Respectivamente, un primer cuartil (Q1) de 3.000 y 4.000; y un tercer cuartil (Q3) de 5.000 para ChatGPT y Perplexity mientras que Bard y ChatSonic muestran un valor de 4.250. Respecto a propiedades de distribución, se aprecian diferencias en las asimetrías y curtosis. La asimetría, indicativa de la inclinación de la distribución, se sitúa en -0.22 para ChatGPT, 0.41 para Bard, -1.04 para Perplexity y -0.23 para ChatSonic. La curtosis, que revela las colas de la distribución, muestra una constancia en los valores negativos, con resultados de -1.73, -1.07, -1.22 y -0.35 para ChatGPT, Bard, Perplexity y ChatSonic respectivamente. Esto representa colas de la distribución más ligeras que las de una distribución normal, lo que significa que hay menos valores extremos en la distribución.

Adicionalmente, se destaca la variable Perplexity, cuya varianza presenta resultados cercanos a cero, lo que sugiere homogeneidad en todas las observaciones. En conjunto, estos resultados aportan una visión detallada sobre las características estadísticas de las variables analizadas, evidenciando similitudes y diferencias en términos de centralidad, dispersión y propiedades de distribución (Tabla 21).

Tabla 21 - Análisis estadístico comparativo de la correspondencia entre el sistema y el mundo real

Variable	Cuento		Número		Porcentaje		Media	Error estándar de la media	Desviación estándar
	total	N	N*	acumulado	Porcentaje	acumulado			
ChatGPT	10	10	0	10	100	100	4.100	0.277	0.876
Bard	10	10	0	10	100	100	3.800	0.249	0.789
Perplexity	10	10	0	10	100	100	4.700	0.153	0.483
ChatSonic	10	10	0	10	100	100	3.700	0.300	0.949

Variable	Coeficiente									
	Varianza	de variación	Mínimo	Q1	Mediana	Q3	Máximo	Rango	IQR	Modo
ChatGPT	0.767	21.36	3.000	3.000	4.000	5.000	5.000	2.000	2.000	5
Bard	0.622	20.76	3.000	3.000	4.000	4.250	5.000	2.000	1.250	3; 4
Perplexity	0.233	10.28	4.000	4.000	5.000	5.000	5.000	1.000	1.000	5

Coeficiente										
Variable	Varianza	de variación	Mínimo	Q1	Mediana	Q3	Máximo	Rango	IQR	Modo
ChatSonic	0.900	25.64	2.000	3.000	4.000	4.250	5.000	3.000	1.250	4

Número			
Variable	para moda	Asimetría	Curtosis
ChatGPT	4	-0.22	-1.73
Bard	4	0.41	-1.07
Perplexity	7	-1.04	-1.22
ChatSonic	4	-0.23	-0.35

Fuente: Elaboración propia.

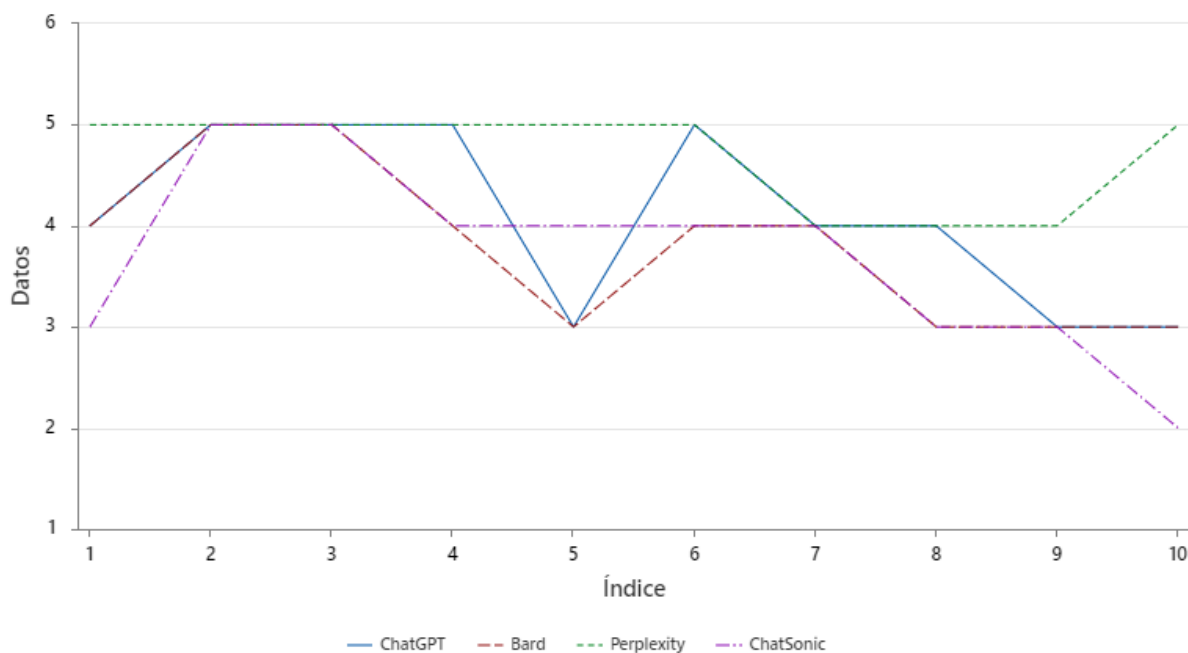


Ilustración 19 - Correspondencia entre el sistema y el mundo real.

Fuente: Elaboración propia.

5.1.2.4 Control y libertad del usuario.

En términos de medidas de tendencia central, la media se encuentra en un rango que va desde 3.800 hasta 4.500, siendo ChatSonic el que posee en valor más bajo y Perplexity su contrario. El error estándar de la media proporciona una estimación de la precisión de la media muestral y varía entre 0.163 y 0.291. La desviación estándar, que mide la dispersión de los datos alrededor de la media, se sitúa en torno a 0.516 y 0.919. La variabilidad de los datos se refleja en la varianza, que oscila entre 0.267 y 0.844 para las diferentes herramientas. El coeficiente de variación, que compara la variabilidad con la media, varía entre 11.74% y 24.18% (Tabla 22). En esta ocasión Google Bard es el poseedor de los valores estadísticos más bajos y ChatSonic de los más altos. Esto representa que hubo una mayor uniformidad en las respuestas de la muestra en cuanto a la calificación de Google Bard, mientras que el caso de ChatSonic es el opuesto.

En cuanto a las medidas que describen la distribución de los datos, se proporcionan valores como el mínimo, el primer cuartil (Q1), la mediana, el tercer cuartil (Q3) y el máximo. El rango, que es la diferencia entre el máximo y el mínimo, varía entre 1.000 y 3.000. El rango intercuartílico (IQR), que mide la dispersión de los valores alrededor de la mediana, se encuentra entre 1.000 y 1.250.

Se destaca que ChatGPT, Bard y ChatSonic tienen el mismo valor de moda (4), lo que sugiere que este valor es el más común en las distribuciones. Cabe resaltar que Perplexity posee una moda de 5. Además, las distribuciones de estas IA presentan asimetría negativa, indicando que se extienden más hacia la izquierda, a excepción de Google Bard que obtuvo una asimetría positiva indicando un sesgo hacia la derecha. Las medidas de curtosis son negativas para ChatGPT y Bard, indicando que las distribuciones son ligeramente menos puntiagudas que una distribución normal. En contraste se encuentran los valores positivos de curtosis de Perplexity AI y ChatSonic que indican distribuciones ligeramente más puntiagudas.

En resumen, el análisis estadístico proporciona una visión completa de las características de los instrumentos examinados (Ilustración 20), incluyendo medidas de tendencia central, dispersión y forma de la distribución. Estas medidas nos ayudan a comprender mejor la distribución y variabilidad de los datos en cada caso.

Tabla 22 - Análisis estadístico comparativo del control y la libertad del usuario

Variable	Conteo		Número		Porcentaje		Error estándar de la media	Desviación estándar	
	total	N	N*	acumulado	Porcentaje acumulado	Media			
ChatGPT	10	10	0	10	100	100	4.300	0.213	0.675
Bard	10	10	0	10	100	100	4.400	0.163	0.516
Perplexity	10	10	0	10	100	100	4.500	0.269	0.850
ChatSonic	10	10	0	10	100	100	3.800	0.291	0.919

Variable	Coeficiente									
	Varianza	de variación	Mínimo	Q1	Mediana	Q3	Máximo	Rango	IQR	Modo
ChatGPT	0.456	15.70	3.000	4.000	4.000	5.000	5.000	2.000	1.000	4
Bard	0.267	11.74	4.000	4.000	4.000	5.000	5.000	1.000	1.000	4
Perplexity	0.722	18.89	3.000	3.750	5.000	5.000	5.000	2.000	1.250	5
ChatSonic	0.844	24.18	2.000	3.000	4.000	4.250	5.000	3.000	1.250	4

Variable	Número para moda	Asimetría	Curtosis
ChatGPT	5	-0.43	-0.28
Bard	6	0.48	-2.28
Perplexity	7	-1.36	0.11
ChatSonic	5	-0.60	0.40

Fuente: Elaboración propia.

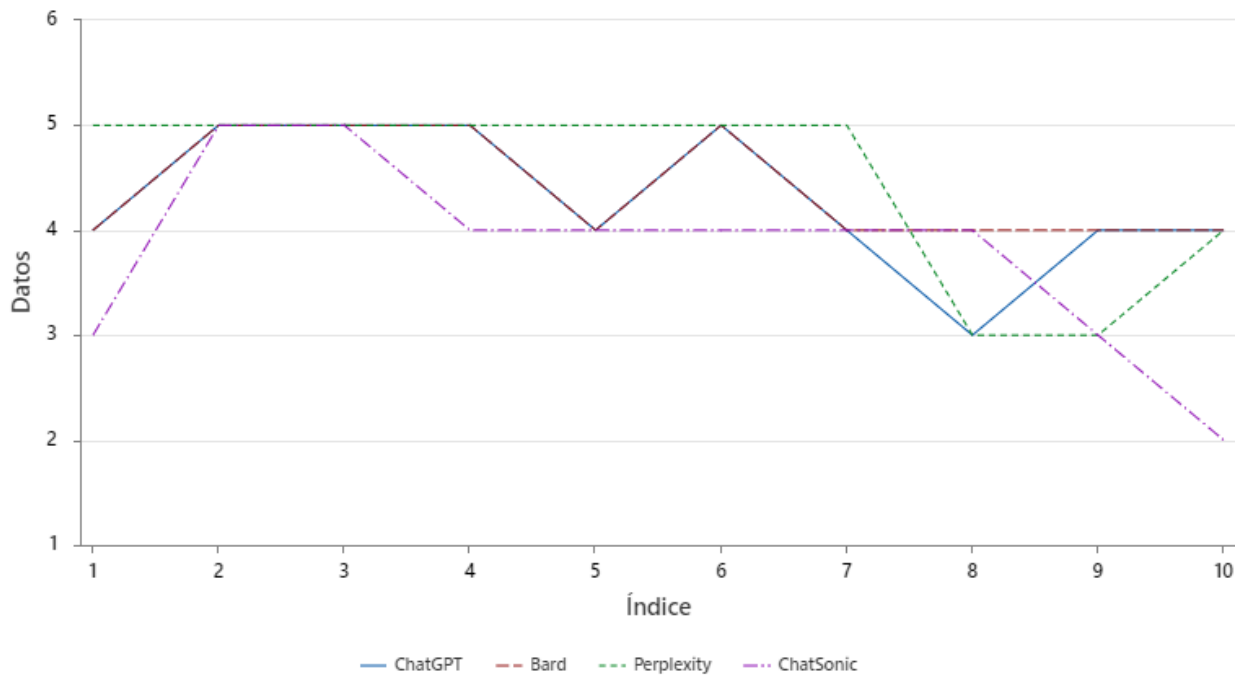


Ilustración 20 - Control y libertad del usuario.

Fuente: Elaboración propia.

5.1.2.5 Consistencia y normas.

El análisis comparativo (Tabla 23) de los datos proporcionados presenta una evaluación detallada de varias variables estadísticas relacionadas con cuatro diferentes agentes conversacionales: ChatGPT, Bard, Perplexity y ChatSonic (Ilustración 21). Se observa que la media, como medida de tendencia central, varía ligeramente entre los agentes. ChatGPT y ChatSonic tienen la misma media (3.900), mientras que Perplexity presenta una media ligeramente más alta (4.500), y Bard tiene una media ligeramente más baja (4.100). En cuanto al error estándar de la media, se observa que Bard posee el valor más bajo (0.568), mientras que ChatSonic tiene el valor más alto (0.994), indicando variabilidad en torno a las medias.

La desviación estándar, que mide la dispersión de los datos respecto a la media, exhibe variaciones entre los agentes. ChatSonic tiene la desviación estándar más alta (0.994), mientras que Perplexity muestra la más baja (0.568). El coeficiente de variación, que compara la desviación estándar con la media, indica que ChatSonic presenta el mayor grado de variabilidad relativa (25.50%).

Los valores de la mediana (Q2) son relativamente consistentes, siendo 4.000 para ChatGPT, Bard y ChatSonic, mientras que Perplexity registra una mediana de 5.000. Los valores de rango y rango intercuartil (IQR) revelan las diferencias en la extensión de los datos. Se nota que

Google Bard obtiene el valor más bajo en el rango intercuartil (0.250), lo que implica una concentración más estrecha de valores en la región intercuartil.

El modo, representando el valor más frecuente, es 5 para Perplexity, 4 para Bard y ChatSonic y de 3 para ChatGPT. La asimetría, indicador de la dirección y grado de asimetría de la distribución, es negativa Perplexity y ChatSonic, siendo más pronunciada para el primero (-1.18 y -0.61 respectivamente). Por otro lado, ChatGPT y Bard tienen valores positivos (0.22 y 0.09), teniendo un valor de casi cero lo que demuestra una simetría casi perfecta. La curtosis, que caracteriza la forma de la distribución, muestra variaciones notables. Bard exhibe una curtosis de 1.50, indicando una distribución puntiaguda y con colas pesadas, mientras que ChatGPT tiene una curtosis negativa (-1.73), señalando una distribución relativamente aplanada.

Tabla 23 - Análisis estadístico comparativo de la consistencia y normas.

Variable	Conteo		Número		Porcentaje		Error estándar de la media	Desviación estándar	
	total	N	N*	acumulado	Porcentaje acumulado	Media			
ChatGPT	10	10	0	10	100	100	3.900	0.277	0.876
Bard	10	10	0	10	100	100	4.100	0.180	0.568
Perplexity	10	10	0	10	100	100	4.500	0.224	0.707
ChatSonic	10	10	0	10	100	100	3.900	0.314	0.994

Variable	Coeficiente									
	Varianza	de variación	Mínimo	Q1	Mediana	Q3	Máximo	Rango	IQR	Modo
ChatGPT	0.767	22.45	3.000	3.000	4.000	5.000	5.000	2.000	2.000	3
Bard	0.322	13.85	3.000	4.000	4.000	4.250	5.000	2.000	0.250	4
Perplexity	0.500	15.71	3.000	4.000	5.000	5.000	5.000	2.000	1.000	5
ChatSonic	0.989	25.50	2.000	3.000	4.000	5.000	5.000	3.000	2.000	4

Variable	Número		
	para moda	Asimetría	Curtosis
ChatGPT	4	0.22	-1.73
Bard	7	0.09	1.50
Perplexity	6	-1.18	0.57
ChatSonic	4	-0.61	-0.16

Fuente: Elaboración propia.

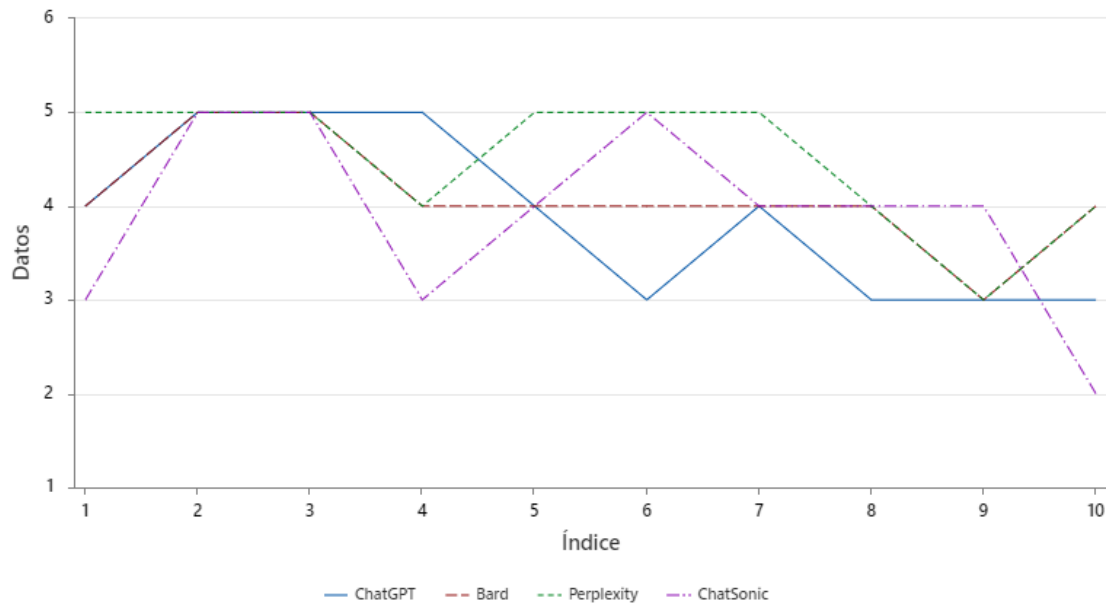


Ilustración 21 - Consistencia y normas.

Fuente: Elaboración propia.

5.1.2.6 Prevención de errores

La media de las variables varía ligeramente, oscilando entre 3.300 y 3.800. En este sentido, Perplexity y ChatSonic presentan las medias más altas, lo que sugiere que tienen un rendimiento ligeramente superior en la prevención de errores en comparación con las otras variables.

Las desviaciones estándar son diferentes entre las variables, con valores que van desde 0.675 hasta 0.966. ChatSonic destaca por tener la desviación estándar más alta, lo que indica una mayor dispersión de los datos en relación con su media. El coeficiente de variación, que proporciona una medida de la variabilidad relativa en relación con la media, es más alto para ChatSonic (26.84%), lo que sugiere una mayor variabilidad en comparación con las demás variables.

