Evaluación contextualizada de accesibilidad y usabilidad de IA generativa en educación mediante tareas DUA y observación moderada

Rosana Montes[1][0000−0002−0183−044X], Noe Zermeño-Mejía[2][0000−0002−7118−6151] Jerónimo Duran-Curiel[2][0000−0002−0368−0035], and Francisco Herrera[1][0000−0002−7283−312X]

[1] Instituto Interuniversitario de Investigación en Data Science and Computational Intelligence, DaSCI. Universidad de Granada, Granada, España
rosana@ugr.es, fherrera@decsai.ugr.es

[2] Centro Universitario de Tlaquepaque. Universidad de Guadalajara. Jalisco, México.
{noe.zermeno,jeronimo.duran}@academicos.udg.mx

Resumen La incorporación de herramientas basadas en inteligencia artificial generativa en educación plantea nuevas oportunidades de inclusión, pero también desafíos relacionados con su accesibilidad y usabilidad para estudiantes de educación superior. Este trabajo presenta una evaluación contextualizada en un entorno educativo real, combinando observación moderada y recogida de datos mediante cuestionarios. Los resultados muestran una valoración generalmente positiva de la experiencia y un elevado grado de autonomía en el acceso a la herramienta. No obstante, se identificaron dificultades relacionadas con la formulación de instrucciones, la comprensión inicial de las consignas y determinadas barreras de accesibilidad en usuarios que emplean tecnologías de asistencia. Los resultados evidencian que la accesibilidad de la IA generativa depende tanto de la tecnología como del contexto de uso y del diseño de las actividades educativas.

Keywords: Inteligencia Artificial Generativa · Usabilidad y experiencias de usuario · Educación Inclusiva.

1. Introducción

Herramientas de inteligencia artificial generativa como ChatGPT, Microsoft Copilot o Gemini han comenzado a incorporarse progresivamente en contextos educativos, profesionales y cotidianos, facilitando tareas de búsqueda de información, generación de contenidos, apoyo al aprendizaje y asistencia personalizada. Sin embargo, emergen también importantes desafíos relacionados con la accesibilidad y la usabilidad real de dichas plataformas. Diferentes estudios y comunidades de usuarios han señalado problemas asociados a la compatibilidad con tecnologías de asistencia, navegación mediante teclado, interfaces dinámicas, sobrecarga cognitiva y consistencia de la interacción conversacional. En entornos de enseñanza, esta problemática adquiere una relevancia especial [1], situación

que resulta especialmente significativa cuando las herramientas son empleadas por estudiantes con perfiles diversos y necesidades heterogéneas de interacción digital.

En este contexto, el presente trabajo analiza la accesibilidad y usabilidad de un sistema conversacional basado en IA generativa utilizado en un entorno educativo real. El estudio se centra en la realización de una evaluación contextualizada mediante pruebas de usabilidad orientadas a identificar barreras de interacción, dificultades de acceso y factores que afectan la experiencia de usuario durante el desarrollo de tareas educativas concretas.

Las principales contribuciones de este trabajo son las siguientes:

- Revisar el estado actual de la accesibilidad y usabilidad en IA generativa conversacional.
- Analizar la limitada transparencia existente en la documentación pública sobre accesibilidad en plataformas LLM.
- Proponer un enfoque de evaluación contextualizado para el análisis de accesibilidad y usabilidad en entornos educativos.
- Presentar una prueba experimental basada en tareas de interacción con un sistema conversacional para el mejor conocimiento de las necesidades de los estudiantes.

El resto del artículo se organiza de la siguiente manera. La Sección 2 presenta el marco teórico y el estado del arte relacionado con accesibilidad digital, usabilidad conversacional e IA generativa en educación. La Sección 3 describe los objetivos y preguntas de investigación. La Sección 4 detalla la metodología empleada y el diseño de la evaluación experimental. La Sección 5 presenta los resultados obtenidos, mientras que la Sección 6 discute las implicaciones y limitaciones del estudio. Finalmente, la Sección 7 recoge las conclusiones y líneas futuras de trabajo.

