

Título del Trabajo:

Precisión de Agentes Conversacionales (AC) basados en Inteligencia Artificial Generativa (AIG) para recuperar Información Técnica para Apoyo en el Suministro Sillas de Ruedas Apropriadas (SRA)

Título del Trabajo en Inglés:

Accuracy of Conversational Agents based on Generative Artificial Intelligence to retrieve Technical Information to Support the Supply of Appropriate Wheelchairs

Nombre: ALBERTO ISAAC

Apellidos: PEREZ SANPABLO

ORCID: 0000-0003-0550-928X

País de Residencia: MEXICO

Área de Investigación: TECNOLÓGICA

Institución a la que Pertenece: INSTITUTO NACIONAL DE REHABILITACION GILLERMO IBARRA IBARRA

Área de Adscripción: Laboratorio de Análisis de Movimiento e Ingeniería de Rehabilitación

Correo Electrónico: albperezinr@gmail.com

Datos de los(as) coautores(as) del Trabajo

Alicia Meneses Peñaloza, Gerardo Rodríguez Reyes, IVETT QUIÑONES URIOSTEGUI, ALDO ALESSI MONTERO, LIDIA NUÑEZ CARRERA, ARTURO GONZALEZ MENDOZA, LEONARDO ELIU ANAYA CAMPOS, VIRGINIA BUEYES ROIZ, PARIS JOAQUIN VELSCO ACOSTA

Rehabilitación Pediátrica, INRLGII, MEXICO, aliciamp1@hotmail.com, 0000-0003-0550-928X

Laboratorio de Ortesis y Prótesis, INRLGII, MEXICO, grodriguezreyes@gmail.com,

LAMIR, INRLGII, MEXICO, iquinonesu@yahoo.com,

LAMIR, INRLGII, MEXICO, aldoalessi@gmail.com,

LAMIR, LGII, MEXICO, lidianc2012@gmail.com,

LAMIR, INRLGII, MEXICO, arturo.gonzalez.inr.lab.movi@gmail.com,

LAMIR, INRLGII, MEXICO, leonardo.anayac@gmail.com,

LAMIR, INRLGII, MEXICO, vbueyes@hotmail.com,

LAMIR, INRLGII, MEXICO, parisvelasco.inr@gmail.com,

Palabras en Español:

Inteligencia Artificial, Silla de Ruedas, Sistema de Apoyo a la Toma de Decisiones

Palabras en Inglés:

Artificial Intelligence, Wheelchair, Decision Support Systems

Título del Trabajo:

Precisión de Agentes Conversacionales (AC) basados en Inteligencia Artificial Generativa (AIG) para recuperar Información Técnica para Apoyo en el Suministro Sillas de Ruedas Apropriadas (SRA)

Título del Trabajo en Inglés:

Accuracy of Conversational Agents based on Generative Artificial Intelligence to retrieve Technical Information to Support the Supply of Appropriate Wheelchairs

Área de Investigación:

Laboratorio de Análisis de Movimiento e Ingeniería de Rehabilitación

Introducción:

El suministro de SRA involucra procesos que consumen tiempo y requieren personal calificado, cuya carencia es un problema. La AIG, cuyo uso es cada vez más popular, puede apoyar a mejorar la eficiencia en tareas como la síntesis de información, pero su desempeño varía según el área y la interacción. Por lo que la evaluación de su desempeño es fundamental. El usuario interactúa con la IAG mediante AC a través de instrucciones o "prompts", cuya ingeniería influye en la precisión de los resultados. Esta es la primera investigación que evalúa AIG en el suministro de SRA. Esta investigación busca desarrollar una metodología innovadora con integración de AIG.

Objetivo:

El objetivo es evaluar la capacidad y precisión de AC para recuperar información técnica específica necesaria para el suministro de SRA, así como el impacto del uso de "prompts" avanzados.

Metodología:

Dos expertos certificados en suministro de SRA seleccionaron 3 AC de acceso libre (ChatGPT 3.5, Bing Chat y Google Bard) y 11 modelos de SRA disponibles localmente. Las consultas se repitieron 2 veces utilizando los mismos prompts estandarizados con todos los AC. Se utilizó un prompt de plantilla con 32 características técnicas, creado a partir del formato para SRA validado por la OMS, para asegurar la recuperación de toda la información técnica necesaria. En la segunda consulta, se añadió un prompt tipo verificación de hechos estandarizado para medir su impacto en la precisión de los datos recuperados. Se evaluó la capacidad de los AC para recuperar la información disponible en línea y seguir instrucciones detalladas. La validación de las respuestas de los AC fue realizada por los 2 expertos quienes evaluaron la cantidad y precisión de la información recuperada. Se calculó el índice de concordancia Kappa entre agentes. Esta investigación no involucró a humanos ni recolectó datos personales.

Resultados:

ChatGPT-3.5 no pudo recuperar información debido a la falta de acceso a internet. Google Bard y Bing Chat sí recuperaron información según el prompt tipo plantilla. Google Bard recuperó el 79±2% de la información con una precisión promedio del 30±13%, mientras que Bing Chat recuperó el 61±31% de la información con una precisión promedio del 41±2%. La adición de prompts de verificación de hechos solo modificó el 16±10% de las primeras respuestas proporcionadas por Bing Chat y no tuvo efecto en Google Bard. El índice de concordancia kappa promedio entre Google Bard y Bing Chat fue

de 0.37 ± 0.19 . Estos resultados demuestran que, aunque los agentes pueden recuperar información técnica específica, la precisión y la capacidad de ajuste de los prompts varían significativamente entre ellos. Los resultados muestran la eficacia relativa y las limitaciones de los AC en la recuperación de información técnica para el suministro de SRA.

Conclusiones:

Los resultados muestran que algunos AC pueden recuperar información relevante para el suministro de SRA, pero su precisión varía, haciendo esencial verificar la interacción y respuestas. Perspectivas de investigación son explorar otras áreas de dispositivos médicos, entrenamiento con bases de conocimiento específicas y uso de ingeniería de prompts.