En términos de distribución, todas las variables muestran una asimetría cercana a cero, lo que sugiere que las distribuciones son relativamente simétricas y no están sesgadas de manera significativa. Excluyendo a ChatGPT (1.18) que posee un sesgo más evidente hacia la derecha. La curtosis, que evalúa la forma de la distribución en relación con la distribución normal, es negativa para todas las variables. Esto indica que las distribuciones son ligeramente aplanadas en comparación con una distribución normal estándar. Asimismo, la excepción es ChatGPT con un valor positivo (0.57) que representa una distribución ligeramente más puntiaguda que el resto.

En relación con las medidas de posición, como los cuartiles y la mediana, las variables exhiben patrones similares. La moda más común para todas las variables es 3, lo que significa que este valor es el más frecuente en cada conjunto de datos. En términos de rango y rango intercuartílico (IQR), las variables ChatGPT y Google Bard muestran valores idénticos, mientras que Perplexity y ChatSonic tienen un rango ligeramente más amplio y un IQR más grande.

En resumen, el análisis comparativo (Tabla 24) revela que las variables ChatGPT, Bard, Perplexity y ChatSonic tienen similitudes en términos de medias y distribución, pero difieren en aspectos como la variabilidad, donde ChatSonic muestra una mayor dispersión. A pesar de estas diferencias, todas las variables parecen mostrar un comportamiento razonablemente similar en cuanto a su capacidad de prevención de errores, con la moda 3 como un punto común en sus distribuciones (Ilustración 22).

Tabla 24 - Análisis estadístico comparativo de la prevención de errores.

Variable	Conteo		Número		Porcentaje		Error estándar de la media	Desviación estándar	
	total	N	N*	acumulado	Porcentaje acumulado	Media			
ChatGPT	10	10	0	10	100	100	3.500	0.224	0.707
Bard	10	10	0	10	100	100	3.300	0.213	0.675
Perplexity	10	10	0	10	100	100	3.800	0.291	0.919
ChatSonic	10	10	0	10	100	100	3.600	0.306	0.966

Variable	Coeficiente									
	Varianza	de variación	Mínimo	Q1	Mediana	Q3	Máximo	Rango	IQR	Modo
ChatGPT	0.500	20.20	3.000	3.000	3.000	4.000	5.000	2.000	1.000	3
Bard	0.456	20.45	2.000	3.000	3.000	4.000	4.000	2.000	1.000	3
Perplexity	0.844	24.18	3.000	3.000	3.500	5.000	5.000	2.000	2.000	3
ChatSonic	0.933	26.84	2.000	3.000	3.500	4.250	5.000	3.000	1.250	3

Variable	Número		
	para moda	Asimetría	Curtosis
ChatGPT	6	1.18	0.57
Bard	5	-0.43	-0.28
Perplexity	5	0.47	-1.81
ChatSonic	4	0.11	-0.62

Fuente: Elaboración propia.

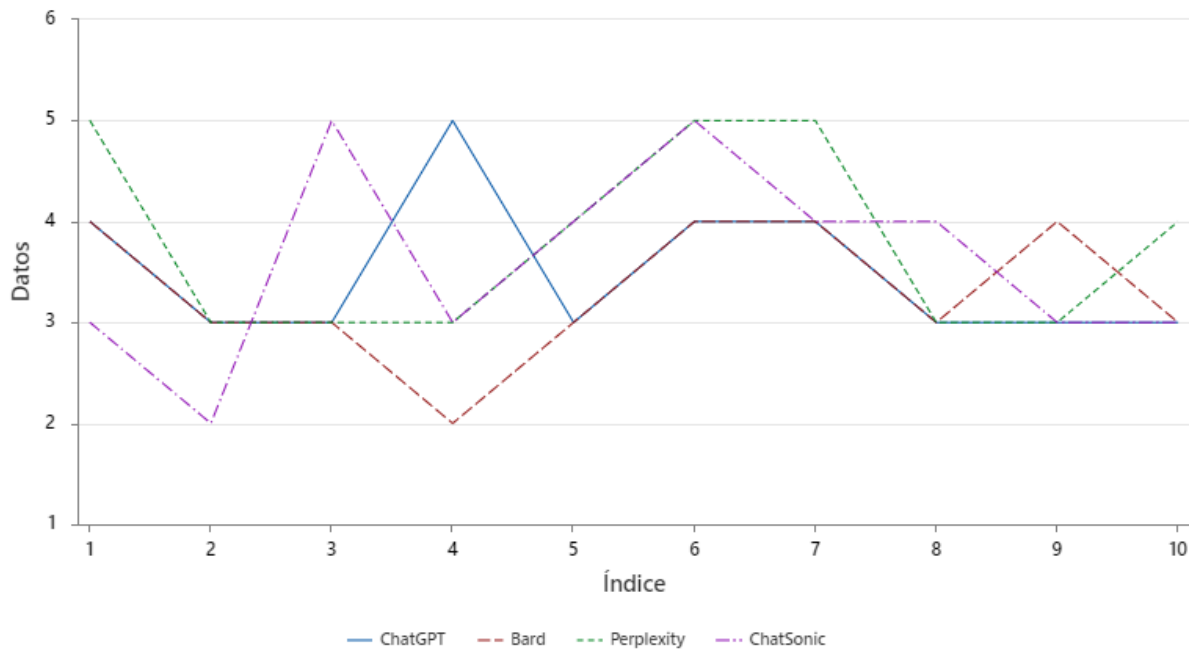


Ilustración 22- Prevención de errores.

Fuente: Elaboración propia.

5.1.2.7 Ayuda y orientación

La media es un indicador clave, todas las variables tienen una media igual a 3.500, excepto Perplexity, que tiene una media de 4.500. Esto sugiere que, en promedio, los valores son similares en ChatGPT, Bard y ChatSonic, mientras que Perplexity tiene un promedio ligeramente más alto.

En términos de dispersión de datos, las desviaciones estándar proporcionan información sobre cuán dispersos están los datos respecto a la media. En este caso, ChatGPT y ChatSonic tienen desviaciones estándar más altas, lo que indica una mayor variabilidad en comparación con Bard y Perplexity, que tienen desviaciones estándar más bajas.

La varianza es otra medida de la dispersión, y Perplexity tiene la varianza más alta (0.700), seguida de ChatSonic (1.100), ChatGPT (1.500) y Google Bard (0.300). Esto significa que Perplexity muestra la mayor variabilidad en sus datos. El coeficiente de variación es una medida relativa de la dispersión en relación con la media. Perplexity tiene el coeficiente de variación más alto (18.59%), lo que indica una mayor variabilidad relativa en comparación con las otras variables.

En cuanto a la distribución de los datos, el rango (diferencia entre el valor máximo y mínimo) es mayor para Perplexity (3.000), seguido de cerca por ChatGPT (2.250), mientras que Bard y ChatSonic tienen rangos más pequeños. El rango intercuartílico (IQR), que mide la

dispersión en el intervalo intercuartílico, varía de mayor a menor en el siguiente orden: ChatGPT (2.250), Bard (1.000), Perplexity (1.250) y ChatSonic (1.500).

Todos los valores de asimetría son relativamente pequeños y cercanos a cero, lo que sugiere que las distribuciones son aproximadamente simétricas. En cuanto a la curtosis, que mide la forma de la distribución, todas las variables parecen tener distribuciones dentro del rango de una distribución normal, con Perplexity mostrando una ligera puntiagudez en comparación con las otras variables.

En resumen, estas estadísticas (Tabla 25) proporcionan una visión detallada de los cuatro agentes conversacionales y sus características. Muestran cómo se distribuyen los datos, su dispersión y otras características importantes de su comportamiento estadístico (Ilustración 23).

Tabla 25 - Análisis estadístico comparativo sobre la ayuda y orientación.

Variable	Conteo		Número		Porcentaje		Error estándar de la media	Desviación estándar	
	total	N	N*	acumulado	Porcentaje acumulado	Media			
ChatGPT	10	10	0	10	100	100	3.700	0.300	0.949
Bard	10	10	0	10	100	100	3.600	0.163	0.516
Perplexity	10	10	0	10	100	100	4.400	0.221	0.699
ChatSonic	10	10	0	10	100	100	3.500	0.307	0.972

Variable	Coeficiente									
	Varianza	de variación	Mínimo	Q1	Mediana	Q3	Máximo	Rango	IQR	Modo
ChatGPT	0.900	25.64	2.000	3.000	4.000	4.250	5.000	3.000	1.250	3
Bard	0.267	14.34	3.000	3.000	4.000	4.000	4.000	1.000	1.000	3
Perplexity	0.489	15.89	3.000	4.000	4.500	5.000	5.000	2.000	1.000	3
ChatSonic	0.944	27.77	2.000	2.750	4.000	4.000	5.000	3.000	1.250	3

Variable	Número		
	para moda	Asimetría	Curtosis
ChatGPT	4	4	-0.23
Bard	4	6	-0.48
Perplexity	5	5	-0.78
ChatSonic	4	5	-0.45

Fuente: Elaboración propia.

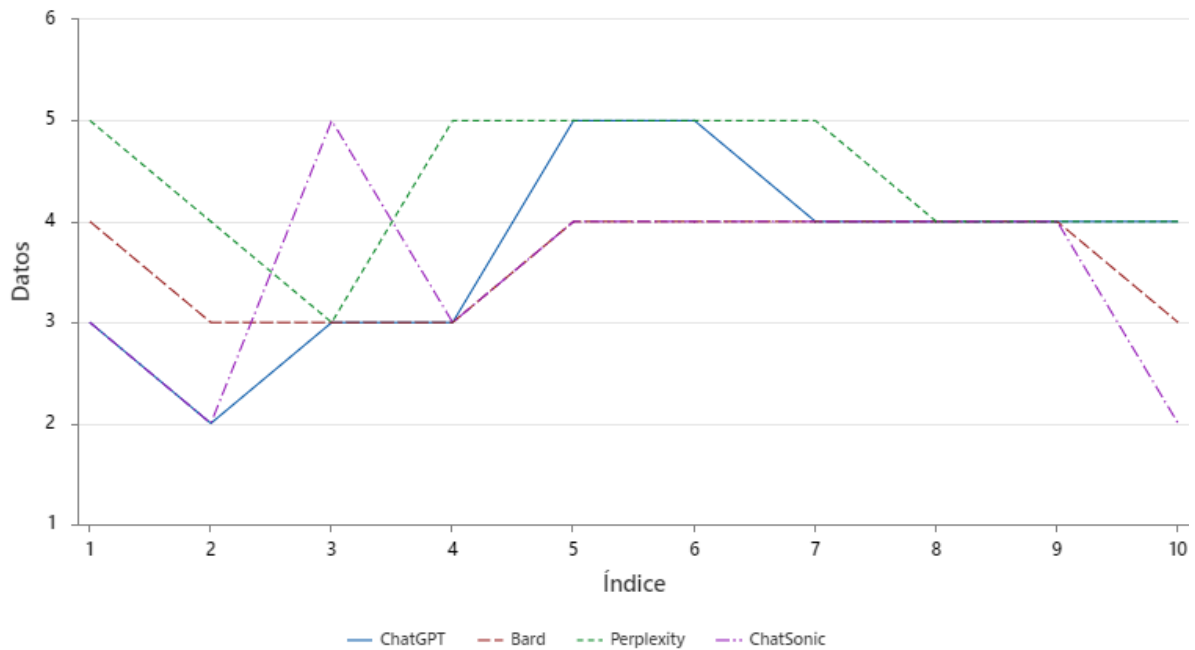


Ilustración 23 - Ayuda y orientación.

Fuente: Elaboración propia.

5.1.2.8 Flexibilidad y eficacia de uso

En el análisis estadístico de la flexibilidad y eficacia de uso (Tabla 26), las medias varían ligeramente, siendo 3.500 para ChatGPT, Google Bard y ChatSonic, y 4.500 para Perplexity. Estos valores promedio están acompañados por errores estándar que oscilan entre 0.224 y 0.500, y desviaciones estándar que van desde 0.548 hasta 1.225.

En términos de variabilidad, las variables también difieren. La varianza para ChatGPT es 1.500, para Bard es 0.300, para Perplexity es 0.700 y para ChatSonic es 1.100. El coeficiente de variación, que mide la relación entre la desviación estándar y la media, muestra que "ChatGPT" tiene la mayor variabilidad relativa, con un 34.99%, mientras que las otras variables tienen valores más bajos, entre 15.65% y 29.97%.

Los cuartiles y rangos también brindan información valiosa. Los valores mínimos son consistentes a través de todas las variables, siendo 2.000. Los primeros cuartiles (Q1) varían entre 2.750 y 3.750, mientras que las medianas (Q2) se sitúan en 3.000, 3.500 y 5.000 respectivamente. Los terceros cuartiles (Q3) oscilan entre 4.000 y 5.000. Los valores máximos para todas las variables son 5.000. Los rangos, que reflejan la amplitud entre el mínimo y el máximo, se sitúan en un rango de 1.000 a 3.000.

En términos de forma de distribución, la asimetría varía de -1.54 a 0.49. Perplexity muestra una asimetría negativa, lo que sugiere que tiene una cola larga en el lado izquierdo de la

distribución. La curtosis, que refleja la forma de la distribución en comparación con una distribución normal, varía de -3.33 a 1.43. Bard tiene una curtosis negativa, lo que implica una distribución más achatada que la normal (Ilustración 24).

Por último, el análisis de moda revela que en ChatGPT la moda es 3, mientras que para Perplexity es 5. En el caso de ChatSonic y Google Bard, hay dos modas: 3 y 4. Esto significa que ambos valores son igualmente frecuentes.

Tabla 26 - Análisis estadístico comparativo sobre la flexibilidad y eficacia de uso.

Variable	Conteo total	N	N*	Número		Porcentaje		Error estándar de la media	Desviación estándar
				acumulado	Porcentaje acumulado	acumulado	Media		
ChatGPT	10	10	0	10	100	100	4.300	0.213	0.675
Bard	10	10	0	10	100	100	3.800	0.133	0.422
Perplexity	10	10	0	10	100	100	4.500	0.167	0.527
ChatSonic	10	10	0	10	100	100	3.400	0.340	1.075

Variable	Varianza	Coefficiente de variación	Mínimo	Q1	Mediana	Q3	Máximo	Rango	IQR	Modo
ChatGPT	0.456	15.70	3.000	4.000	4.000	5.000	5.000	2.000	1.000	4
Bard	0.178	11.10	3.000	3.750	4.000	4.000	4.000	1.000	0.250	4
Perplexity	0.278	11.71	4.000	4.000	4.500	5.000	5.000	1.000	1.000	4; 5
ChatSonic	1.156	31.62	2.000	2.750	3.000	4.250	5.000	3.000	1.500	3

Variable	Número para moda	Asimetría	Curtosis
ChatGPT	5	-0.43	-0.28
Bard	8	-1.78	1.41
Perplexity	5	-0.00	-2.57
ChatSonic	4	0.32	-0.88

Fuente: Elaboración propia.

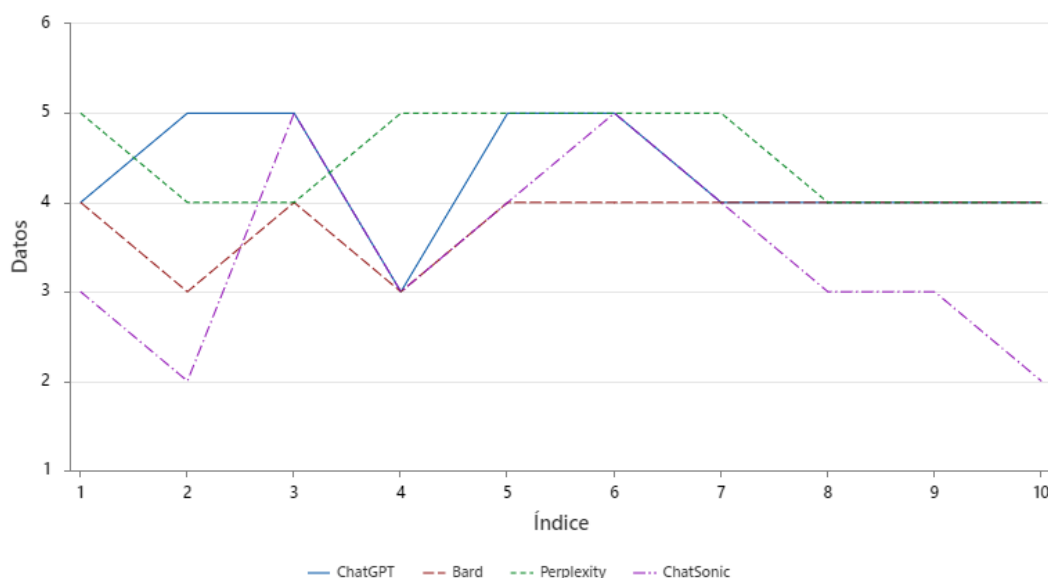


Ilustración 24 - Flexibilidad y eficacia de uso.

Fuente: Elaboración propia.

5.1.2.9 Diseño estético, minimalista y atractivo

Se realizó un análisis estadístico comparativo del diseño estético, minimalista y atractivo utilizando los datos recopilados (Tabla 27). Los valores clave incluyen la media y la desviación estándar de las herramientas evaluadas: ChatGPT, Bard, Perplexity y ChatSonic. Las medias muestran ligeras diferencias entre las variables, con valores promedio de 4.333 para ChatGPT, 3.667 para Google Bard, 4.667 para Perplexity y 3.833 para ChatSonic. Las desviaciones estándar son consistentes en todas las variables, lo que indica que las respuestas de los evaluadores están relativamente uniformes.

Además, se ha considerado el coeficiente de variación, que relaciona la desviación estándar con la media. Esta medida muestra un nivel similar de dispersión relativa en todas las variables. Se ha explorado la mediana y el rango intercuartílico (IQR) como indicadores de la tendencia central y la dispersión de los valores más representativos. La primera oscila entre valores de 3.500 a 5.000, y la segunda entre valores de 0.500 a 2.250.

El análisis de asimetría y curtosis ofrece información sobre la forma y simetría de las distribuciones. La variable Perplexity presenta la asimetría más negativa, lo que sugiere una inclinación hacia valores inferiores. En términos de curtosis, Perplexity se destaca con una distribución más apuntada, mientras que ChatGPT y Bard presentan una distribución con el menor apuntamiento.

En última instancia, estos resultados indican que las evaluaciones de diseño estético, minimalista y atractivo tienden a ser similares en términos de medias y desviaciones estándar. Sin embargo, existen diferencias en la forma de las distribuciones y la dispersión de los valores (Ilustración 25). Es fundamental recordar que el diseño estético y atractivo puede ser subjetivo, lo que agrega un matiz adicional a la interpretación de estos resultados numéricos.