2. Marco teórico y estado del arte

2.1. Accesibilidad digital e inteligencia artificial

La accesibilidad digital persigue que todas las personas puedan percibir, comprender, navegar e interactuar con los sistemas digitales, independientemente de sus capacidades o del uso de tecnologías de apoyo. Diversos estudios han señalado que las herramientas de IA generativa pueden introducir nuevas barreras relacionadas con interfaces, navegación y compatibilidad con tecnologías de asistencia [4,2]. Aunque la tecnología destaca en escala y eficiencia, a menudo carece de la sabiduría y el contexto necesarios para generar interacciones verdaderamente empáticas y fiables [5].

Aunque las Pautas de Accesibilidad para el Contenido Web (WCAG 2.2) [3] constituyen el principal referente para evaluar la accesibilidad técnica de los entornos digitales, la interacción con sistemas de inteligencia artificial introduce retos adicionales relacionados con la comprensión de las respuestas generadas, la formulación de consultas eficaces y la autonomía de los usuarios durante la

realización de tareas educativas. En consecuencia, la evaluación de estas herramientas debe integrar tanto aspectos técnicos de accesibilidad como factores de usabilidad asociados a la interacción en contexto.

Desde esta perspectiva, el presente trabajo considera la accesibilidad como una propiedad que emerge de la interacción entre el estudiante, la herramienta de inteligencia artificial y la actividad de aprendizaje. Este enfoque justifica la utilización de una metodología basada en observación moderada durante tareas educativas reales, permitiendo identificar barreras, necesidades de apoyo y oportunidades de mejora que no pueden detectarse mediante auditorías técnicas aisladas.

2.2. Usabilidad y experiencia de usuario en sistemas conversacionales

La usabilidad se entiende como el grado en que un producto puede ser utilizado para alcanzar objetivos específicos con eficacia, eficiencia y satisfacción en un contexto de uso determinado [6]. En los sistemas conversacionales basados en IA generativa, esta definición se amplía, ya que la calidad de la interacción depende tanto de la interfaz como de la capacidad del usuario para formular consultas, interpretar las respuestas y mantener un diálogo efectivo con el sistema.

En este contexto, Borsci et al. [7] proponen la *Chatbot Usability Scale* (BUS-15), una escala específica para evaluar la usabilidad percibida en asistentes conversacionales, incorporando dimensiones como la claridad de las respuestas, la facilidad de interacción, la percepción de control, la coherencia conversacional y la utilidad percibida. De forma complementaria, trabajos recientes destacan la necesidad de considerar aspectos propios de la interacción con IA, como la transparencia del sistema, la gestión de la incertidumbre, la recuperación ante errores y la comunicación de sus limitaciones [8].

La literatura también ha explorado metodologías alternativas para evaluar la experiencia de usuario, tanto mediante el análisis de reseñas espontáneas en plataformas digitales [9] como mediante estudios de uso en contextos reales [10]. En conjunto, estos trabajos coinciden en que la evaluación de la usabilidad de sistemas basados en IA no puede limitarse a indicadores tradicionales, sino que debe considerar factores como la confianza, la carga cognitiva, la interpretabilidad y el contexto de interacción.

En entornos de educación inclusiva, la usabilidad está estrechamente vinculada a la accesibilidad, ya que una herramienta técnicamente accesible puede seguir presentando dificultades para comprender las respuestas, elaborar consultas eficaces o completar una actividad de aprendizaje de forma autónoma. Por ello, en este trabajo ambas dimensiones se evalúan de forma integrada mediante observación moderada durante tareas educativas contextualizadas, permitiendo identificar barreras y necesidades de apoyo difíciles de detectar mediante evaluaciones exclusivamente técnicas o cuestionarios posteriores.

2.3. Accesibilidad declarada en plataformas basadas en IA generativa

A pesar del rápido crecimiento de las plataformas LLM, la información pública disponible sobre sus características de accesibilidad continúa siendo limitada, heterogénea y poco estandarizada. Esta situación dificulta la comparación objetiva entre herramientas. Es por ello que algunas compañías tecnológicas han comenzado a incorporar referencias explícitas a accesibilidad en su documentación oficial. Microsoft, por ejemplo, dispone de recursos específicos sobre accesibilidad para Copilot [11] y destaca el potencial de Microsoft 365 Copilot para favorecer entornos más inclusivos mediante herramientas de asistencia y productividad [12]. OpenAI también ha presentado iniciativas relacionadas con accesibilidad visual y multimodalidad [13,18].

Sin embargo, foros especializados continúan reportando dificultades de accesibilidad en interfaces conversacionales basadas en IA. La ausencia de métricas homogéneas, auditorías públicas sistemáticas y evaluaciones centradas en usuarios dificulta determinar hasta qué punto estas herramientas resultan realmente accesibles en contextos de uso reales, especialmente en entornos educativos inclusivos.

2.4. IA generativa y educación inclusiva

La IA generativa ofrece un potencial significativo para favorecer entornos de aprendizaje más inclusivos mediante contenidos adaptados, asistencia personalizada y apoyo multimodal. Sin embargo, la evidencia indica que estas tecnologías no pueden automatizar completamente la accesibilidad y continúan requiriendo supervisión humana para garantizar una experiencia realmente inclusiva [14].

Su incorporación al ámbito educativo también puede introducir nuevas barreras cuando las herramientas no consideran criterios de accesibilidad, mediación pedagógica y adaptación a la diversidad del alumnado [1]. Entre los principales riesgos identificados se encuentran los sesgos algorítmicos, la limitada adaptación a perfiles diversos, la falta de formación docente y la dependencia tecnológica [4].

Asimismo, se ha observado que modelos como ChatGPT generan, por defecto, respuestas con un nivel de legibilidad propio de estudiantes universitarios, lo que puede incrementar la carga cognitiva de alumnado con dificultades de aprendizaje o baja alfabetización y limitar la comprensión de los contenidos [15].

En este contexto, el Diseño Universal para el Aprendizaje (DUA) constituye un marco adecuado para orientar el uso educativo de la IA generativa, al promover múltiples formas de representación, acción y expresión que permiten adaptar la interacción a las necesidades de un alumnado diverso.

3. Objetivo y preguntas de investigación

El objetivo de este trabajo es proponer y aplicar una metodología de evaluación contextualizada que permita analizar la accesibilidad y la usabilidad de

herramientas de inteligencia artificial generativa durante actividades reales de aprendizaje en educación inclusiva, mediante observación moderada y un instrumento estructurado basado en los principios del Diseño Universal para el Aprendizaje.

A partir de este objetivo se plantean las siguientes preguntas de investigación:

PI1. ¿Qué barreras de accesibilidad y usabilidad experimentan los estudiantes al utilizar herramientas de inteligencia artificial generativa durante actividades educativas reales?

PI2. ¿En qué medida la observación moderada permite identificar necesidades de apoyo y patrones de interacción que no pueden detectarse mediante evaluaciones exclusivamente técnicas o cuestionarios posteriores?

PI3. ¿Es la metodología propuesta un procedimiento transferible para evaluar herramientas de inteligencia artificial generativa en contextos de educación inclusiva?

4. Metodología

4.1. Enfoque metodológico

El presente estudio adopta un enfoque exploratorio y contextualizado para evaluar la accesibilidad y usabilidad de un sistema conversacional basado en IA generativa en un entorno educativo real. La metodología combina observación moderada de la interacción, análisis de tareas y recogida de percepciones de los participantes durante el uso de la herramienta.

La elección de una prueba de usabilidad con usuarios reales responde a la necesidad de analizar la experiencia completa de interacción humano–IA en situaciones próximas al contexto cotidiano de uso. Las pruebas de usabilidad contextualizadas identifican problemas de interacción, barreras cognitivas y dificultades de uso que difícilmente emergen mediante aproximaciones exclusivamente técnicas o evaluaciones automatizadas [16,17]. Este enfoque resulta especialmente relevante en sistemas conversacionales basados en IA generativa, donde la calidad de la experiencia depende de factores dinámicos asociados a la comprensión del lenguaje natural, la continuidad del diálogo y la interpretación de respuestas probabilísticas.