Tabla 27 - Análisis estadístico comparativo sobre el diseño.

Variable	Conteo			Numero acumulado	Porcentaje			Media	Error estándar de la media	Desviación estándar
	total	N	N*		Porcentaje	acumulado	Media			
ChatGPT	10	10	0	10	100	100	4.400	0.221	0.699	
Bard	10	10	0	10	100	100	3.800	0.200	0.632	
Perplexity	10	10	0	10	100	100	4.500	0.269	0.850	
ChatSonic	10	10	0	10	100	100	3.600	0.340	1.075	

Variable	Coeficiente									
	Varianza	de variación	Mínimo	Q1	Mediana	Q3	Máximo	Rango	IQR	Modo
ChatGPT	0.489	15.89	3.000	4.000	4.500	5.000	5.000	2.000	1.000	5
Bard	0.400	16.64	3.000	3.000	4.000	4.000	5.000	2.000	1.000	4
Perplexity	0.722	18.89	3.000	3.750	5.000	5.000	5.000	2.000	1.250	5

Coeficiente										
Variable	Varianza	de variación	Mínimo	Q1	Mediana	Q3	Máximo	Rango	IQR	Modo
ChatSonic	1.156	29.86	2.000	2.750	4.000	4.250	5.000	3.000	1.500	4

Número			
Variable	para moda	Asimetría	Curtosis
ChatGPT	5	-0.78	-0.15
Bard	6	0.13	0.18
Perplexity	7	-1.36	0.11
ChatSonic	4	-0.32	-0.88

Fuente: Elaboración propia.

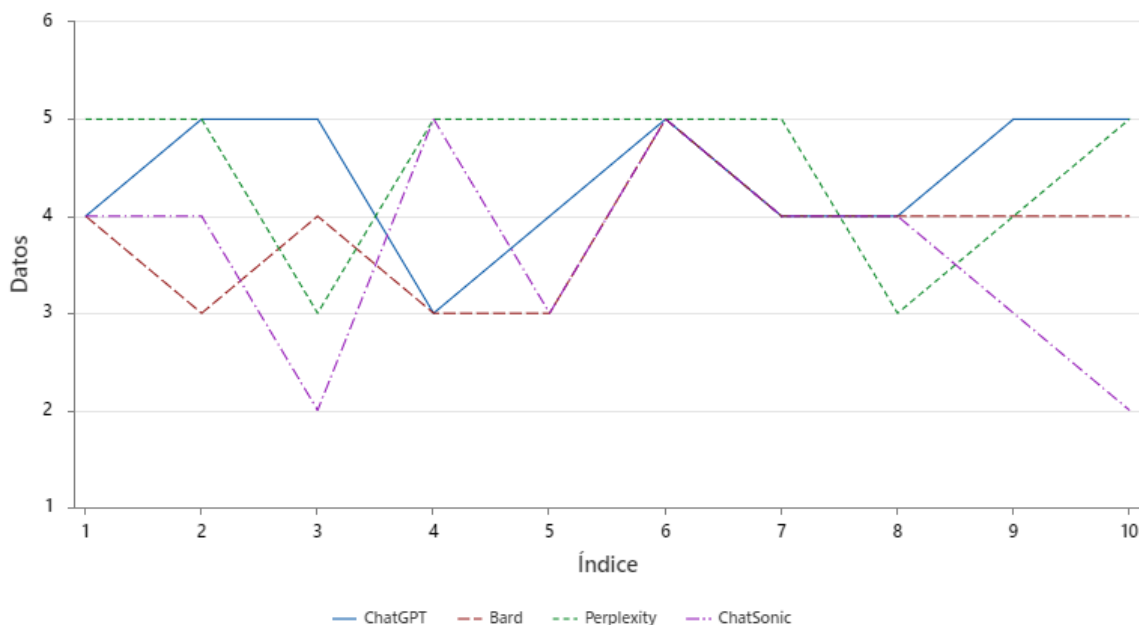


Ilustración 25 - Diseño estético, minimalista y atractivo.

Fuente: Elaboración propia.

5.1.2.10 Ayudar a los usuarios a reconocer, diagnosticar y recuperarse de los errores.

En este análisis (Tabla 28) se examinaron varias medidas estadísticas para evaluar la capacidad de los agentes conversacionales para ayudar a los usuarios a lidiar con errores. En términos de medidas centrales, se encontró que los métodos ChatGPT, Bard y ChatSonic tienen medias similares alrededor de 3.67, mientras que Perplexity tiene una media ligeramente mayor, de 4.00. Esto sugiere que, en promedio, los métodos se desempeñan de manera comparable en la tarea en cuestión, provocando un resultado satisfactorio. Sin embargo, se observó que el error estándar de la media es menor para Bard en comparación con los otros métodos, lo que implica una mayor consistencia en las puntuaciones.

En cuanto a la dispersión de los datos, se analizó la desviación estándar y la varianza. Se observó que ChatSonic y ChatGPT tienen las desviaciones estándar más altas (1.049 y 1.033,

respectivamente), lo que indica una mayor variabilidad en las puntuaciones. Por otro lado, Bard y Perplexity tienen desviaciones estándar menores (0.816 y 0.894, respectivamente), lo que sugiere una mayor cohesión en sus resultados. Las varianzas siguen un patrón similar.

La distribución y la tendencia central también se examinaron a través de la mediana y el rango. La mediana, que representa el valor medio en un conjunto de datos ordenados, fue constante en 4.000 para ChatGPT y Perplexity, y ligeramente más baja en 3.500 para ChatSonic y Bard. El rango, que indica la diferencia entre los valores máximo y mínimo, fue mayor para ChatGPT y ChatSonic (3.000), y menor para Bard y Perplexity (2.000).

La forma de la distribución fue evaluada mediante la asimetría y la curtosis. Se encontró que ChatGPT tiene una asimetría negativa, lo que sugiere una distribución sesgada hacia la izquierda, mientras que Bard presenta una asimetría positiva, indicando un sesgo hacia la derecha. Perplexity y ChatSonic tienen asimetría cercana a cero, lo que indica distribuciones más simétricas. La curtosis también varía, siendo positiva para ChatGPT, lo que implica distribuciones más pesadas en las colas en comparación con una distribución normal. Por otro lado, Bard, Perplexity y ChatSonic muestran valores de curtosis más bajos, sugiriendo formas de distribución más cercanas a la normalidad.

En resumen, el análisis estadístico (Ilustración 26) revela que los métodos difieren en términos de dispersión, tendencia central y forma de distribución en relación con su capacidad para ayudar a los usuarios a enfrentar y resolver errores. Estas medidas proporcionan una visión más detallada de cómo se desempeñan los métodos en la tarea específica y pueden ayudar a tomar decisiones informadas sobre su implementación.

Tabla 28 - Análisis estadístico comparativo sobre la ayuda a los usuarios para lidiar con errores.

Variable	Conteo		Número		Porcentaje		Error estándar de la media	Desviación Estándar	
	total	N	N*	acumulado	Porcentaje	acumulado			Media
ChatGPT	10	10	0	10	100	100	3.800	0.249	0.789
Bard	10	10	0	10	100	100	3.600	0.221	0.699
Perplexity	10	10	0	10	100	100	3.900	0.277	0.876
ChatSonic	10	10	0	10	100	100	3.300	0.300	0.949

Variable	Coeficiente									
	Varianza	de variación	Mínimo	Q1	Mediana	Q3	Máximo	Rango	IQR	Modo
ChatGPT	0.622	20.76	2.000	3.750	4.000	4.000	5.000	3.000	0.250	4
Bard	0.489	19.42	3.000	3.000	3.500	4.000	5.000	2.000	1.000	3
Perplexity	0.767	22.45	3.000	3.000	4.000	5.000	5.000	2.000	2.000	3
ChatSonic	0.900	28.75	2.000	2.750	3.000	4.000	5.000	3.000	1.250	3

Variable	Número para moda	Asimetría	Curtosis
ChatGPT	7	-1.29	2.98
Bard	5	0.78	-0.15
Perplexity	4	0.22	-1.73
ChatSonic	4	0.23	-0.35

Fuente: Elaboración propia.

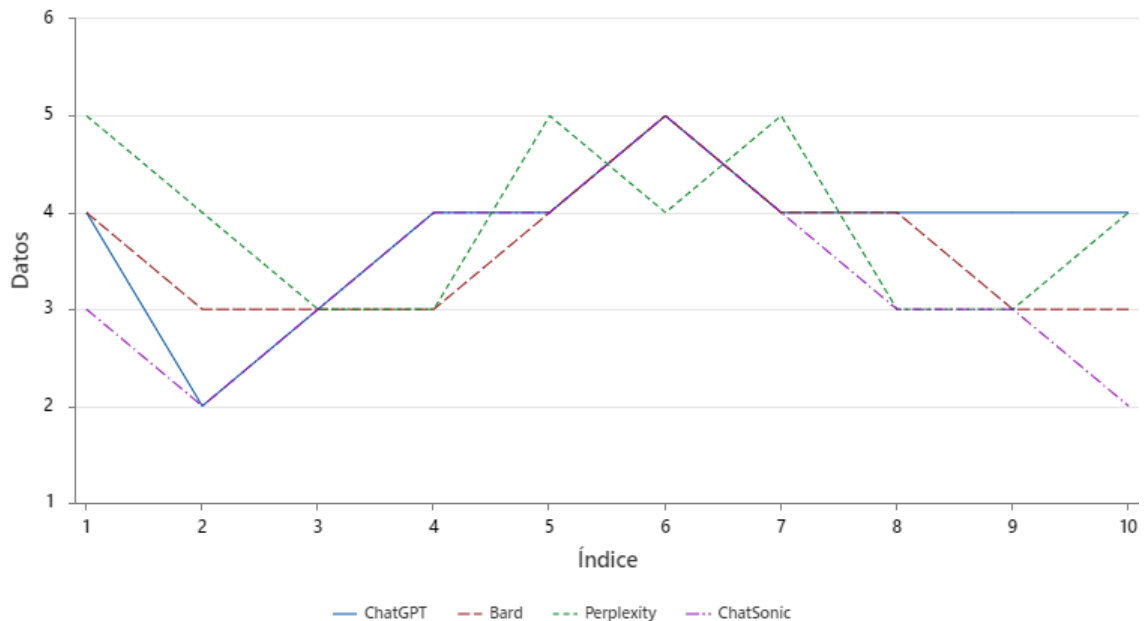


Ilustración 26 - Ayudar a los usuarios a reconocer, diagnosticar y recuperarse de los errores.

Fuente: Elaboración propia.

5.1.2.11 Conservación del contexto

Se calcularon diversas medidas estadísticas (Tabla 29) que ofrecen una visión completa de la distribución y la tendencia central de los datos (Ilustración 27).

Al examinar las medidas de tendencia central, se observa que las medias varían ligeramente entre las diferentes variables. ChatGPT tiene una media de aproximadamente 4.333, Bard tiene una media de alrededor de 3.833, Perplexity tiene una media de alrededor de 4.500, y ChatSonic tiene una media de alrededor de 4.167. Estos valores indican el valor promedio en cada conjunto de datos. Siendo estos resultados satisfactorios o totalmente satisfactorios para los usuarios.

En términos de dispersión, la desviación estándar es una medida clave que indica cuánto se dispersan los valores individuales alrededor de la media. Las desviaciones estándar varían

entre 0.211 (ChatGPT), 0.401 (Bard), 0.224 (Perplexity) y 0.307 (ChatSonic). Estos valores revelan la variabilidad de los datos en relación con la media.

La distribución de los datos también se examina a través de medidas como la asimetría y la curtosis. La asimetría proporciona información sobre la simetría de la distribución, donde valores positivos indican una cola hacia la derecha y valores negativos indican una cola hacia la izquierda. Por otro lado, la curtosis mide la "puntiagudez" de la distribución en comparación con una distribución normal. Los valores de asimetría varían entre 0.97 (ChatGPT), -1.44 (Bard), 0.00 (Perplexity) y -0.31 (ChatSonic), mientras que los valores de curtosis son -1.87 (ChatGPT), 3.60 (Bard), -3.33 (Perplexity) y -0.10 (ChatSonic).

Tabla 29 - Análisis estadístico comparativo sobre la conservación del contexto.

Variable	Conteo		Numero		Porcentaje		Error estándar de la media	Desviación estándar	
	total	N	N*	acumulado	Porcentaje	acumulado			Media
ChatGPT	10	10	0	10	100	100	4.100	0.180	0.568
Bard	10	10	0	10	100	100	3.700	0.300	0.949
Perplexity	10	10	0	10	100	100	4.400	0.163	0.516
ChatSonic	10	10	0	10	100	100	3.900	0.277	0.876

Variable	Coeficiente									
	Varianza	de variación	Mínimo	Q1	Mediana	Q3	Máximo	Rango	IQR	Modo
ChatGPT	0.322	13.85	3.000	4.000	4.000	4.250	5.000	2.000	0.250	4
Bard	0.900	25.64	2.000	3.500	4.000	4.000	5.000	3.000	0.500	4
Perplexity	0.267	11.74	4.000	4.000	4.000	5.000	5.000	1.000	1.000	4
ChatSonic	0.767	22.45	2.000	3.750	4.000	4.250	5.000	3.000	0.500	4

Variable	Número		
	para moda	Asimetría	Curtosis
ChatGPT	7	0.09	1.50
Bard	7	-1.21	0.95
Perplexity	6	0.48	-2.28
ChatSonic	6	-1.02	1.83

Fuente: Elaboración propia.

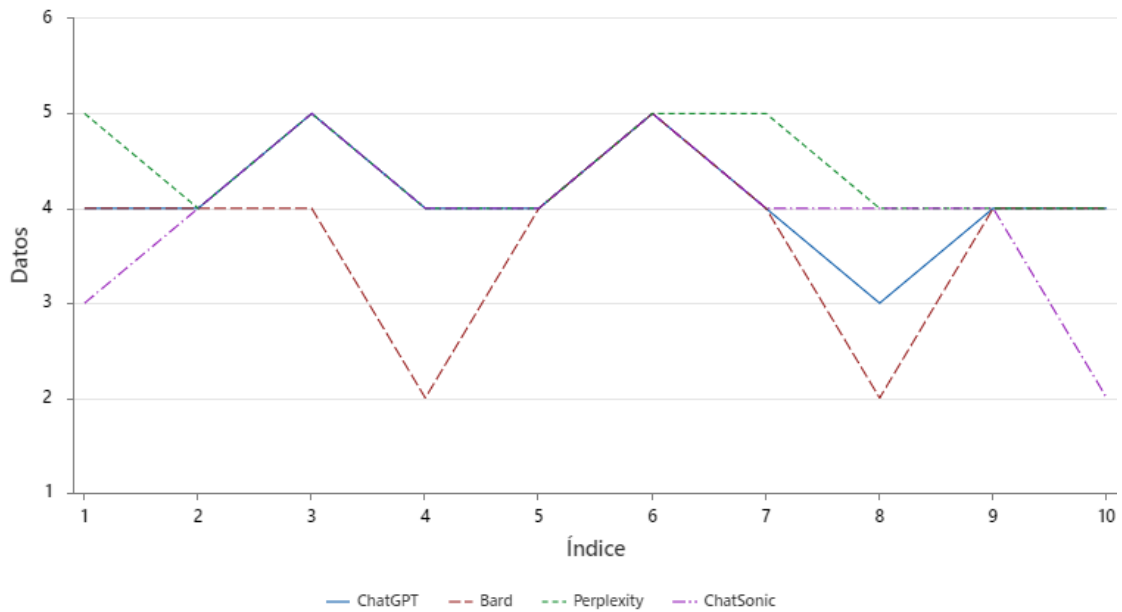


Ilustración 27 - Conservación del contexto.

Fuente: Elaboración propia.

5.1.2.12 *Fiabilidad*

Se llevó a cabo un análisis estadístico de los cuatro chatbots (Ilustración 28). Para cada variable, se calcularon varias medidas de tendencia central y dispersión (Tabla 30). En términos de medidas de tendencia central, la variable ChatGPT tiene una media de 3.000, mientras que Bard tiene una media ligeramente inferior de 2.667. Por su parte, Perplexity y ChatSonic presentan medias de 4.000 y 3.667, respectivamente. Los errores estándar asociados con estas medias son relativamente pequeños, con valores entre 0.258 y 0.422.

En cuanto a la dispersión de los datos, las variables ChatGPT y Perplexity tienen desviaciones estándar idénticas de 0.632 y varianzas correspondientes de 0.400. Por otro lado, Bard y ChatSonic exhiben desviaciones estándar de 0.816 y 1.033, junto con varianzas de 0.667 y 1.067, respectivamente. Los coeficientes de variación, que indican la proporción de la desviación estándar con respecto a la media, varían entre 21.08% y 30.62%, lo que señala diferentes niveles de variabilidad en las distribuciones de datos.

En lo que respecta a la forma de las distribuciones, las medianas para las variables ChatGPT, Bard, Perplexity y ChatSonic son 3.000, 2.500, 4.000 y 4.000, respectivamente. Los valores de rango, que muestran la diferencia entre los valores mínimo y máximo, varían entre 2.000 y 3.000. El rango intercuartílico (IQR), que indica la dispersión de los valores centrales,

fluctúa entre 0.500 y 1.500. Además, se ha identificado el valor modal, o moda, para cada variable, siendo 3 para ChatGPT, 2 para Bard y 4 para Perplexity y ChatSonic.

La asimetría, una medida de la distribución de los datos en relación con la media, varía entre -0.67 y 0.86. Valores negativos sugieren una distribución más sesgada hacia la izquierda, mientras que valores positivos indican un sesgo hacia la derecha. La curtosis, que evalúa la forma de los picos en la distribución, tiene valores entre -0.30 y 2.50. Un valor de curtosis de 0 señala una distribución normal, valores negativos indican una distribución achatada y valores positivos sugieren una distribución más puntiaguda.

Tabla 30 - Análisis estadístico comparativo sobre la fiabilidad.

Variable	Conteo total	N	N*	Número acumulado	Porcentaje	Porcentaje acumulado	Media	Error estándar de la media	Desviación estándar
ChatGPT	10	10	0	10	100	100	3.100	0.180	0.568
Bard	10	10	0	10	100	100	3.000	0.258	0.816
Perplexity	10	10	0	10	100	100	3.900	0.233	0.738
ChatSonic	10	10	0	10	100	100	3.600	0.267	0.843

Variable	Varianza	Coefficiente de variación	Mínimo	Q1	Mediana	Q3	Máximo	Rango	IQR	Modo
ChatGPT	0.322	18.31	2.000	3.000	3.000	3.250	4.000	2.000	0.250	3
Bard	0.667	27.22	2.000	2.000	3.000	4.000	4.000	2.000	2.000	3
Perplexity	0.544	18.92	3.000	3.000	4.000	4.250	5.000	2.000	1.250	4
ChatSonic	0.711	23.42	2.000	3.000	4.000	4.000	5.000	3.000	1.000	4

Variable	Número para moda	Asimetría	Curtosis
ChatGPT	7	0.09	1.50
Bard	4	0.00	-1.39
Perplexity	5	0.17	-0.73
ChatSonic	5	-0.39	0.37

Fuente: Elaboración propia.

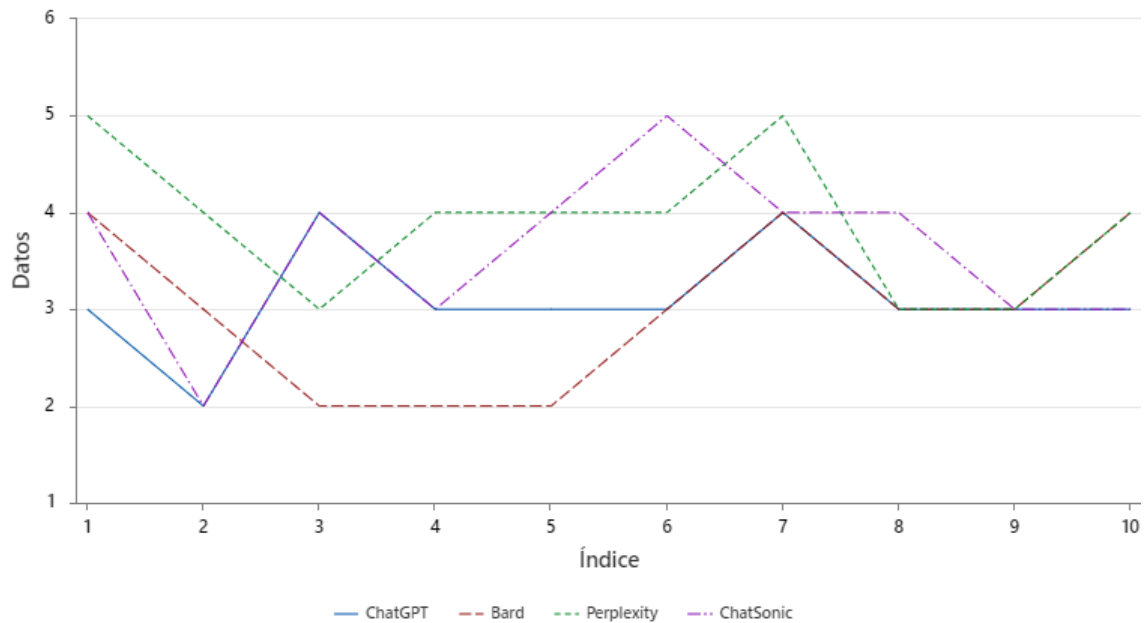


Ilustración 28 - Fiabilidad.

Fuente: Elaboración propia.

5.1.3 ENCUESTA DE ACEPTACIÓN

La encuesta efectuada (Anexo 13) tenía como propósito conocer el nivel de aceptación por parte de la muestra, para implementar este tipo de tecnología.

5.1.3.1 Análisis cuantitativo

En el contexto de la ejecución de la encuesta de aceptación, se identificaron varias incertidumbres persistentes. Como parte de la evaluación, se formularon una serie de interrogantes con el propósito de medir el grado de aplicabilidad percibido por los participantes en relación con los chatbots que se habían empleado previamente en el estudio.

En este sentido, se concluye que, de los diez sujetos de prueba analizados (Tabla 31), el cien por ciento manifestó una genuina predisposición para utilizar chatbots en el futuro, tanto aquellos que fueron presentados en el estudio como aquellos que puedan encontrar posteriormente. Estos individuos consideran que los chatbots representarían una herramienta adicional para sus actividades diarias, dado a las facilidades de búsqueda que estas representan.

Tabla 31 - Empleo de chatbots como herramienta en sus actividades cotidianas.

<u>Emplearía chatbots como herramienta en sus actividades cotidianas.</u>	<u>Conteo</u>	<u>Porcentaje</u>	<u>Conteo Acumulado</u>	<u>Porcentaje Acumulado</u>
Si	10	100.00	10	100.00
N=	10			

Fuente: Elaboración propia.

Además, entre los 10 sujetos de prueba, 8 de ellos expresaron su intención de emplear los chatbots como una herramienta de trabajo, basándose en las facilidades que estos demostraron durante las pruebas previas (Tabla 32). En contraste, los dos sujetos restantes manifestaron una actitud más reservada, caracterizada por un "tal vez", debido a complicaciones surgidas durante las pruebas o la percepción de que los chatbots no proporcionaban información suficientemente precisa o relevante.

Tabla 32 - Implementarlos facilitará su trabajo.

Implementarlos facilitará su trabajo	Conteo	Porcentaje	Conteo Acumulado	Porcentaje Acumulado
Si	8	80.00	8	80.00
Tal vez	2	20.00	10	100.00
N=	10			

Fuente: Elaboración propia.

En cuanto a la frecuencia con la que planeaban implementar los chatbots, entre los diez voluntarios que participaron en las pruebas y adquirieron un mayor entendimiento sobre los diferentes chatbots evaluados (Tabla 33), cuatro indicaron que los utilizarían de manera "frecuente", mientras que tres de ellos afirmaron que los emplearían "muy frecuentemente". Por otro lado, los tres voluntarios restantes expresaron su intención de utilizarlos "ocasionalmente". Estas decisiones se basaron en la experiencia previa que tuvieron con los chatbots en el estudio.

Tabla 33 - Frecuencia con el que los implementaría en el trabajo.

Frecuencia con el que los implementaría en el trabajo.	Conteo	Porcentaje	Conteo Acumulado	Porcentaje Acumulado
Frecuentemente	4	40.00	4	40.00
Muy frecuentemente	3	30.00	7	70.00
Ocasionalmente	3	30.00	10	100.00
N=	10			

Fuente: Elaboración propia.

De acuerdo con las respuestas recopiladas en la encuesta, la valoración del nivel de recomendación para la incorporación de chatbots en una unidad de biomédica, realizada por los diez sujetos de prueba, refleja diversas percepciones (Tabla 34). En este sentido, dos de ellos mantienen una posición "neutral" al respecto, mientras que tres consideran que es "probable" que la recomendación sea válida, basándose en su experiencia previa. Por último, los cinco encuestados restantes mostraron una inclinación hacia calificar la recomendación

como "muy probable", evidenciando así un consenso en la creencia de que la adopción de chatbots podría contribuir significativamente a mejorar y agilizar las operaciones en esta área biomédica.

Tabla 34 - Recomendación para su uso en unidad de biomédica.

Recomendación para su uso en unidad de biomédica	Conteo	Porcentaje	Conteo Acumulado	Porcentaje Acumulado
Neutral	2	20.00	2	20.00
Probable	3	30.00	5	50.00
Muy Probable	5	50.00	10	100.00
N=	10			

Fuente: Elaboración propia.

En lo que respecta a los resultados relativos al grado de confianza, se observan tendencias claras (Tabla 35). De los diez sujetos de prueba, siete manifestaron una posición de "neutralidad" en cuanto a la fiabilidad de la información proporcionada por los chatbots, mientras que los tres sujetos restantes expresaron "confianza" en la precisión y calidad de la información entregada y evaluada durante las pruebas con los chatbots

Tabla 35 - Nivel de confianza en chatbots.

Nivel de confianza	Conteo	Porcentaje	Conteo Acumulado	Porcentaje Acumulado
Neutral	7	70.00	7	70.00
Confianza	3	30.00	10	100.00
N=	10			

Fuente: Elaboración propia.

En cuanto a la inquietud manifestada sobre la privacidad, que se erigió como un tema recurrente y central en las discusiones durante las pruebas, es importante destacar que de los diez sujetos entrevistados, ocho de ellos expresaron que no experimentan preocupaciones al respecto. En contraste, los dos participantes restantes reconocieron que sí sienten inquietud, ya que perciben que la cuestión de la privacidad tiene implicaciones tanto para los datos del hospital como para los datos personales de los usuarios que interactúan con los diversos agentes conversacionales.

Tabla 36 - Preocupación por la privacidad.

Preocupación por la privacidad	Conteo	Porcentaje	Conteo Acumulado	Porcentaje Acumulado
No	8	80.00	8	80.00
Si	2	20.00	10	100.00
N=	10			

Fuente: Elaboración propia.

En el presente análisis (Ilustración 31), se sometieron a evaluación diversas medidas estadísticas con el propósito de evaluar la eficacia de los agentes conversacionales en la asistencia a los usuarios para gestionar errores. Con respecto a las medidas de centralidad, se encontró que las variables confianza en la capacidad de los chatbots, privacidad y seguridad, disponibilidad de trabajo y nivel de confianza presentan medias cercanas alrededor de 3.500, Mientras que disposición a obtener información exhibe una media ligeramente más alta, alcanzando 4.2000. Esta convergencia sugiere que, en términos promedio, los distintos métodos demuestran un desempeño equiparable en la tarea analizada, obteniendo resultados satisfactorios. Sin embargo, se observa que el error estándar de la media es menor para las variables confianza en la capacidad de los chatbots y nivel de confianza en comparación con los otros métodos, lo que implica una mayor consistencia en las puntuaciones.

En relación con la dispersión de los datos, se exploraron la desviación estándar y la varianza. Se identificó que la variable privacidad y seguridad presenta la desviación estándar más elevadas (1.354), lo cual indica una mayor variabilidad en las puntuaciones. Por otro lado, disposición a obtener información y disponibilidad a trabajar con chatbots comparten un valor inferior de 0.632, lo que sugiere una mayor cohesión en sus resultados. Mientras que nivel de confianza y confianza en la capacidad de los chatbots exhiben una desviación estándar de 0.483 y 0.422 respectivamente. Las varianzas siguen un patrón similar.

La mediana, que representa el valor medio en un conjunto de datos ordenados, sitúa a disposición a obtener información y disponibilidad a trabajar con chatbots con el valor más alto de esta categoría, alcanzando 4.000. Le sigue privacidad y seguridad con 3.500, finalmente confianza en su capacidad y nivel de confianza permanecen constantes en 3.000. El rango, indicador de la diferencia entre los valores máximo y mínimo, presenta un mayor valor para privacidad y seguridad (4.000) y valores menores para disposición a obtener información y disponibilidad de trabajar con 2.000 para ambas variables. Por otro lado, confianza en su capacidad y nivel de Confianza obtuvieron un valor de 1.000.

La distribución se evaluó mediante la asimetría y la curtosis. Se determinó que disposición a obtener información y privacidad y seguridad exhiben una asimetría negativa, sugiriendo una distribución sesgada hacia la izquierda. Mientras que disponibilidad de trabajar con chatbots, confianza en su capacidad y nivel de confianza presentan una asimetría positiva, indicando un sesgo hacia la derecha. Tanto disposición a obtener información como disponibilidad de trabajo revelan una asimetría cercana a cero, lo que señala distribuciones

más simétricas. La curtosis varía para las variables, sugiriendo los valores positivos una mayor concentración de valores alrededor de la media con colas más pronunciadas, mientras que una curtosis negativa implica una distribución más amplia y menos concentración de valores extremos.

El análisis estadístico (Tabla 3) puso en evidencia las diferencias en términos de dispersión, centralidad y forma de distribución con respecto a la capacidad de los métodos para asistir a los usuarios en la identificación y solución de errores. Estas medidas proporcionan una visión detallada, lo que puede contribuir a tomar decisiones informadas.

Tabla 37 - Análisis estadístico de variables cuantitativas sobre la aceptación.

Variable	Conteo total	N	N*	Número acumulado	Porcentaje	Porcentaje acumulado	Media	Error estándar de la media
Disposición a obtener información	10	10	0	10	100	100	4.200	0.200
Disponibilidad de trabajar con chatbots	10	10	0	10	100	100	3.800	0.200
Confianza en la capacidad	10	10	0	10	100	100	3.200	0.133
Privacidad y seguridad	10	10	0	10	100	100	3.500	0.428
Nivel de confianza	10	10	0	10	100	100	3.300	0.153

Variable	Desviación estándar	Varianza	Coficiente de variación	Mínimo	Q1	Mediana	Q3	Máximo
Disposición a obtener información	0.632	0.400	15.06	3.000	4.000	4.000	5.000	5.000
Disponibilidad de trabajar con chatbots	0.632	0.400	16.64	3.000	3.000	4.000	4.000	5.000
Confianza en la capacidad	0.422	0.178	13.18	3.000	3.000	3.000	3.250	4.000
Privacidad y seguridad	1.354	1.833	38.69	1.000	2.750	3.500	5.000	5.000
Nivel de confianza	0.483	0.233	14.64	3.000	3.000	3.000	4.000	4.000

Variable	Rango	IQR	Modo	Número para moda	Asimetría	Curtosis
Disposición a obtener información	2.000	1.000	4	6	-0.13	0.18
Disponibilidad de trabajar con chatbots	2.000	1.000	4	6	0.13	0.18
Confianza en la capacidad	1.000	0.250	3	8	1.78	1.41
Privacidad y seguridad	4.000	2.250	3; 5	3	-0.50	-0.47
Nivel de confianza	1.000	1.000	3	7	1.04	-1.22

Fuente: Elaboración propia.

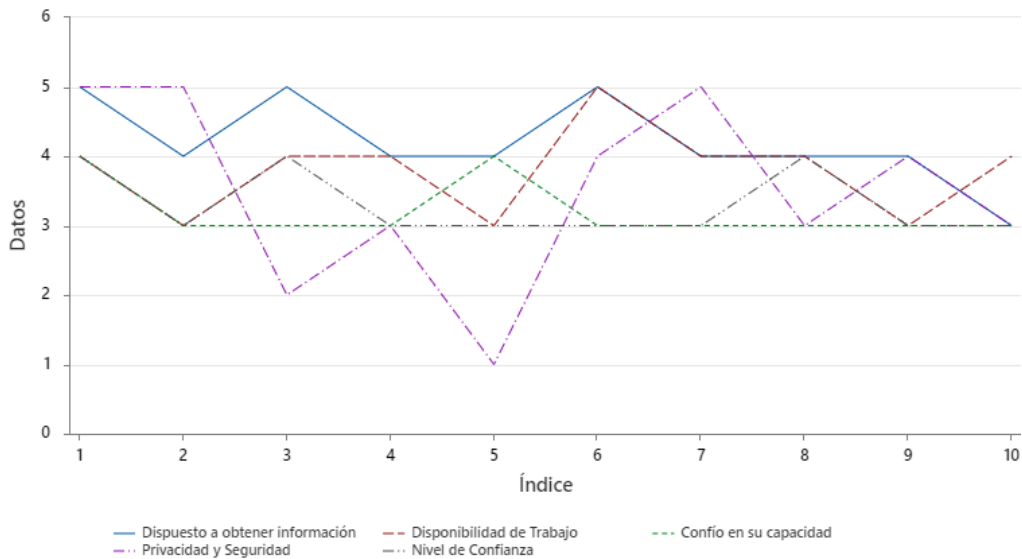


Ilustración 29 - Datos cuantitativos sobre la aceptación.

Fuente: Elaboración propia

Con el fin de mejorar la comprensión de las calificaciones otorgadas, se procedió a obtener el promedio de cada componente de la heurística para los cuatro tipos de agentes conversacionales. Estos resultados se presentan en una comparativa visual (Ilustración 30).

Es crucial señalar en esta ocasión que Perplexity se destaca notablemente al lograr la calificación promedio más alta en cada uno de los elementos sometidos a evaluación. Este hecho señala su sobresaliente rendimiento en todos los aspectos considerados en el marco de la evaluación.

	ChatGPT	Google Bard	Perplexity AI	ChatSonic
Visibilidad del estado del sistema	4.300	4.400	4.500	3.800
Correspondencia entre el sistema y el mundo real	3.900	4.100	4.500	3.900
Control y libertad del usuario	3.500	3.500	3.800	3.600
Consistencia y normas	3.700	3.600	4.400	3.500
Prevención de errores	3.500	3.300	3.800	3.600
Ayuda y orientación	3.700	3.600	4.400	3.500
Flexibilidad y eficacia de uso	4.300	3.800	4.500	3.400
Diseño estético, minimalista y atractivo	4.400	3.800	4.500	3.600
Ayuda a los usuarios a reconocer, diagnosticar y recuperarse de los errores	3.800	3.600	3.900	3.300
Conservación del contexto	4.100	3.700	4.400	3.900
Fiabilidad	3.100	3.000	3.900	3.600

Ilustración 30 – Comparativa de la media de las características evaluadas en la heurística.

Fuente: Elaboración propia.

Al considerar los promedios de los diversos componentes que conforman la heurística, se derivó un promedio general que, a su vez, determinó la calificación final asignada a cada chatbot. En este contexto, es importante destacar que Perplexity AI sobresalió al obtener la calificación más elevada, con un puntaje de 4.24. Seguido por ChatGPT, que obtuvo una calificación de 3.85. En el tercer lugar, encontramos a Google Bard con una calificación de 3.67, y con una pequeña diferencia, en el cuarto lugar, se ubicó ChatSonic con una calificación de 3.61 (Ilustración 31).

De esta manera, se pone de manifiesto que, a pesar de la prominencia de otros chatbots evaluados en las pruebas, emergieron aspectos distintivos que conferían un mayor valor a Perplexity AI, consolidándolo como el favorito entre los cuatro chatbots utilizados en el estudio.

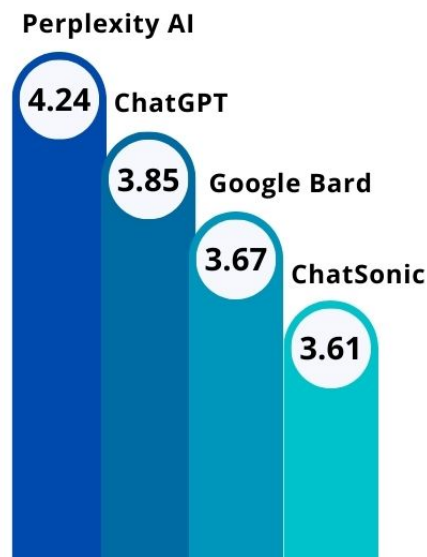


Ilustración 31 - Calificación general de cada chatbot de acuerdo con la media.

Fuente: Elaboración propia.