La metodología se alinea con las recomendaciones existentes para la evaluación de interfaces, que subrayan la necesidad de considerar no solo la efectividad funcional del sistema, sino también aspectos relacionados con transparencia, interpretabilidad, recuperación ante errores y experiencia global de interacción conversacional [7,8]. Asimismo, el estudio toma como referencia conceptual las dimensiones de eficacia, eficiencia y satisfacción propuestas por la norma ISO 9241-11 [6], adaptándolas al análisis de sistemas conversacionales en contextos educativos inclusivos.

Dado el carácter exploratorio del estudio, el objetivo no es obtener resultados generalizables, sino identificar barreras y necesidades de adaptación que contribuyan a una evaluación más contextualizada de la IA generativa en educación.

4.2. Contexto educativo y participantes

En la prueba participaron estudiantes pertenecientes a diferentes perfiles educativos ($n = 84$), con edades comprendidas entre 18 y 23 años. La selección de participantes respondió a un enfoque de diversidad funcional y heterogeneidad de experiencia tecnológica, considerando distintos niveles de familiaridad con herramientas digitales y sistemas basados en IA generativa.

Con el objetivo de contextualizar adecuadamente la interacción observada, se recogió información relacionada con experiencia previa en el uso de asistentes conversacionales, frecuencia de utilización de herramientas de IA, dispositivos empleados durante la sesión y modalidades habituales de acceso e interacción digital. Asimismo, se contempló la posible utilización de tecnologías de asistencia y la presencia de necesidades de apoyo comunicadas voluntariamente por los participantes o identificadas durante la observación contextual de la actividad.

Dado el carácter exploratorio y contextualizado del estudio, el objetivo no consistió en obtener una muestra estadísticamente representativa, sino en observar patrones de interacción, barreras de accesibilidad y necesidades de adaptación relevantes para el análisis de sistemas conversacionales basados en IA generativa en entornos educativos inclusivos.

4.3. Diseño de la prueba

La evaluación experimental se diseñó como una prueba de usabilidad moderada y contextualizada orientada a observar la interacción real entre estudiantes y un sistema conversacional basado en IA generativa en un entorno educativo cotidiano. A diferencia de evaluaciones realizadas en contextos de laboratorio o mediante tareas artificiales altamente controladas, el estudio se planteó en una situación próxima al uso habitual que podría producirse durante el inicio de una actividad académica ordinaria.

La actividad se realizó con herramientas conversacionales de IA generativa disponibles para el alumnado. El uso de la herramienta fue registrado de forma autorreportada, identificándose principalmente ChatGPT y Gemini, además de menciones puntuales a Copilot, DeepSeek y Be My Eyes. La actividad fue diseñada para requerir capacidades de búsqueda, comprensión, reformulación y comunicación de información mediante lenguaje natural.

La consigna propuesta a los participantes consistió en explicar un concepto tecnológico cotidiano —como el uso de *cookies* en Internet— utilizando lenguaje comprensible y adaptado a contextos cercanos al alumnado.

El diseño de la actividad tomó como referencia principios del Diseño Universal para el Aprendizaje (DUA), proporcionando diferentes formas de acceso a la información, distintas posibilidades de interacción y múltiples opciones de expresión de resultados. Esta aproximación permitió observar no solo la capacidad técnica de acceso a la herramienta, sino también aspectos relacionados con comprensión de consignas, autonomía en la interacción, necesidad de apoyos y capacidad de interpretación crítica de las respuestas generadas por la IA.

Desde el punto de vista observacional, la prueba se estructuró en torno a tres dimensiones principales de análisis:

- **Acceso:** condiciones de interacción inicial con la herramienta, incluyendo navegación, uso de tecnologías asistivas, modalidades de entrada y posibles barreras técnicas de acceso.
- **Comprensión:** capacidad para interpretar consignas, comprender las respuestas generadas por la IA y mantener una interacción conversacional coherente durante la realización de la tarea.
- **Expresión:** formas utilizadas por los participantes para elaborar, reformular o comunicar las respuestas obtenidas, así como estrategias de validación o revisión crítica de la información generada por el sistema.

La prueba fue moderada por los investigadores para observar el comportamiento de los participantes y registrar incidencias relacionadas con accesibilidad, interacción y experiencia de uso. Esta modalidad permitió identificar dificultades emergentes, necesidades de adaptación y estrategias espontáneas de interacción.