5.2.1.2 Análisis cualitativo

Basándonos en las percepciones de los encuestados, las características que sobresalen como elementos clave para la utilidad de un chatbot se pueden resumir de la siguiente manera: la capacidad de presentar documentación técnica, proporcionar acceso a referencias bibliográficas, ofrecer soluciones a problemas específicos, eliminar las barreras del idioma, y mantener una alta precisión en las respuestas (Ilustración 32). Estas cualidades contribuyen significativamente a la facilidad de uso y promueven una interacción más efectiva con los usuarios. Esto, a su vez, contribuye a la creación de un entorno de usuario amigable y una funcionalidad sólida del chatbot en su conjunto.

Esta información derivada de las respuestas de los encuestados es esencial para comprender y mejorar el rendimiento y la eficacia de los chatbots, lo que puede llevar a una experiencia de usuario más enriquecedora y productiva en el futuro.



Ilustración 32 - Características más valiosas de los chatbots según la opinión.

Fuente: Elaboración propia.

Asimismo, dentro del ámbito de la ingeniería clínica, se ha identificado una serie de características esenciales que contribuyen a la efectividad de un chatbot. Estas conclusiones se derivan tanto de las entrevistas con los encuestados como de las respuestas proporcionadas en los formularios correspondientes (Ilustración 33). Entre las respuestas que se destacan en este contexto, se enfatiza la necesidad de que el chatbot ofrezca soluciones técnicas de manera directa y precisa para los problemas específicos, que sea intuitivo en su funcionamiento, capaz de proporcionar referencias y fuentes bibliográficas confiables, así como ejemplos de prácticas previamente implementadas. Además, se subraya la importancia de facilitar la búsqueda de manuales y, notablemente, la mención más frecuente se refiere a la inclusión de imágenes de referencia de procedimientos o videos ilustrativos.

Estos hallazgos destacan los atributos cruciales para la funcionalidad óptima de un chatbot en el contexto de la ingeniería clínica. Estas características no solo mejoran la experiencia del usuario, sino que también aumentan la utilidad y la eficacia del chatbot en esta área especializada.

La consideración y aplicación de estas características puede ser crucial para el desarrollo de futuros chatbots en ingeniería clínica, ya que se adaptan específicamente a las necesidades y expectativas de los profesionales en este campo.

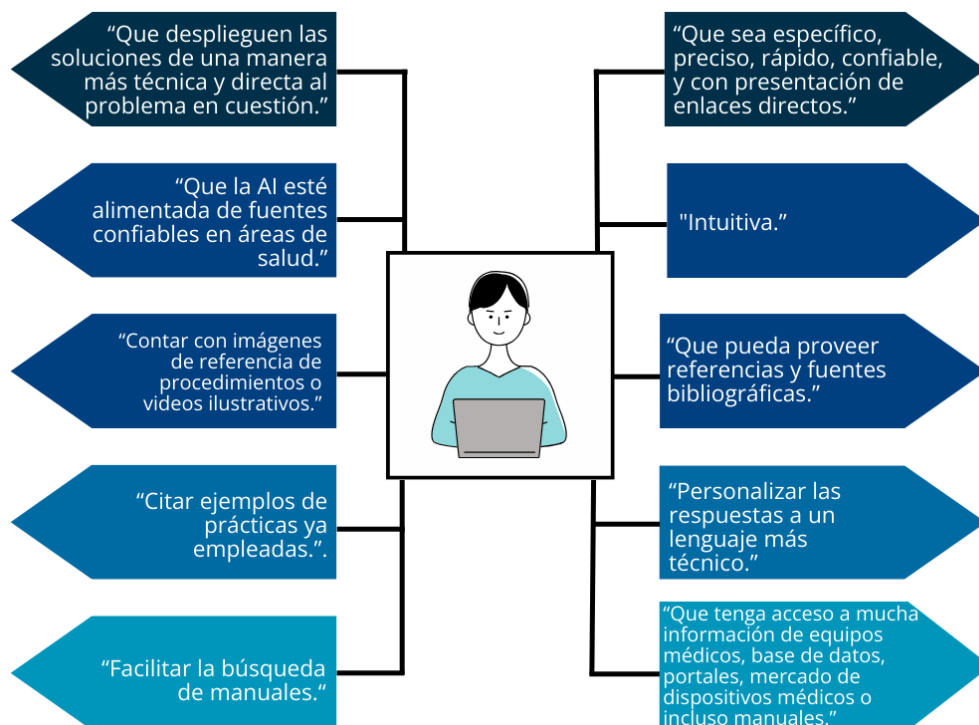


Ilustración 33 - Características esenciales para que un chatbot sea efectivo en el ámbito de la ingeniería clínica.

Fuente: Elaboración propia.

En el proceso de construir la encuesta de aceptación, se planteó una pregunta particularmente intrigante: después de utilizar las cuatro herramientas, ¿cuál de ellas los entrevistados recomendarían en función de las características evaluadas en la heurística de usabilidad? Como resultado, dos de ellas emergieron como las más destacadas: ChatGPT, con un 10% de preferencia, y Perplexity, con un notable 90% (Ilustración 34). Esta última herramienta, en particular, se ha destacado como la favorita para aplicaciones en el campo de la ingeniería clínica, debido a su exhaustividad y la cumplimentación de un mayor número de cualidades apreciadas por los usuarios.

Esta preferencia resaltada por los entrevistados hacia Perplexity puede sugerir una mayor idoneidad de esta herramienta en el ámbito de la ingeniería clínica, lo que podría tener implicaciones significativas en la elección de herramientas en futuros proyectos y aplicaciones dentro de este campo.

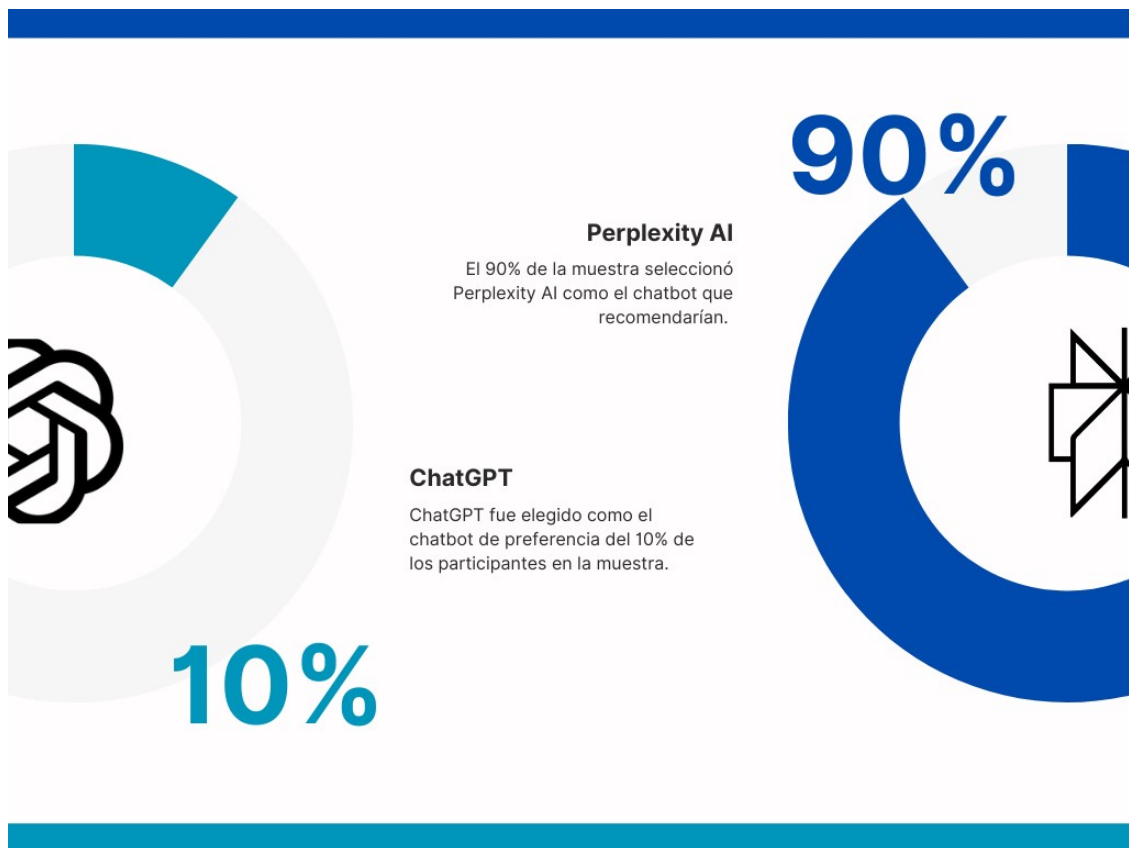


Ilustración 34 - Chatbot más recomendado según la muestra.

Fuente: Elaboración propia.

Así como se han identificado aspectos positivos en la utilización de los chatbots, durante la ejecución de las distintas pruebas con estas herramientas, los encuestados también se han enfrentado a diversos desafíos y dificultades (Ilustración 35). Entre estos obstáculos, se han destacado principalmente:

- 1) Falta de fundamentación bibliográfica: Algunas respuestas proporcionadas por los chatbots carecían de respaldo en forma de bibliografía o referencias, lo que generaba incertidumbre en cuanto a la fiabilidad de la información.
- 2) Falta de precisión: En ocasiones, las respuestas resultaban ser demasiado generales, lo que resultaba insatisfactorio en temas que requerían especificidad y precisión.
- 3) Limitación de información: En algunos casos, la información se mostraba restringida, ya sea debido a preocupaciones de confidencialidad por parte de las empresas propietarias de equipos médicos o simplemente porque no estaba disponible.

- 4) Dependencia contextual: En varias instancias, la obtención de respuestas adecuadas dependía en gran medida del contexto de la pregunta, lo que podía dificultar la localización de la información precisa.
- 5) Necesidad de habilidad en la formulación de preguntas: La habilidad para formular preguntas adecuadas se destacó como un factor crucial para obtener respuestas pertinentes y precisas.

Estos desafíos señalados proporcionan una visión más completa de las complejidades asociadas con la utilización de chatbots en el ámbito de la ingeniería clínica. Su resolución puede requerir un enfoque más matizado y una mejora continua de estas herramientas para satisfacer las necesidades de los usuarios de manera más efectiva.

La consideración y abordaje de estas dificultades son fundamentales para el desarrollo futuro de chatbots en este campo y pueden llevar a mejoras significativas en la calidad de las respuestas y la satisfacción de los usuarios.

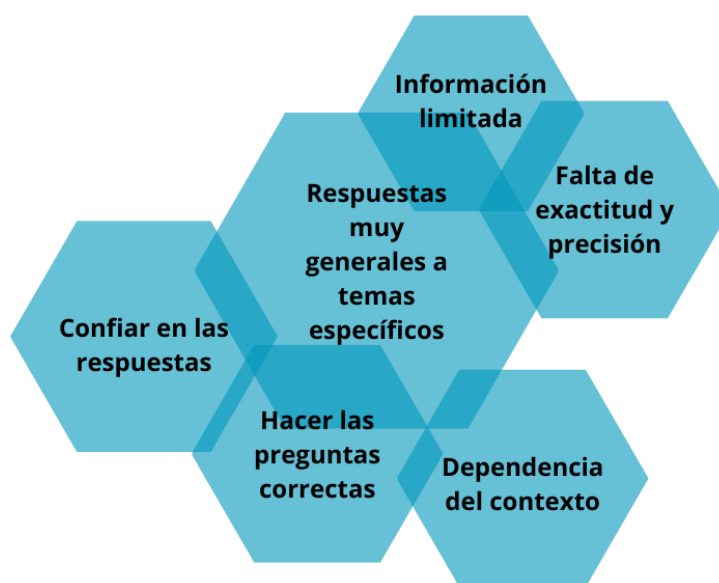


Ilustración 35 - Dificultad al utilizar chatbots de acuerdo con la opinión de la muestra.

Fuente: Elaboración propia.

5.1.4 COSTOS OPERACIONALES

5.1.4.1 ChatGPT

La iteración GPT-3.5 de ChatGPT se ofrece de manera totalmente gratuita, sin restricciones de uso ni limitaciones en términos de extensión textual. Sin embargo, cabe destacar que existe una iteración posterior denominada GPT-4.0, la cual conlleva un costo de 20 dólares estadounidenses (Tabla 38). No obstante, es relevante resaltar que esta versión

mejorada está en la actualidad disponible solamente para utilización dentro del territorio de los Estados Unidos de América.

Tabla 38 - Costos operacionales ChatGPT.

	Plan gratis	ChatGPT Plus
Mensual	\$0	\$20

*\$ dólares americanos

Fuente: Elaboración propia.

5.1.4.2 *Google Bard*

En lo que concierne a Google Bard, su acceso es de costo alguno. La entrada a esta herramienta está abierta a cualquier individuo mediante el sitio web oficial de Google Bard o por medio de la plataforma WhatsApp, sin necesidad de suscripciones ni desembolsos monetarios.

Cabe señalar que Google Bard se encuentra en fase experimental, por lo exento cual no se descarta la eventual introducción de tarifas en un futuro. No obstante, en la actualidad, Bard prevalece como una utilidad de libre disposición al alcance de cualquier persona (Tabla 39).

Tabla 39 - Costos operacionales de Google Bard.

	Etapas experimental
Mensual	\$0

*\$ dólares americanos

Fuente: Elaboración propia.

5.1.4.3 *Perplexity AI*

Los precios asociados a Perplexity varían según la edición empleada (Tabla 40). La variante de Perplexity de carácter gratuito está a disposición de los usuarios sin requerir pago alguno. Por otro lado, se encuentra la iteración remunerada de Perplexity denominada Perplexity Pro, la cual implica un desembolso de 20 dólares estadounidenses.

Aquellos que opten por la suscripción de Perplexity Pro obtendrán prioridad en el acceso al modelo, intervalos de respuesta más ágiles, una experiencia más fluida y un aumento en el límite de utilización.

En resumen, la edición sin costos de Perplexity es de acceso gratuito, en contraste con la variante de pago, Perplexity Pro, que requiere un desembolso de 20 dólares estadounidenses mensuales.

Tabla 40 - Costos operacionales de Perplexity.

	Free Plan	Pro
Mensual	\$0	\$20
Anual	\$0	\$200

*\$ dólares americanos

Fuente: Elaboración propia.

5.1.4.4 ChatSonic

ChatSonic se ofrece de manera gratuita, permitiendo su utilización sin imponer cargos. Sin embargo, en la edición de gratuidad existe un límite fijado en 10,000 palabras. En el caso de requerir una producción textual más extensa, existe la opción de adquirir uno de sus esquemas de remuneración (Tabla 41).

El nivel elemental de suscripción tiene un costo de 16 dólares al mes y habilita la generación de hasta 47,500 palabras. La modalidad de suscripción más avanzada, con un precio de 99 dólares al mes, posibilita la creación de hasta 200,000 palabras. ChatSonic se constituye en una utilidad de confección de texto impulsada por IA, propiciando la creación eficiente y sencilla de contenido de alta calidad.

Tabla 41 - Costos operacionales de ChatSonic.

	Free Trial	Unlimited	Business	Enterprise
Mensual	\$0	\$16	\$19	Elaborar una evaluación de gastos por organización.
Anual	\$0	\$192	\$152.04	

*\$ dólares americanos

Fuente: Elaboración propia.

VI. CONCLUSIONES

- Se realizó una evaluación acerca de la posible adopción de cuatro herramientas de inteligencia artificial y su impacto en la labor de los ingenieros clínicos. Se determinó que los agentes conversacionales pueden desempeñar un papel como herramienta auxiliar que agiliza el acceso a información. Debido a que estas herramientas tienen la capacidad de proporcionar rúbricas de evaluación de desempeño, manuales y listas de posibles fallas y soluciones a los desafíos que se presenten.
 - a) La posible incorporación de herramientas de inteligencia artificial, como los chatbots, en el entorno de trabajo de los ingenieros clínicos ha demostrado un impacto positivo. Esto se refleja en una mejora significativa en la eficiencia en la toma de decisiones. Debido a que los chatbots facilitan un acceso más rápido a información relevante. Esto no solo optimiza la gestión de dispositivos y sistemas médicos, sino que también contribuye a una atención más eficaz y segura para los pacientes, al tiempo que reduce la carga de trabajo y el estrés de los ingenieros clínicos.
 - b) La percepción de la muestra de ingenieros biomédicos sobre la utilidad de las herramientas de inteligencia artificial es positiva. Han reconocido que estas herramientas les brindan un soporte valioso para conocer lineamientos para mantenimientos preventivos, la creación de herramientas complementarias, resolución de problemas y la indagación de manuales.
 - c) La aplicación de una heurística efectiva ha llevado a la identificación de Perplexity AI como el chatbot más eficaz para los ingenieros biomédicos en la zona noroccidental de Honduras en 2023. Este chatbot ha demostrado ser altamente adaptable a las necesidades específicas de la profesión, optimizando la asistencia técnica.

VII. RECOMENDACIONES

- Llevar a cabo evaluaciones adicionales con nuevas implementaciones de inteligencia artificial. Asimismo, se sugiere incrementar el número de estas implementaciones para su análisis y considerar la posibilidad de someterlas a un proceso de entrenamiento para la obtención de datos más específicos y pertinentes en el contexto de la ingeniería clínica. Este enfoque permitirá una mayor exploración y validación de las capacidades de las inteligencias artificiales en dicho campo.
- Se aconseja realizar evaluaciones rigurosas que incluyan análisis detallados del rendimiento de los agentes conversacionales en situaciones prácticas y en escenarios más cercanos a la aplicación en cuestión.
- Se sugiere fomentar una mayor colaboración entre profesionales y expertos en el área con el propósito de facilitar el intercambio de ideas, conocimientos y recursos. Además, esta colaboración puede ser fundamental para una comprensión más profunda y precisa de la implementación de las pruebas en cuestión. La sinergia entre individuos con experiencia en el campo podría enriquecer significativamente la planificación y ejecución de las pruebas, mejorando así la calidad de los resultados obtenidos.

VIII. IMPLEMENTACIÓN

Los chatbots o agentes conversacionales pueden ser integrados como herramientas auxiliares en unidades de ingeniería biomédica o relacionados, dentro de instituciones públicas o privadas. Su implementación demanda únicamente de una infraestructura tecnológica básica. Esto facilitaría la ejecución de tareas repetitivas, además de proporcionar una perspectiva suplementaria respecto a los problemas presentados. La aplicabilidad de estos sistemas podría contribuir a elevar la calidad de los servicios proporcionados.

IX. EVOLUCIÓN DEL TRABAJO ACTUAL

A lo largo del desarrollo de la investigación, se identificaron ciertas restricciones relacionadas con la implementación de las pruebas. Esto se debe a que, a pesar de la importancia de la ingeniería clínica en la región costa norte, existe una escasez de profesionales en ingeniería que se dediquen específicamente a este campo. Como resultado, la evaluación se ha visto limitada en cierta medida, a pesar de esto, se obtuvieron resultados favorables.

Se anticipa que en el futuro se disponga de un mayor número de inteligencias artificiales para permitir otras comparaciones en relación con la solución preferida identificada en este estudio. Esto servirá para determinar si la herramienta mejor calificada en la investigación actual sigue siendo la opción más precisa y eficaz. Además, se contempla la realización de otras modalidades de simulaciones para evaluar la eficiencia y eficacia de la aplicación de estas tecnologías en la ingeniería clínica. Estas acciones futuras contribuirán a enriquecer aún más la comprensión de la utilidad de las inteligencias artificiales en este campo.

BIBLIOGRAFÍA

- 1) Abid Haleem, Mohd Javid, & Ravi Pratap Singh. (2022). *An era of ChatGPT as a significant futuristic support tool: A study on features, abilities, and challenges*. <https://doi.org/10.1016/j.tbench.2023.100089>
- 2) Acosta, J. N., Falcone, G. J., Rajpurkar, P., & Topol, E. J. (2022). Multimodal biomedical AI. *Nature Medicine*, 28(9), 1773–1784. <https://doi.org/10.1038/s41591-022-01981-2>
- 3) Avila, H. F., González, M. M., & Licea, S. M. (2020). LA ENTREVISTA Y LA ENCUESTA: ¿MÉTODOS O TÉCNICAS DE INDAGACIÓN EMPÍRICA? *Didáctica y Educación ISSN 2224-2643*, 11(3), Article 3.
- 4) Banco de Desarrollo de América Latina y el Caribe. (2021). *¿Cómo puede la inteligencia artificial mejorar la salud de los latinoamericanos?* <https://www.caf.com/es/conocimiento/visiones/2021/09/como-puede-la-inteligencia-artificial-mejorar-la-salud-de-los-latinoamericanos/>
- 5) Branditechure. (2023). *Download Google Bard Logo PNG and Vector (PDF, SVG, Ai, EPS) Free*. <https://branditechure.agency/brand-logos/download/google-bard/>
- 6) Chandel, S., Yuying, Y., Yujie, G., Razaque, A., & Yang, G. (2019). Chatbot: Efficient and Utility-Based Platform. In K. Arai, S. Kapoor, & R. Bhatia (Eds.), *Intelligent Computing* (pp. 109–122). Springer International Publishing. https://doi.org/10.1007/978-3-030-01174-1_9
- 7) Chua, I. S., Gaziel-Yablowitz, M., Korach, Z. T., Kehl, K. L., Levitan, N. A., Arriaga, Y. E., Jackson, G. P., Bates, D. W., & Hassett, M. (2021). Artificial intelligence in oncology: Path to implementation. *Cancer Medicine*, 10(12), 4138–4149. <https://doi.org/10.1002/cam4.3935>
- 8) Comisión Europea. (2019, April 8). *Ethics guidelines for trustworthy AI | Shaping Europe's digital future*. <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>
- 9) Corvalán, J. G. (2019). El impacto de la inteligencia artificial en el trabajo. *Revista de Direito Econômico e Socioambiental*, 10, 37. <https://doi.org/10.7213/rev.dir.econ.soc.v10i1.25870>
- 10) Cuesta Ríos, C. (2021, July). *Retos legales y éticos de las inteligencias artificiales en el diagnóstico de enfermedades* [Info:eu-repo/semantics/bachelorThesis]. E.T.S.I de Sistemas Informáticos (UPM). <https://oa.upm.es/68212/>

- 11) De Ramón Fernández, A., Ruiz Fernández, D., & Prieto Sánchez, M. T. (2019). A decision support system for predicting the treatment of ectopic pregnancies. *International Journal of Medical Informatics*, 129, 198–204. <https://doi.org/10.1016/j.ijmedinf.2019.06.002>
- 12) Diction, S. &. (2023, August 18). Branding perplexity.ai. *Smith & Diction*. <https://medium.com/smith-diction/branding-perplexity-ai-70eb2cb2ef48>
- 13) García Benítez, V., & Ruvalcaba-Gómez, E. A. (2021). Análisis de las estrategias nacionales de inteligencia artificial en América Latina: Estudio de los enfoques de ética y de derechos humanos. *Revista de Gestión Pública*, 10(1), 5–32.
- 14) Gausemeier, J., & Moehring, S. (2002). VDI 2206- A New Guideline for the Design of Mechatronic Systems. *IFAC Proceedings Volumes*, 35(2), 785–790. [https://doi.org/10.1016/S1474-6670\(17\)34035-1](https://doi.org/10.1016/S1474-6670(17)34035-1)
- 15) Google. (2023). *Preguntas frecuentes sobre Bard*. <https://bard.google.com/faq>
- 16) Hernández-Sampieri, R., & Mendoza, C. (2018). *Metodología de la investigación: Las rutas cuantitativa, cualitativa y mixta*. McGraw Hill.
- 17) Hitakshi. (2023, April 20). *What is ChatSonic: The Most Advanced AI Chatbot in 2023*. MLYearning. <https://www.mlyearning.org/chatsonic/>
- 18) Ibáñez, C. R., & Soto, J. M. (2006). Ingeniería Clínica: Introducción, percepción y práctica en el área metropolitana de Medellín. *CES Medicina*.
- 19) Insider Intelligence, M. (2022). *Chatbot market in 2022: Stats, trends, and companies in the growing AI chatbot industry*. Insider Intelligence. <https://www.insiderintelligence.com/insights/chatbot-market-stats-trends/>
- 20) Janiesch, C., Zschech, P., & Heinrich, K. (2021). Machine learning and deep learning. *Electronic Markets*, 31(3), 685–695. <https://doi.org/10.1007/s12525-021-00475-2>
- 21) Jiang, Z., Xu, F. F., Araki, J., & Neubig, G. (2020). How Can We Know What Language Models Know? *Transactions of the Association for Computational Linguistics*, 8, 423–438. https://doi.org/10.1162/tacl_a_00324
- 22) Jiménez, C., Rusu, C., Roncagliolo, S., Inostroza, R., & Rusu, V. (2012). Evaluating a Methodology to Establish Usability Heuristics. *2012 31st International Conference of the Chilean Computer Science Society*, 51–59. <https://doi.org/10.1109/SCCC.2012.14>

- 23) Juaréz, P. F. (2018). La importancia de la técnica de la entrevista en la investigación en comunicación y las ciencias sociales. *Investigación documental. Ventajas y limitaciones. Sintaxis*, 1, Article 1. <https://doi.org/10.36105/stx.2018n1.07>
- 24) Karczewski, K. J., & Snyder, M. P. (2018). Integrative omics for health and disease. *Nature Reviews Genetics*, 19(5), 299–310. <https://doi.org/10.1038/nrg.2018.4>
- 25) Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., ... Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, 102274. <https://doi.org/10.1016/j.lindif.2023.102274>
- 26) Kavya, R., Christopher, J., Panda, S., & Lazarus, Y. B. (2021). Machine Learning and XAI approaches for Allergy Diagnosis. *Biomedical Signal Processing and Control*, 69, 102681. <https://doi.org/10.1016/j.bspc.2021.102681>
- 27) Kelleher, J. D. (2019). *Deep Learning*. MIT Press.
- 28) Kojima, T., Gu, S. S., Reid, M., Matsuo, Y., & Iwasawa, Y. (2022). *Large Language Models are Zero-Shot Reasoners*. <https://doi.org/10.48550/arXiv.2205.11916>
- 29) Langevin, R., Lordon, R., Avrahami, T., Cowan, B., Hirsch, T., & Hsieh, G. (2021). *Heuristic Evaluation of Conversational Agents*. 15. <https://doi.org/10.1145/3411764.3445312>
- 30) Liu, G. K.-M. (2023). Transforming Human Interactions with AI via Reinforcement Learning with Human Feedback (RLHF). *Massachusetts Institute of Technology*. <https://computing.mit.edu/wp-content/uploads/2023/06/Transforming-Human-Interactions-with-AI-via-Reinforcement-Learning-with-Human-Feedback-RLHF.pdf>
- 31) Lundberg, S. M., Nair, B., Vavilala, M. S., Horibe, M., Eisses, M. J., Adams, T., Liston, D. E., Low, D. K.-W., Newman, S.-F., Kim, J., & Lee, S.-I. (2018). Explainable machine-learning predictions for the prevention of hypoxaemia during surgery. *Nature Biomedical Engineering*, 2(10), 749–760. <https://doi.org/10.1038/s41551-018-0304-0>
- 32) Martínez-García, D. N., Dalgo-Flores, V. M., Herrera-López, J. L., Analuisa-Jiménez, E. I., & Velasco-Acurio, E. F. (2019). Avances de la inteligencia artificial en salud. *Dominio de Las Ciencias*, 5(3), 603. <https://doi.org/10.23857/dc.v5i3.955>

- 33) Maubert, A., Birtwisle, L., Bernard, J. L., Benizri, E., & Bereder, J. M. (2019). Can machine learning predict resectability of a peritoneal carcinomatosis? *Surgical Oncology*, 29, 120–125. <https://doi.org/10.1016/j.suronc.2019.04.008>
- 34) Meyer, J. G., Urbanowicz, R. J., Martin, P. C. N., O'Connor, K., Li, R., Peng, P.-C., Bright, T. J., Tatonetti, N., Won, K. J., Gonzalez-Hernandez, G., & Moore, J. H. (2023). ChatGPT and large language models in academia: Opportunities and challenges. *BioData Mining*, 16(1), 20. <https://doi.org/10.1186/s13040-023-00339-9>
- 35) Microsoft. (2023a). *¿Qué es Microsoft Forms? - Soporte técnico de Microsoft.* <https://support.microsoft.com/es-es/office/-qu%C3%A9-es-microsoft-forms-6b391205-523c-45d2-b53a-fc10b22017c8>
- 36) Microsoft. (2023b). *Software de hojas de cálculo Microsoft Excel.* <https://www.microsoft.com/es/microsoft-365/excel>
- 37) Minitab. (2023). *Herramientas estadísticas, de análisis de datos y de mejora de procesos | Minitab.* <https://www.minitab.com/es-mx/>
- 38) Nielsen, J. (1994). Heuristic evaluation. In *Usability Inspection Methods* (Nielsen, J., and Mack, R.L.). John Wiley & Sons.
- 39) Noah, B., Keller, M. S., Mosadeghi, S., Stein, L., Johl, S., Delshad, S., Tashjian, V. C., Lew, D., Kwan, J. T., Jusufagic, A., & Spiegel, B. M. R. (2018). Impact of remote patient monitoring on clinical outcomes: An updated meta-analysis of randomized controlled trials. *Npj Digital Medicine*, 1(1), 20172. <https://doi.org/10.1038/s41746-017-0002-4>
- 40) OMS. (2023). *La OMS propugna un uso seguro y ético de la inteligencia artificial para la salud.* <https://www.who.int/es/news/item/16-05-2023-who-calls-for-safe-and-ethical-ai-for-health>
- 41) OpenAI. (2022). *Introducing ChatGPT.* <https://openai.com/blog/chatgpt>
- 42) Peng, J., Zou, K., Zhou, M., Teng, Y., Zhu, X., Zhang, F., & Xu, J. (2021). An Explainable Artificial Intelligence Framework for the Deterioration Risk Prediction of Hepatitis Patients. *Journal of Medical Systems*, 45(5), 61. <https://doi.org/10.1007/s10916-021-01736-5>
- 43) Perplexity AI. (2023). *Perplexity AI.* <https://www.perplexity.ai/>
- 44) Rivas, P., Holzmayer, K., Hernandez, C., & Grippaldi, C. (2018). Excitement and Concerns about Machine Learning-Based Chatbots and Talkbots: A Survey. *2018 IEEE*

- International Symposium on Technology and Society (ISTAS)*, 156–162.
<https://doi.org/10.1109/ISTAS.2018.8638280>
- 45) Rodríguez-Blanco, A., Simón-Cuevas, A., Hojas-Mazo, W., & Perea-Ortega, J. (2016). Extracción de datos enlazados desde información no estructurada aplicando técnicas de PLN y ontologías. *IWSW 2016. CEUR Workshop Proceedings*, 1797, 80–91.
- 46) Rouhiainen, L. (2018). *Inteligencia artificial*.
- 47) Sabol, P., Sinčák, P., Hartono, P., Kočan, P., Benetinová, Z., Blichárová, A., Verbóová, L., Štammová, E., Sabolová-Fabianová, A., & Jašková, A. (2020). Explainable classifier for improving the accountability in decision-making for colorectal cancer diagnosis from histopathological images. *Journal of Biomedical Informatics*, 109, 103523.
<https://doi.org/10.1016/j.jbi.2020.103523>
- 48) Saldarriaga, O. D. (2008). EXPERIENCIAS Y VIVENCIAS EN LA INGENIERÍA CLÍNICA. *Revista Ingeniería Biomédica*, 2(4), 10–14.
- 49) Sarmiento-Ramos, J. L. (2020). Aplicaciones de las redes neuronales y el deep learning a la ingeniería biomédica. *Revista UIS Ingenierías*, 19(4), 1–18.
<https://doi.org/10.18273/revuin.v19n4-2020001>
- 50) Stahl, B. C. (2021). *Artificial Intelligence for a Better Future: An Ecosystem Perspective on the Ethics of AI and Emerging Digital Technologies*. <https://doi.org/10.1007/978-3-030-69978-9>
- 51) Statista. (2023, March 23). *Infografía: OpenAI: los sectores que ya utilizan su software*. Statista Daily Data. <https://es.statista.com/grafico/29555/empresas-y-organizaciones-de-todo-el-mundo-que-utilizan-productos-de-openai>
- 52) Steinhubl, S. R., Wolff-Hughes, D. L., Nilsen, W., Iturriaga, E., & Califf, R. M. (2019). Digital clinical trials: Creating a vision for the future. *Npj Digital Medicine*, 2(1), 126, s41746-019-0203-0. <https://doi.org/10.1038/s41746-019-0203-0>
- 53) Sussman, L., Garcia-Robledo, J. E., Ordóñez-Reyes, C., Forero, Y., Mosquera, A. F., Ruíz-Patiño, A., Chamorro, D. F., & Cardona, A. F. (2022). Integration of artificial intelligence and precision oncology in Latin America. *Frontiers in Medical Technology*, 4, 1007822.
<https://doi.org/10.3389/fmedt.2022.1007822>
- 54) Trejos, O. (2022). *Lógica de programación*. Ediciones de la U.

- 55) Vecteezy. (2023). *ChatGPT Logo—Chat gpt Icon on Black and Green Background 21608788 Vector Art at Vecteezy*. <https://www.vecteezy.com/vector-art/21608788-chatgpt-logo-chat-gpt-icon-on-black-and-green-background>
- 56) Wall, J., & Krummel, T. (2020). The digital surgeon: How big data, automation, and artificial intelligence will change surgical practice. *Journal of Pediatric Surgery, 55*, 47–50. <https://doi.org/10.1016/j.jpedsurg.2019.09.008>
- 57) Werutsky, G., Barrios, C. H., Cardona, A. F., Albergaria, A., Valencia, A., Ferreira, C. G., Rolfo, C., De Azambuja, E., Rabinovich, G. A., Sposetti, G., Arrieta, O., Dienstmann, R., Rebelatto, T. F., Denninghoff, V., Aran, V., & Cazap, E. (2021). Perspectives on emerging technologies, personalised medicine, and clinical research for cancer control in Latin America and the Caribbean. *The Lancet Oncology, 22*(11), e488–e500. [https://doi.org/10.1016/S1470-2045\(21\)00523-4](https://doi.org/10.1016/S1470-2045(21)00523-4)
- 58) Writesonic. (2023). *ChatSonic—A ChatGPT Alternative With Superpowers and GPT-4*. https://writesonic.com/chat?ref=na84&fbclid=IwAR1XjKMaBGsS8-o_5msgBQDhn0AEY8TTrPtkfTmn5EimKxZqTNEKZwpN9qw&gclid=CjwKCAjwtuOIBhBREiwA7agf111C4dXcA65mDjQ9FZ7AoE16xaJrGcHYt2rn_oMP1D--zsMxa-WIDRoCgJgQAvD_BwE
- 59) Wu, H., Chen, W., Xu, S., & Xu, B. (2021). Counterfactual Supporting Facts Extraction for Explainable Medical Record Based Diagnosis with Graph Network. *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 1942–1955*. <https://doi.org/10.18653/v1/2021.naacl-main.156>
- 60) Wu, J. T., Leung, K., & Leung, G. M. (2020). Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: A modelling study. *The Lancet, 395*(10225), 689–697. [https://doi.org/10.1016/S0140-6736\(20\)30260-9](https://doi.org/10.1016/S0140-6736(20)30260-9)
- 61) Ye, Y., You, H., & Du, J. (2023). Improved Trust in Human-Robot Collaboration With ChatGPT. *IEEE Access, 11*, 55748–55754. <https://doi.org/10.1109/ACCESS.2023.3282111>
- 62) Zeroual, I., & Lakhouaja, A. (2018). Data science in light of natural language processing: An overview. *Procedia Computer Science, 127*, 82–91. <https://doi.org/10.1016/j.procs.2018.01.101>
- 63) Zhou, Z.-H. (2021). *Machine Learning*. Springer Nature.

ANEXOS

Anexo 1- Tres categorías de problemas éticos de la inteligencia artificial

1. Problemas derivados del aprendizaje automático	
Privacidad y protección de datos	Falta de privacidad
	Uso indebido de datos personales
	Problemas de seguridad
Fiabilidad	Falta de datos de calidad
	Falta de exactitud de los datos
	Problemas de integridad
Transparencia	Falta de responsabilidad
	Falta de transparencia
	Prejuicios y discriminación
	Falta de precisión de las recomendaciones predictivas
	Falta de precisión de las recomendaciones no individuales
	Daños a la integridad física
2. Vivir en un mundo digital	
Aspectos económicos	Desaparición de empleos
	Concentración de poder económico
	Coste de la innovación
Justicia y equidad	Propiedad controvertida de los datos
	Impacto negativo en el sistema judicial
	Falta de acceso a los servicios públicos
	Violación de los derechos humanos fundamentales de los usuarios finales
	Violación de los derechos humanos fundamentales en la cadena de suministro
	Impacto negativo en los grupos vulnerables
	Injusticia
Libertad	Falta de acceso y libertad de información
	Pérdida de capacidad de decisión humana
	Pérdida de libertad y autonomía individual
Problemas sociales más amplios	Relaciones de poder desiguales
	Asimetrías de poder
	Impacto negativo en la democracia
	Problemas de control y uso de datos y sistemas
	Falta de consentimiento informado
	Falta de confianza
	Posibilidad de uso militar

2. Vivir en un mundo digital	
Problemas sociales más amplios	Impacto negativo en la salud
	Reducción del contacto humano
	Impacto negativo en el medio ambiente
Problemas de incertidumbre	Efectos adversos imprevistos y no deseados
	Priorización de los problemas "equivocados"
	Potencial de uso delictivo y malintencionado
3. Problemas metafísicos	
	Conciencia de las máquinas
	"Despertar" de la IA
	Agentes morales autónomos
	Superinteligencia
	Singularidad
	Cambios en la naturaleza humana

Fuente: (Stahl, 2021)

Anexo 2 - La heurística de Nielsen comparada con la heurística final creada para el agente conversacional por Langevin et. al.