El estudio no se orientó a realizar una categorización clínica o diagnóstica, sino a analizar la diversidad de formas de interacción observadas durante el uso de la herramienta. Se consideraron el empleo de tecnologías de asistencia, la necesidad de apoyos adicionales, los diferentes ritmos de interacción y las variaciones en comprensión o autonomía durante la tarea.

4.4. Instrumento de recogida

El instrumento de recogida de datos diseñado para este estudio combina elementos de observación contextual y percepción subjetiva del usuario, permitiendo analizar de forma integrada aspectos de accesibilidad, usabilidad e interacción humano–IA en entornos educativos inclusivos.

El formulario —implementado mediante *Google Forms*— incorpora variables relacionadas con autonomía, comprensión de consignas, necesidad de apoyos, barreras de interacción, adaptación del lenguaje, uso de tecnologías asistivas (AT, *Assistive Technologies*) y capacidad de revisión crítica de las respuestas generadas por IA. Esta aproximación resulta especialmente adecuada para contextos NEAE, ya que no se limita a evaluar únicamente el cumplimiento técnico de criterios de accesibilidad, sino que considera también factores cognitivos, comunicativos y pedagógicos asociados a la experiencia real de uso.

Asimismo, el instrumento facilita la identificación de necesidades de adaptación y proporciona evidencia contextualizada sobre cómo las herramientas de IA generativa pueden favorecer o dificultar la participación autónoma de estudiantes con perfiles funcionales diversos.

El instrumento final estuvo compuesto por 22 ítems, incluyendo variables de identificación contextual y percepción de uso. El tiempo estimado de cumplimentación tras la prueba es de aproximadamente 10–15 minutos.

4.5. Procedimiento

La moderación de las sesiones fue realizada por colaboradores externos vinculados al ámbito universitario, siguiendo un protocolo común previamente definido. Antes del inicio de la actividad, los participantes recibieron información

general sobre el propósito del estudio, el carácter voluntario de la participación, el tratamiento anonimizado de los datos recogidos, la posibilidad de abandonar la prueba en cualquier momento y el tiempo aproximado requerido para completar la actividad.

La tarea propuesta consistió en interactuar con el sistema conversacional para resolver una actividad educativa sencilla basada en comprensión y reformulación de información mediante lenguaje natural. Durante la interacción, los participantes pudieron utilizar libremente sus estrategias habituales de acceso e interacción, incluyendo tecnologías de asistencia, modalidades alternativas de entrada o apoyos personales cuando resultara necesario.

El procedimiento de recogida de información combinó dos fuentes complementarias de datos:

- **Formulario del participante:** orientado a recoger percepciones subjetivas relacionadas con facilidad de uso, comprensión de respuestas, autonomía durante la tarea, utilidad percibida y dificultades experimentadas durante la interacción con la herramienta.
- **Formulario de observación del moderador:** destinado a registrar variables contextuales relevantes para la evaluación de accesibilidad y usabilidad, incluyendo tipo de dispositivo utilizado, entorno de interacción, modalidades de acceso, uso de tecnologías de asistencia, experiencia tecnológica previa y necesidades de apoyo observadas o comunicadas voluntariamente durante la sesión.

Además de las respuestas obtenidas mediante formularios, los moderadores registraron incidencias relevantes relacionadas con barreras de interacción, dificultades de comprensión, necesidad de reformulación de consignas, problemas técnicos o estrategias espontáneas de adaptación utilizadas por los participantes durante el desarrollo de la actividad.

4.6. Ética y reproducibilidad

El estudio fue desarrollado en un contexto educativo de educación superior real mediante una actividad supervisada orientada a analizar la interacción entre estudiantes y herramientas basadas en IA generativa. La participación se realizó de forma informada y voluntaria, utilizando procedimientos de consentimiento adaptados al contexto de la evaluación. Con el objetivo de favorecer la transparencia y reproducibilidad del trabajo, los materiales metodológicos empleados —incluyendo protocolo de sesión, definición de tareas, formularios de observación y modelo de consentimiento informado— se encuentran disponibles en el repositorio público del proyecto [19].