Heurística de Nielsen	Heurística de la fase 4
<p>Visibilidad del estado del sistema El sistema debe mantener siempre informados a los usuarios de lo que ocurre, mediante una retroalimentación adecuada en un plazo razonable.</p>	<p>Visibilidad del estado del sistema El sistema debe mantener siempre informados a los usuarios sobre lo que está ocurriendo, mediante una retroalimentación adecuada en un tiempo razonable, sin abrumar al usuario.</p>
<p>Correspondencia entre el sistema y el mundo real El sistema debe hablar el idioma de los usuarios, con palabras, frases y conceptos que les resulten familiares, en lugar de términos orientados al sistema. Debe seguir las convenciones del mundo real para que la información aparezca en un orden natural y lógico.</p>	<p>Correspondencia entre el sistema y el mundo real El sistema debe entender y hablar el lenguaje de los usuarios -con palabras, frases y conceptos familiares para el usuario y una voz adecuada- en lugar de términos orientados al sistema o terminología confusa. La información debe aparecer en un orden natural y lógico. Incluya elementos de diálogo que creen una conversación fluida mediante aperturas, orientación en mitad de la conversación y salidas elegantes.</p>
<p>Control y libertad del usuario Los usuarios suelen elegir funciones del sistema por error y necesitarán una "salida de emergencia" claramente señalizada para abandonar el estado no deseado sin tener que pasar por un largo diálogo. Soporta deshacer y rehacer.</p>	<p>Control y libertad del usuario A menudo, los usuarios eligen funciones del sistema por error y necesitan una opción para salir del estado no deseado sin tener que pasar por un diálogo prolongado. Supone deshacer y rehacer.</p>

Heurística de Nielsen	Heurística de la fase 4
<p>Consistencia y normas Los usuarios no deben preguntarse si distintas palabras, situaciones o acciones significan lo mismo. Siga las convenciones de la plataforma.</p>	<p>Consistencia y normas Los usuarios no deben preguntarse si distintas palabras, opciones o acciones significan lo mismo. Siga las convenciones de la plataforma para el diseño de elementos visuales y de interacción. Los usuarios también deben poder recibir respuestas coherentes, aunque comuniquen la misma función de múltiples formas (y modalidades). Dentro de la interacción, el sistema debe tener una voz, un estilo de lenguaje y una personalidad coherentes.</p>
<p>Prevención de errores Aún mejor que un buen mensaje de error es un diseño cuidadoso que evite que se produzca un problema. Elimine las condiciones propensas a errores o compruébelas y ofrezca a los usuarios una opción de confirmación antes de que lleven a cabo la acción.</p>	<p>Prevención de errores Incluso mejor que unos buenos mensajes de error es un diseño cuidadoso de la conversación y la interfaz para reducir la probabilidad de que se produzca un problema en primer lugar. Prepárate para pausas, interrupciones de la conversación, fallos en el diálogo, callejones sin salida o desvíos. Prevenga o elimine de forma proactiva posibles situaciones propensas a errores, y consulte y conforme con los usuarios antes de que cometan una acción.</p>
<p>Reconocer en lugar de recordar Minimice la carga de memoria del usuario haciendo visibles los objetos, las acciones y las opciones. El usuario no debe tener que recordar información de una parte del diálogo a otra. Las instrucciones de uso del sistema deben ser visibles o fácilmente recuperables siempre que sea necesario.</p>	<p>Ayuda y orientación El sistema debe guiar al usuario a lo largo del diálogo aclarando las capacidades del sistema. Las funciones de ayuda deben ser fáciles de recuperar y buscar, centrarse en la tarea del usuario, enumerar pasos concretos a realizar y no ser demasiado extensas. Las acciones y opciones deben ser visibles cuando proceda.</p>
<p>Flexibilidad y eficacia de uso Los aceleradores, invisibles para el usuario principiante, pueden a menudo acelerar la interacción para el usuario experto, de modo que el sistema pueda atender tanto a los usuarios inexpertos como a los experimentados. Permiten a los usuarios adaptar acciones frecuentes.</p>	<p>Flexibilidad y eficacia de uso Apoye interacciones flexibles en función del contexto de uso proporcionando a los usuarios la modalidad de entrada y salida y los hardware adecuados (o preferidos). Además, para que el sistema sea eficiente, se deben proporcionar aceleradores, como abreviaturas de comandos, que los principiantes no puedan ver pero que aceleren las interacciones de los expertos.</p>
<p>Diseño estético y minimalista Los diálogos no deben contener información irrelevante o que rara vez se necesite. Cada unidad de información adicional en un diálogo compite con las unidades de información relevantes y disminuye su visibilidad relativa.</p>	<p>Diseño estético, minimalista y atractivo Los diálogos no deben contener información irrelevante o que rara vez se necesite. Los elementos de interacción deben ser los necesarios para atraer al usuario y cumplir el objetivo del sistema. Las interfaces deben permitir interacciones breves y ampliar la conversación si el usuario lo desea.</p>

Heurística de Nielsen	Heurística de la fase 4
<p>Ayudar a los usuarios a reconocer, diagnosticar y recuperarse de los errores Los mensajes de error deben expresarse en lenguaje sencillo (sin códigos), indicar con precisión el problema y sugerir constructivamente una solución.</p>	<p>Ayudar a los usuarios a reconocer, diagnosticar y recuperarse de los errores Los mensajes de error deben expresarse en lenguaje sencillo (sin códigos), indicar con precisión el problema y sugerir una solución constructiva.</p>
<p>Ayuda y documentación Aunque es mejor que el sistema pueda utilizarse sin documentación, puede ser necesario proporcionar ayuda y documentación. Dicha información debe ser fácil de buscar, centrarse en la tarea del usuario, enumerar los pasos concretos a seguir y no ser demasiado extensa.</p>	
	<p>Conservación del contexto Mantener la preservación del contexto en relación con el tema de conversación dentro de la sesión y, si es posible, entre sesiones. Permitir que el usuario haga referencia a mensajes anteriores para interacciones posteriores a fin de respaldar las expectativas implícitas del usuario respecto a las conversaciones.</p>
	<p>Fiabilidad El sistema debe transmitir confianza garantizando la privacidad de los datos del usuario y siendo transparente y sincero con él. El sistema no debe afirmar falsamente que es humano.</p>

Fuente: (Langevin et al. 2021).

Anexo 3 - Encuesta de diagnóstico aplicada.

1. ¿Tiene acceso a una computadora en su lugar de trabajo? *

Si

No

2. ¿Tiene acceso a internet/WiFi en su lugar de trabajo? *

Si

No

3. ¿Alguna vez ha trabajado con chatbots? *

Un chatbot es un software basado en inteligencia artificial capaz de mantener una conversación en tiempo real por texto o por voz. Ejemplo: ChatGPT, Google Bard, Perplexity, etc.

Si

No

4. ¿Ha interactuado con otras inteligencias artificiales? *

Ejemplo: Siri, Alexa, Bixby, etc.

Si

No

5. Califique cada una de las siguientes afirmaciones según a su nivel de acuerdo o desacuerdo. *

	Totalmente en desacuerdo (1)	En desacuerdo (2)	Ni de acuerdo ni en desacuerdo (3)	De acuerdo (4)	Totalmente de acuerdo (5)
Estoy dispuesto a trabajar con chatbots.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Tengo un alto nivel de confianza en los chatbots.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Se me facilita tener un diálogo con los chatbots ya que preservan el contexto de la conversación.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Considero que los chatbots pueden ser una solución eficaz para realizar tareas rutinarias y repetitivas.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

6. ¿Cuál es su nivel de conocimiento en el uso de tecnologías de la información y la comunicación? *

- Novato
- Principiante
- Intermedio
- Avanzado
- Experto

7. ¿Considera que la infraestructura tecnológica en su lugar de trabajo le permitiría trabajar con chatbots? *

- Si
- No
- Necesita mejorar

8. ¿Qué factores influyen en su decisión de utilizar o no utilizar un chatbot? *

- Disponibilidad y acceso a la herramienta
- Fiabilidad y precisión de las respuestas
- Conveniencia y facilidad de uso
- Otras

9. ¿Considera que implementar el uso de chatbots ayudaría a facilitar su trabajo? *

- Si
- No
- Tal vez

Fuente: Elaboración propia.

Anexo 4 – Cuarta prueba con el chatbot ChatGPT.

The screenshot shows a web browser window with a ChatGPT chat interface. The prompt is: "Dame de el ejemplo de una rúbrica para mantenimiento de una bomba de infusión". The response is: "Por supuesto, aquí tienes un ejemplo de cómo podría ser una rúbrica para evaluar el mantenimiento de una bomba de infusión. Esta rúbrica se basa en criterios generales y puede ser adaptada según las especificaciones de tu organización y el modelo de la bomba. Cada criterio se evalúa en una escala de cumplimiento, que podría ser "Cumple", "Parcialmente Cumple" o "No Cumple"."

Rúbrica de Mantenimiento de Bomba de Infusión

Equipo Evaluado: [Nombre o modelo de la bomba de infusión]
Fecha de Evaluación: [Fecha de la evaluación]
Técnico Responsable: [Nombre del técnico responsable]

Criterio	Cumple	Parcialmente Cumple	No Cumple	Regenerate
----------	--------	---------------------	-----------	------------

Send a message

Free Research Preview. ChatGPT may produce inaccurate information about people, places, or facts. ChatGPT August 3 Version

Fuente: Elaboración propia.

Anexo 5 -Décima prueba con el chatbot ChatGPT.

The screenshot shows a web browser window with a ChatGPT chat interface. The prompt is: "Las lámparas UV (ultravioleta) son dispositivos utilizados para desinfectar superficies y ambientes al emitir radiación ultravioleta que puede dañar el ADN de microorganismos, como bacterias, virus y otros patógenos. Sin embargo, es importante destacar que el uso de lámparas UV para la desinfección en un quirófano u otros entornos de atención médica debe realizarse con precaución y siguiendo pautas específicas para garantizar la seguridad y la efectividad. Aquí hay una metodología general para el uso de lámparas UV de desinfección en un quirófano:". The response is: "1. Preparación y seguridad: * Asegúrate de que no haya personas presentes en el quirófano durante el proceso de desinfección con lámparas UV. * Usa equipo de protección personal (EPP), como gafas de protección y guantes, para evitar la exposición directa a la radiación UV. 2. Limpieza previa: * Limpia y descontamina las superficies del quirófano antes de usar la lámpara UV. La desinfección por UV es más efectiva en superficies limpias. 3. Posicionamiento: * Coloca la lámpara UV de manera que la radiación cubra todas las áreas que deseas"

1. **Preparación y seguridad:**

- * Asegúrate de que no haya personas presentes en el quirófano durante el proceso de desinfección con lámparas UV.
- * Usa equipo de protección personal (EPP), como gafas de protección y guantes, para evitar la exposición directa a la radiación UV.

2. **Limpieza previa:**

- * Limpia y descontamina las superficies del quirófano antes de usar la lámpara UV. La desinfección por UV es más efectiva en superficies limpias.

3. **Posicionamiento:**

- * Coloca la lámpara UV de manera que la radiación cubra todas las áreas que deseas

Send a message

Free Research Preview. ChatGPT may produce inaccurate information about people, places, or facts. ChatGPT August 3 Version

Fuente: Elaboración propia.

Anexo 6 - Segunda prueba con el chatbot Google Bard.

The screenshot shows the Google Bard chatbot interface in a browser window. The user's question is: "¿Cuál es el mantenimiento preventivo de una incubadora para cuidados neonatales?". The chatbot's response is: "El mantenimiento preventivo de una incubadora para cuidados neonatales es esencial para garantizar que el equipo funcione correctamente y que los bebés que se encuentran en ella estén seguros y protegidos. El mantenimiento preventivo debe realizarse de acuerdo con las instrucciones del fabricante, pero en general, incluye los siguientes pasos:" followed by a bulleted list of maintenance steps: "Limpieza diaria", "Inspección semanal", "Mantenimiento mensual", and "Mantenimiento anual". The interface includes a sidebar with chat history, a search bar, and a Windows taskbar at the bottom.

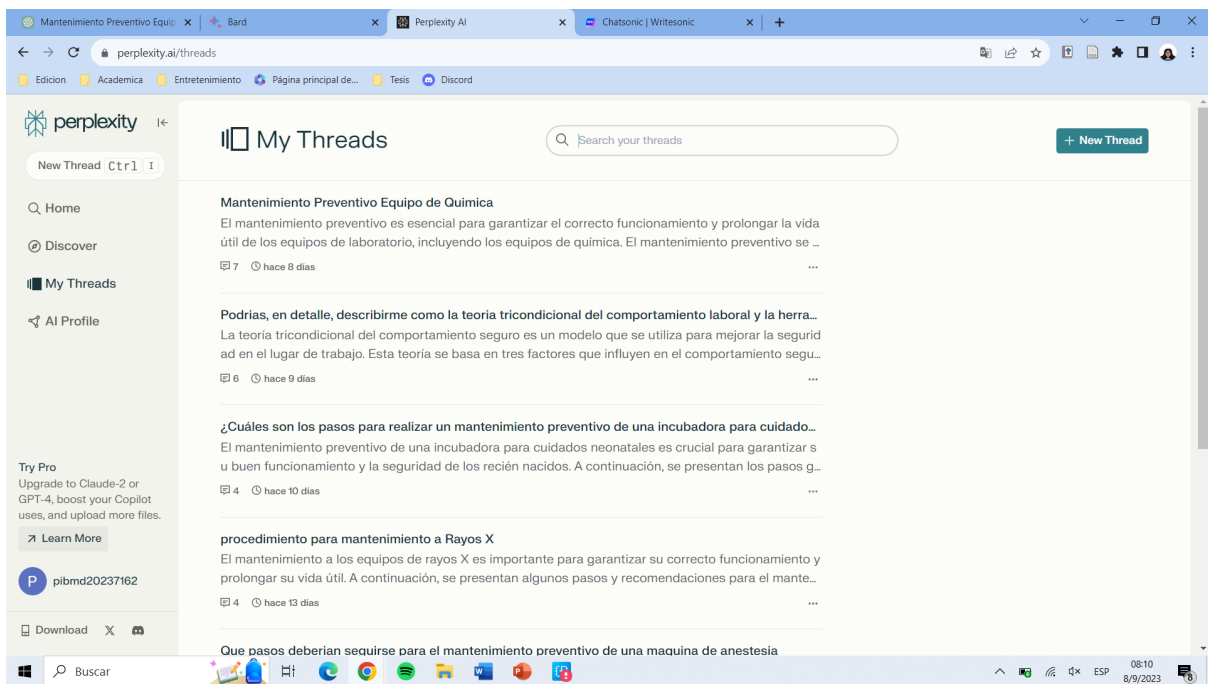
Fuente: Elaboración propia.

Anexo 7 - Quinta prueba con el chatbot Google Bard.

The screenshot shows the Google Bard chatbot interface in a browser window. The user's question is: "como realizar el balance del brazo de suspensión del microscopio quirurgico". The chatbot's response is: "El balance del brazo de suspensión del microscopio quirúrgico es un procedimiento importante para asegurar que el microscopio esté correctamente equilibrado y que el cirujano pueda operar con comodidad. Para realizar el balance del brazo de suspensión, siga estos pasos:" followed by a numbered list of steps: "1. Coloque el microscopio en su posición de trabajo.", "2. Ajuste la altura del microscopio para que esté a la altura cómoda del cirujano.", "3. Ajuste la inclinación del brazo de suspensión para que el microscopio esté alineado con la línea de visión del cirujano.", "4. Ajuste el ángulo del brazo de suspensión para que el microscopio esté en la posición deseada." The interface includes a sidebar with chat history, a search bar, and a Windows taskbar at the bottom.

Fuente: Elaboración propia.

Anexo 8 - Historial de búsquedas con el chatbot Perplexity AI.



Fuente: Elaboración propia.

Anexo 9 - Sexta prueba con el chatbot Perplexity AI.



Fuente: Elaboración propia.

Anexo 10 - Segunda prueba con el chatbot ChatSonic.

ChatSonic interface showing a user query: "por qué la imagen de mi ultrasonido no se proyecta en TV via HDMI". The chatbot response lists three possible reasons:

1. **Compatibilidad de resolución:** Asegúrate de que la resolución de salida del dispositivo de ultrasonido sea compatible con la resolución admitida por tu televisor. Si la resolución no coincide, es posible que la imagen no se muestre correctamente en la TV [3].
2. **Configuración de salida de video:** Verifica la configuración de salida de video en el dispositivo de ultrasonido. Algunos dispositivos pueden requerir que se cambie el formato de salida de video para que sea compatible con el televisor. Consulta el manual del dispositivo de ultrasonido para obtener instrucciones sobre cómo ajustar la configuración de salida de video.
3. **Cable HDMI defectuoso:** Asegurate de utilizar un cable HDMI de alta velocidad y en buen estado. Si el cable está dañado o no es de alta velocidad, es posible que no transmita correctamente la señal de video [2].

Fuente: Elaboración propia.

Anexo 11 - Novena prueba con el chatbot ChatSonic.

ChatSonic interface showing a user query: "Me puedes hacer un protocolo de mantenimiento del equipo Marca Abbott Cell dyn 1800, en forma de tabla, que indique descripción de actividad, tiempo de ejecución y periodicidad". The chatbot response provides a table:

Actividad de Mantenimiento	Tiempo de Ejecución	Periodicidad
Limpieza externa del equipo	15 minutos	Diario
Inspección visual de cables y conexiones	10 minutos	Semanal
Calibración del equipo	1 hora	Mensual
Verificación del rendimiento	30 minutos	Trimestral
Cambio de consumibles (reactivos, diluyentes, etc.)	30 minutos	Mensual
Actualización del software	1 hora	Anual

Es importante tener en cuenta que estos tiempos son aproximados y pueden variar dependiendo de la experiencia del técnico. Es fundamental seguir las instrucciones del fabricante para cada actividad de mantenimiento y contar con el equipo de protección personal adecuado.