Durante la evaluación no se realizaron categorizaciones diagnósticas individuales ni se recopilaron datos personales innecesarios. Las observaciones se centraron en patrones de interacción, barreras de accesibilidad, necesidades de apoyo y experiencia de uso durante el desarrollo de la actividad.

5. Resultados

La muestra obtenida estuvo compuesta por 84 participantes con un contexto educativo heterogéneo y representativo de diferentes perfiles académicos. De ellos, 83 no reportaron necesidades específicas de apoyo educativo (NEAE), mientras que una persona indicó presentar NEAE y utilizar tecnologías de apoyo (AT). Debido a esta distribución claramente desequilibrada, no resulta metodológicamente adecuado realizar comparaciones entre grupos.

Los resultados resumidos en la Tabla 1, muestran que la interacción con sistemas conversacionales no depende únicamente del acceso a la plataforma, sino también de la capacidad para interpretar la tarea, formular peticiones claras y utilizar críticamente la respuesta generada. En cambio, las mayores dificultades se concentraron en dimensiones cognitivas y comunicativas, especialmente en la comprensión inicial de la consigna y en la formulación de instrucciones dirigidas a la IA. Estas barreras serán descritas en la Sección 5.1.

En consecuencia, se identifican como necesidades relevantes el diseño de consignas más explícitas, la inclusión de ejemplos, la fragmentación de las tareas en pasos y la orientación básica en estrategias de *prompting*.

Cuadro 1. Resultados de la prueba de usabilidad moderada.

Dimensión observada	**Sin apoyo**	**Con apoyo**	**No logrado**
Acceso a la herramienta	75 (90.4 %)	8 (9.6 %)	0 (0.0 %)
Comprensión de la consigna	55 (66.3 %)	23 (27.7 %)	5 (6.0 %)
Formulación de la petición a la IA	43 (51.8 %)	32 (38.6 %)	8 (9.6 %)
Comprensión de la respuesta generada	69 (83.1 %)	13 (15.7 %)	1 (1.2 %)
Uso de la respuesta para completar la tarea	65 (78.3 %)	15 (18.1 %)	3 (3.6 %)

Entre los participantes, se incorporó al experimento una estudiante con discapacidad visual usuaria de lector de pantalla. A pesar de ser ciega de nacimiento, realizó la actividad sin el apoyo de otra persona para acceder a la herramienta. Las barreras que menciona en el uso de la IA en tareas educativas se centra en la creación de elementos visuales (gráficos, imágenes o tablas), la falta de adaptación para lector de pantallas de herramientas en la interfaz que impide la navegación entre las funciones o su traducción al español.

5.1. Discusión y barreras emergentes

A continuación, se presentan las principales observaciones realizadas durante la prueba de usabilidad moderada y los patrones de interacción identificados entre los participantes. La mayoría logró ingresar de manera autónoma, por lo que la principal barrera no se ubicó en el acceso técnico, sino en las etapas posteriores de comprensión, interacción y uso pedagógico de la IA generativa.

Respecto a la comprensión de la consigna, algunos estudiantes requirieron apoyo para interpretar con claridad la actividad. Esto sugiere la necesidad de

diseñar consignas más explícitas, estructuradas y contextualizadas, ya que una instrucción ambigua puede trasladarse a la interacción con la IA.

La formulación de la petición a la IA fue la dimensión con mayor necesidad de acompañamiento. Varios estudiantes elaboraron instrucciones demasiado generales, lo que generó respuestas amplias y poco focalizadas. Esto evidencia que el *prompting* constituye una competencia clave para obtener respuestas claras, específicas y útiles.

En relación con la comprensión de la respuesta generada, los participantes mostraron un desempeño favorable; sin embargo, algunas respuestas resultaron extensas y demandaron procesos adicionales de lectura, selección y síntesis. La mayoría logró aprovechar la información para completar la tarea, aunque algunos requirieron apoyo para transformarla en un producto concreto. Esto indica que la claridad de la petición inicial influye directamente en la utilidad de la respuesta y que la IA generativa requiere mediación pedagógica para promover la revisión, adaptación y apropiación crítica del contenido.