Fuente: Elaboración propia.

Anexo 12 - Heurística aplicada.

1. ¿Qué chatbot utilizó? *

- ChatGPT
- Google Bard
- ChatSonic
- Perplexity

2. Califique cada una de las siguientes características según corresponda. *

	Totalmente insatisfecho (1)	Insatisfecho (2)	Ni satisfecho ni insatisfecho (3)	Satisfecho (4)	Totalmente satisfecho (5)
Visibilidad del estado del sistema	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Correspondencia entre el sistema y el mundo real	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Control y libertad del usuario	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Consistencia y normas	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Prevención de errores	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ayuda y orientación	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Flexibilidad y eficacia de uso	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Diseño estético, minimalista y atractivo	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ayudar a los usuarios a reconocer, diagnosticar y recuperarse de los errores	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Conservación del contexto	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Fiabilidad	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Fuente: Elaboración propia.

Anexo 13 - Encuesta de aceptación aplicada al finalizar las pruebas.

1. ¿Haría uso de los chatbots en sus labores diarias? *

Si

No

2. ¿Considera que implementar el uso de chatbots ayudaría a facilitar su trabajo? *

Si

No

Tal vez

3. ¿Qué tan frecuentemente está dispuesto a implementar el uso de chatbots en su trabajo? *

- Muy frecuentemente
- Frecuentemente
- Ocasionalmente
- Rara vez
- Nunca

4. ¿Con qué probabilidad recomendaría el uso de chatbots en una unidad de biomédica? *

- 1: Muy improbable
- 2: Poco probable
- 3: Neutral
- 4: Probable
- 5: Muy probable

1	2	3	4	5
---	---	---	---	---

Muy improbable

Muy probable

5. ¿Qué dificultades encontró al momento de utilizar los chatbots? *

6. Si tuviera que recomendar alguno de estos chatbots, ¿cuál elegiría? *

- ChatGPT
- Google Bard
- ChatSonic
- Perplexity

7. ¿Qué característica de los chatbots le pareció más valiosa? *

8. Califique cada una de las afirmaciones según su nivel de acuerdo o desacuerdo. *

	Totalmente en desacuerdo (1)	En desacuerdo (2)	Ni de acuerdo ni en desacuerdo (3)	De acuerdo (4)	Totalmente de acuerdo (5)
Estoy dispuesto a obtener información técnica o resolver problemas a través de chatbots.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Es importante contar con la disponibilidad de chatbots en el área biomédica.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Confío en la capacidad de un chatbot para proporcionar información precisa y relevante.

La privacidad y seguridad de los datos son una preocupación significativa para mí.

9. ¿Cuál es su nivel de confianza en los chatbots? *

- 1: Total desconfianza
- 2: Desconfianza
- 3: Neutral
- 4: Confianza
- 5: Total confianza

1	2	3	4	5
---	---	---	---	---

Total desconfianza

Total confianza

10. ¿Tiene alguna preocupación en cuanto a la privacidad y seguridad de los datos al interactuar con chatbots? *

- Si
- No

11. Explique su preocupación. *

12. En las actividades del Departamento de Biomédica, ¿En qué contexto o situaciones considera que los chatbots son particularmente útiles? *

13. ¿Qué funciones o características considera esenciales para que un chatbot sea efectivo en el ámbito de la ingeniería clínica? *

Fuente: Elaboración propia.

Anexo 14 - Matriz bibliográfica

Autores del estudio	Año	Título de la investigación	País	Área de conocimiento	Propósito de investigación
Abid Haleem, Mohd Javid, & Ravi Pratap Singh	2022	An era of ChatGPT as a significant futuristic support tool: A study on features, abilities, and challenges.	India	Ingeniería mecánica	Estudio sobre las características y capacidades específicas de ChatGPT.
Acosta, J. N., Falcone, G. J., Rajpurkar, P., & Topol, E. J.	2022	Multimodal biomedical AI.	Estados Unidos de América	Medicina e informática biomédica	Artículo de revisión que describe las principales aplicaciones posibles y los retos técnicos y analíticos que plantean los datos biomédicos.
Ávila, H. F., González, M. M., & Licea, S. M.	2020	La entrevista y la encuesta: ¿métodos o técnicas de indagación empírica?	Cuba	Ciencias pedagógicas	Argumentar entre la entrevista y la encuesta. Defiende la encuesta, como alternativo a la entrevista.
Chandel, S., Yuying, Y., Yujie, G., Razaque, A., Yang, G.	2019	Chatbot: Efficient and Utility-Based Platform	China	Ingeniería y ciencias computacionales	Estudio sobre el análisis de la tecnología de los chatbots e investigación de su desarrollo. Se discuten las características más destacadas de las técnicas de chatbot, y se analiza su importancia.
Chua, I. S., Gaziel-Yablowitz, M., Korach, Z. T., Kehl, K. L., Levitan, N. A., Arriaga, Y. E., Jackson, G. P., Bates, D. W., & Hassett, M.	2021	Artificial intelligence in oncology: Path to implementation.	Estados Unidos de América	Medicina	Estudio sobre los pasos para la implementación de la IA en la medicina oncológica.
Comisión Europea.	2019	Ethics guidelines for trustworthy AI	Unión europea	Informática	Establecer directrices para lograr una IA digna de confianza. Además de una lista de principios éticos, pretende ofrecer orientaciones sobre la forma en que dichos principios pueden ponerse en práctica en los sistemas sociotécnicos.
Corvalán, J. G.	2019	El impacto de la inteligencia artificial en el trabajo.	Brasil	Derecho económico y socioambiental	Análisis sobre como El progreso de la inteligencia artificial y la robótica está afectando el empleo, lo que lleva a la necesidad de que las personas se adapten a un entorno laboral en constante cambio.

Autores del estudio	Año	Título de la investigación	País	Área de conocimiento	Propósito de investigación
De Ramón Fernández, A., Ruiz Fernández, D., & Prieto Sánchez, M. T.	2019	A decision support system for predicting the treatment of ectopic pregnancies.	España	Informática y medicina	Desarrollar un sistema basado en algoritmos de inteligencia artificial para ayudar al personal clínico a elegir el tratamiento para la paciente.
García Benítez, V. H., & Ruvalcaba-Gómez, E. A.	2021	Análisis de las estrategias nacionales de inteligencia artificial en América Latina: Estudio de los enfoques de ética y de derechos humanos	México	Administración pública	Analizar las diferencias y similitudes de las características de las estrategias nacionales en materia de IA en el sector público, de los países de la región de América Latina: Argentina, Brasil, Chile, Colombia, México y Uruguay.
Hernández-Sampieri, R. & Mendoza, C	2018	Metodología de la investigación. Las rutas cuantitativa, cualitativa y mixta	México	Metodología de la investigación	Libro de texto relativo al estudio de la investigación se basa en tres enfoques: cuantitativo, cualitativo y mixto.
Janiesch, C., Zschech, P. & Heinrich, K.	2021	Machine learning and deep learning	Alemania	Ciencia de datos	Artículo de resumen de los fundamentos del aprendizaje automático y del aprendizaje profundo para generar una comprensión más amplia de la base metódica de los sistemas inteligentes actuales.
Jiang, Z., Xu, F. F., Araki, J., & Neubig, G.	2020	How can we know what language models know?	Estados Unidos de América	Informática	Artículo que trata de estimar con mayor precisión el conocimiento contenido en los modelos de lenguaje mediante el descubrimiento automático de mejores mensajes para utilizar en consultas.
Juárez, P. F.	2018	La importancia de la técnica de la entrevista en la investigación en comunicación y las ciencias sociales. Investigación documental. Ventajas y limitaciones.	México	Investigación para la comunicación	Análisis sobre la importancia de la entrevista para la investigación cualitativa y documental en las ciencias sociales.

Autores del estudio	Año	Título de la investigación	País	Área de conocimiento	Propósito de investigación
Karczewski, K. J., & Snyder, M. P.	2018	Integrative omics for health and disease.	Estados Unidos de América	Medicina	Revisión bibliográfica donde se destacan ejemplos de cómo la integración de datos provenientes de tecnologías ómicas permiten comprender, diagnosticar y tratar enfermedades. Se reconocen los desafíos técnicos y otros que surgen en la aplicación clínica de estas tecnologías integradoras.
Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., ... Kasneci, G.	2023	ChatGPT for good? On opportunities and challenges of large language models for education.	Alemania	Inteligencia artificial	Documento de opinión que presenta los posibles beneficios y retos de las aplicaciones educativas de los LLM desde el punto de vista de los estudiantes y los profesores.
Kavya, R., Christopher, J., Panda, S., & Lazarus, Y. B.	2021	Machine Learning and XAI approaches for Allergy Diagnosis.	India	Procesamiento de señales biomédicas	Presenta un sistema de diagnóstico de alergias asistido por ordenador capaz de gestionar las comorbilidades. Este sirve de fuente de información para que los médicos pueden utilizarlo para confirmar sus diagnósticos.
Kelleher, J. D.	2019	Deep learning	Irlanda	Información y comunicación	Libro de texto que funge como una introducción accesible y concisa, pero exhaustiva, a la tecnología fundamental en el corazón de la revolución de la inteligencia artificial.
Kojima, T., Gu, S. S., Reid, M., Matsuo, Y., & Iwasawa, Y.	2022	Large Language Models are Zero-Shot Reasoners.	Japón	Ciencia de datos	Estudio de base para las pruebas de razonamiento más exigentes. Destaca la importancia de explorar y analizar cuidadosamente el enorme conocimiento de cero que esconden los LLM antes de crear conjuntos de datos de ajuste.

Autores del estudio	Año	Título de la investigación	País	Área de conocimiento	Propósito de investigación
Langevin, R., Lordon, R., Avrahami, T., Cowan, B., Hirsch, T., & Hsieh, G.	2021	Heuristic Evaluation of Conversational Agents	Estados Unidos de América		Estudio para propuesta de un conjunto de heurísticas para agentes conversacionales adaptadas de la heurística de Nielsen.
Lundberg, S. M., Nair, B., Vavilala, M. S., Horibe, M., Eisses, M. J., Adams, T., Liston, D. E., Low, D. K.-W., Newman, S.-F., Kim, J., & Lee, S.-I.	2018	Explainable machine-learning predictions for the prevention of hypoxaemia during surgery	Estados Unidos de América	Ciencias Computacionales	Descripción sobre el desarrollo y las pruebas de un sistema basado en el aprendizaje automático que predice el riesgo de hipoxemia y ofrece explicaciones sobre los factores de riesgo en tiempo real durante la anestesia general.
Martínez-García, D. N., Dalgo-Flores, V. M., Herrera-López, J. L., Analuisa-Jiménez, E. I., & Velasco-Acurio, E. F	2019	Avances de la inteligencia artificial en salud.	Ecuador	Informática médica	Artículo de revisión, donde se evidencia el éxito en la búsqueda de IA y los beneficios sin precedentes.
Maubert, A., Birtwisle, L., Bernard, J. L., Benizri, E., & Bereder, J. M.	2019	Can machine learning predict resectability of a peritoneal carcinomatosis? Surgical Oncology.	Francia	Ginecología	El uso del Machine Learning para predecir carcinomatosis peritoneal.
Meyer, J. G., Urbanowicz, R. J., Martin, P. C. N., O'Connor, K., Li, R., Peng, P.-C., Bright, T. J., Tatonetti, N., Won, K. J., Gonzalez-Hernandez, G., & Moore, J. H.	2023	ChatGPT and large language models in academia: Opportunities and challenges.	Estados Unidos de América	Minería de datos	Discusión sobre los LLMs desde la perspectiva de lo académico con respecto a sus limitaciones y utilidad para la escritura académica, la educación y la programación.
Noah, B., Keller, M. S., Mosadeghi, S., Stein, L., Johl, S., Delshad, S., Tashjian, V. C., Lew, D., Kwan, J. T., Jusufagic, A., & Spiegel, B. M. R. (2018	Impact of remote patient monitoring on clinical outcomes: An updated meta-analysis of randomized controlled trials.	Estados Unidos de América.	Medicina Remota.	Monitorización remota de pacientes en los resultados clínicos.

Autores del estudio	Año	Título de la investigación	País	Área de conocimiento	Propósito de investigación
Peng, J., Zou, K., Zhou, M., Teng, Y., Zhu, X., Zhang, F., & Xu, J.	2021	An Explainable Artificial Intelligence Framework for the Deterioration Risk Prediction of Hepatitis Patients	China	Gastroenterología	Predicción del riesgo de deterioro de los pacientes con hepatitis mediante inteligencias artificiales.
Rivas, P., Holzmayer, K., Hernandez, C., & Grippaldi, C.	2018	Excitement and Concerns about Machine Learning-Based Chatbots and Talkbots: A Survey	Estados Unidos de América	Ciencias Computacionales	En este artículo presentamos los resultados de un experimento diseñado para encuestar la percepción de la gente sobre estos bots y cuánto confían en ellos.
Rouhiainen, L.	2018	Inteligencia artificial: 101 cosas que debes saber hoy sobre NUUESTRO FUTURO .	España	Ciencia y tecnología	Este libro incluye temas relacionados con la inteligencia artificial y también da acceso a un gran número de recursos, ideas y consejos que ayudan a entender cómo la inteligencia artificial va a cambiar nuestra vida.
Sabol, P., Sinčák, P., Hartono, P., Kočan, P., Benetinová, Z., Blichárová, A., Verbóová, L., Štammová, E., Sabolová-Fabianová, A., & Jašková, A.	2020	Explainable classifier for improving the accountability in decision-making for colorectal cancer diagnosis from histopathological images.	Eslovaquia y Japón	Ciencias Computacionales	Este artículo expone acerca de la estructura matemática del clasificador, que no está diseñado para ser utilizado como una herramienta de diagnóstico completamente automatizada sino como un sistema de apoyo para expertos médicos.
Sarmiento-Ramos, J. L.	2020	Aplicaciones de las redes neuronales y el deep learning a la ingeniería biomédica.	Colombia	Ingeniería de Sistemas e Informática	Revisión actualizada de las principales aplicaciones de las redes neuronales y el deep learning a la ingeniería biomédica en las ramas de la ómica, la imagenología, las interfaces cerebro-máquina y hombre-máquina, y la gestión y administración de la salud pública.

Autores del estudio	Año	Título de la investigación	País	Área de conocimiento	Propósito de investigación
Stahl, B. C.	2021	Artificial Intelligence for a Better Future: An Ecosystem Perspective on the Ethics of AI and Emerging Digital Technologies.	Estados Unidos de América	Ciencias Computacionales	Este libro presenta una respuesta innovadora a la pregunta al presentar una perspectiva diferente sobre la IA y sus consecuencias éticas.
Steinhubl, S. R., Wolff-Hughes, D. L., Nilsen, W., Iturriaga, E., & Califf, R. M.	2019	Digital clinical trials: Creating a vision for the future.	Estados Unidos de América	Ciencias Computacionales	Los ensayos clínicos son el mecanismo central para la evaluación imparcial de los avances propuestos en salud, atención médica y evaluación de opciones comparativas para enfoques de prevención, diagnóstico y tratamiento.
Sussman, L., Garcia-Robledo, J. E., Ordóñez-Reyes, C., Forero, Y., Mosquera, A. F., Ruíz-Patiño, A., Chamorro, D. F., & Cardona, A. F.	2022	Integration of artificial intelligence and precision oncology in Latin America.	Estados Unidos de América y Colombia	Oncología	Esta revisión presenta la IA en la atención del cáncer en América Latina, así como las ventajas y resultados prometedores que ha mostrado en este contexto sociodemográfico.
Trejos, O.	2022	Lógica de programación.	Colombia	Sistemas e informática	Expone los conceptos fundamentales para aprender a programar.
Wall, J., & Krummel, T.	2020	The digital surgeon: How big data, automation, and artificial intelligence will change surgical practice.	Estados Unidos de América	Automatización e inteligencia artificial en cirugía pediátrica	Demuestra como los cirujanos deben estar preparados para adoptar modalidades de entrenamiento más inteligentes.
Werutsky, G., Barrios, C. H., Cardona, A. F., Albergaria, A., Valencia, A., Ferreira, C. G., Rolfo, C., De Azambuja, E., Rabinovich, G. A., Sposetti, G., Arrieta, O., Dienstmann, R., Rebelatto, T. F., Denninghoff, V., Aran, V., & Cazap, E.	2021	Perspectives on emerging technologies, personalized medicine, and clinical research for cancer control in Latin America and the Caribbean	Estados Unidos de América	Automatización e inteligencia artificial en investigación clínica para el control del cáncer	En América Latina y el Caribe, la adopción de herramientas digitales de salud se encuentra en etapas tempranas y la calidad de los registros de cáncer, los registros de salud electrónicos y las bases de datos estructuradas son problemáticas.

Autores del estudio	Año	Título de la investigación	País	Área de conocimiento	Propósito de investigación
Wu, H., Chen, W., Xu, S., & Xu, B.	2021	Counterfactual Supporting Facts Extraction for Explainable Medical Record Based Diagnosis with Graph Network.	China	Informática médica	Este artículo propone un método de extracción de hechos de apoyo de un gráfico de granularidad múltiple contrafactual (CMGE) para extraer hechos de apoyo del propio EMR irregular sin bases de conocimiento externas.
Wu, J. T., Leung, K., & Leung, G. M.	2020	Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: A modelling study.	China	Estadística.	Modelado matemático de la posible propagación del brote de nCoV-2019.
Ye, Y., You, H., & Du, J.	2023	Improved Trust in Human-Robot Collaboration with ChatGPT.	Estados Unidos de América	Ciencias de Computacionales	Este documento explora el impacto de ChatGPT en la confianza en una tarea de ensamblaje de colaboración entre humanos y robots.
Zeroual, I., & Lakhouaja, A.	2018	Data science in light of natural language processing: An overview.	Marruecos	Ciencia de datos	El objetivo de este resumen es arrojar luz sobre el papel vital de los datos en diversos campos y ofrecer una mejor comprensión de los datos a la luz de la PNL. Describe lo que ocurre con los datos durante su ciclo de vida: fases de construcción, procesamiento, análisis y exploración.
Zhou, Z.-H.	2021	Machine Learning.	China	Ciencia de datos	Ofrece una introducción completa e imparcial a casi todos los aspectos del aprendizaje automático.

Fuente: Elaboración propia.