De forma transversal, el moderador desempeñó un papel de apoyo y mediación, proporcionando orientación gradual cuando los estudiantes encontraban dificultades. Este acompañamiento resultó especialmente útil para concretar objetivos, formular consultas más precisas y adaptar las respuestas a las necesidades de la tarea. En consecuencia, se observaron interacciones más efectivas y respuestas generalmente más claras, pertinentes y comprensibles.

Dentro del contexto de la accesibilidad para personas con discapacidad visual que utilizan herramientas de apoyo, como lectores de pantalla, aún existen varias limitaciones que no permiten el completo aprovechamiento del uso de la IA. Es importante resaltar que aún falta adaptar las interfaces para que sean realmente accesibles conforme a las recomendaciones WAI —*Web Accessibility Initiative* como iniciativa para la Accesibilidad en la Web del W3C.

6. Limitaciones del estudio

El estudio presenta limitaciones propias de su carácter exploratorio y contextualizado. En primer lugar, la muestra no fue seleccionada mediante procedimientos probabilísticos, por lo que los resultados no deben interpretarse como generalizables a todos los contextos educativos. Su valor principal reside en identificar patrones de interacción, barreras recurrentes y necesidades de apoyo durante una experiencia real de uso. En segundo lugar, la actividad evaluada se centra en una tarea educativa breve basada en comprensión, reformulación y explicación de información mediante lenguaje natural. Por ello, las barreras observadas podrían variar en tareas de mayor complejidad, en actividades disciplinares distintas o en contextos con perfiles funcionales más específicos. Finalmente, parte de los datos proceden de registros de observación y respuestas autorreportadas, lo que puede introducir variaciones en la interpretación de las categorías. No obstante, la depuración de registros y la normalización de respuestas permiten organizar la información de manera consistente para una primera lectura empírica de la accesibilidad y usabilidad de la herramienta.

7. Conclusiones

Los resultados permiten concluir que la IA generativa puede actuar como un apoyo relevante para favorecer la comprensión, organización y reformulación de información en contextos educativos inclusivos. Sin embargo, el acceso técnico a la herramienta no garantiza una interacción autónoma ni una experiencia plenamente accesible. Las principales barreras se concentraron en la comprensión inicial de la consigna y en la formulación de la petición a la IA. Esto muestra que el uso educativo de sistemas conversacionales requiere acompañamiento pedagógico, orientación en la construcción de instrucciones y estrategias de revisión crítica de las respuestas generadas. Desde la perspectiva del Diseño Universal para el Aprendizaje, los hallazgos refuerzan la necesidad de diseñar actividades con múltiples formas de representación, interacción y expresión. Las explicaciones breves, los ejemplos cotidianos, las analogías, los apoyos visuales y el desglose por pasos favorecen la participación de estudiantes con distintos niveles de experiencia tecnológica. Como línea futura, se propone ampliar la muestra, comparar diferentes herramientas de IA generativa, incorporar perfiles funcionales más diversos y analizar con mayor profundidad la relación entre accesibilidad técnica, carga cognitiva, autonomía y calidad de la interacción humano–IA en tareas educativas.

Agradecimientos Esta investigación es parte del proyecto "Inteligencia Artificial Ética, Responsable y de Propósito General: Aplicaciones en Escenarios de Riesgo" (IAFER), Exp. TSI-100927-2023-1, financiado en el marco del Plan de Recuperación, Transformación y Resiliencia de la Unión Europea- Next Generation EU, a través del Ministerio para la Transformación Digital y de la Función Pública y está parcialmente financiada por el Ministerio de Ciencia, Innovación y Universidades de España en el marco del proyecto PID2023-150070NB-I00. Asimismo, el Centro Universitario de Tlaquepaque de Universidad de Guadalajara, México financia a los autores Durán y Zermeño.

Declaración sobre el uso de IA Durante la preparación del manuscrito se utilizaron modelos de lenguaje para mejorar la redacción. Los autores revisaron el contenido y asumen plena responsabilidad sobre el artículo.

Referencias

1. UNESCO, "Guía para el uso de la IA generativa en educación e investigación," Guía oficial UNESCO, 2024. Disponible en: https://unesdoc.unesco.org/ark:/48223/pf0000386693
2. C. Schwab and M. Abowd, "From Inaccessible to Inclusive: How the New PDF Accessibility Remediation Solution Helps Institutions Compliantly Address Accessibility Requirements,"
3. W3C, "Web Content Accessibility Guidelines (WCAG) 2.2," W3C Recommendation, 2024. [Online]. Disponible en: https://www.w3.org/TR/WCAG22/
4. A. Ahmed, M. Fresco, F. Forsberg, and H. Grotli, "From Code to Compliance: Assessing ChatGPT's Utility in Designing an Accessible Webpage—A Case Study," *arXiv preprint*, 2025. [Online]. Disponible en: https://arxiv.org/abs/2501.03572

5. D. Shingleton, "The missing link in AI: Human Insight loops for customer-centric conversations," Keynote en el Conversational AI & Customer Experience Summit, 2026. [Online]. Disponible en: https://www.usertesting.com/blog/human-feedback-conversational-ai
6. International Organization for Standardization, *Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs)—Part 11: Guidance on Usability, ISO 9241-11:1998*, Geneva, Switzerland, 1998.
7. S. Borsci, A. Malizia, M. Schmettow, F. Van Der Velde, G. Tariverdiyeva, D. Balaji, and A. Chamberlain, "The Chatbot Usability Scale: the Design and Pilot of a Usability Scale for Interaction with AI Chatbots," 2022. [Online]. Disponible en: https://research.utwente.nl/en/publications/the-chatbot-usability-scale-the-design-and-pilot-of-a-usability-s/
8. UX studio, "Assessing the 10 Usability Principles for AI Interfaces," Jul. 2024. [Online]. Disponible en: https://www.uxstudioteam.com/ux-blog/10-usability-principles-for-ai
9. R. Alabduljabbar, "User-centric AI: evaluating the usability of generative AI applications through user reviews on app stores," *PeerJ Comput Sci.*, p. 10:e2421, 2024. [Online]. Disponible en: https://pmc.ncbi.nlm.nih.gov/articles/PMC11623163/
10. Nielsen Norman Group, "AI as a UX Assistant," Oct. 2023. [Online]. Disponible en: https://www.nngroup.com/articles/ai-roles-ux/
11. Microsoft Support. Herramientas de accesibilidad para Microsoft Copilot, 2024. [Online]. Disponible en: https://support.microsoft.com/es-es/topic/herramientas-de-accesibilidad-para-microsoft-copilot
12. Microsoft. Driving inclusion and accessibility with Microsoft 365 Copilot, 2025. [Online]. Disponible en: https://www.microsoft.com/en-us/microsoft-cloud/blog/2025/01/13/driving-inclusion-and-accessibility-with-microsoft-365-copilot/
13. OpenAI Community. The ChatGPT web interface is not accessible to blind users, 2024. [Online]. Disponible en: https://community.openai.com/t/the-chatgpt-web-interface-is-not-accessible-to-blind-users/
14. L. Lundqvist, "Optimizing accessibility in ux design through the integration of artificial intelligence," Master's thesis, Umeå University, 2024. [Online]. Disponible en: https://urn.kb.se/resolve?urn=urn:nbn:se:umu:diva-226365
15. O. Z. Lavoie-Gagné, F. R. Chandini, and L. Khanna, "Assessing the Usability of ChatGPT Responses Compared to Other Online Information in Hand Surgery," *Hand (New York)*, 2025. [Online]. Disponible en: https://doi.org/10.1177/15589447251329584
16. Contentsquare, "8 métodos de pruebas de usabilidad que funcionan (tipos y ejemplos)," Oct. 2024. [Online]. Disponible en: https://contentsquare.com/es-es/guias/test-de-usabilidad/metodos/
17. QuestionPro, "Métodos para la evaluación de usabilidad: ¿cuál elegir?" Oct. 2023. [Online]. Disponible en: https://www.questionpro.com/blog/es/metodos-para-la-evaluacion-de-usabilidad/
18. OpenAI. Be My Eyes and OpenAI, 2023. [Online]. Disponible en: https://openai.com/index/be-my-eyes/
19. R. Montes, "Accesibilidad-IA: Materiales metodológicos para evaluación de accesibilidad y usabilidad de IA generativa," 2026. [Online]. Disponible en: https://github.com/rosanamontes/Accesibilidad-IA.