

Revisión Sistemática de Literatura: [redactar un título representativo con su idea de investigación]

[Nombres y apellidos]

Universidad Nacional de Loja, Facultad de la Energía, las Industrias y los Recursos Naturales No Renovables, Carrera de Computación, Av. Pío Jaramillo Alvarado y Av. Reinaldo Espinosa, código postal 110111, Loja, Ecuador.

[]@unl.edu.ec

Resumen

Antecedentes: [redactar los antecedentes o motivación de la presente investigación exploratoria]

Objetivo: [redactar el objetivo de la RSL]

Métodos: [redactar el procedimiento de manera general y que metodología se usó]

Resultados: [redactar de manera general que es lo que se encontró a nivel cuantitativo]

1. Introducción

[Extender los antecedentes, citar artículos que fundamenten esta parte teórica]

[Trabajos relacionados, que otros trabajos similares o revisiones sistemáticas de literatura existen, citar esos trabajos]

[Redactar el objetivo de la RSL y los aportes y beneficios que se obtendrán]

2. Metodología

Se utilizó la metodología propuesta por Kitchenham [1][2], el análisis PICOC (Tabla 2) del método PRISMA [3] y la guía práctica para desarrollar un Revisión Sistemática de Literatura (RSL) propuesta en [4], que en conjunto y de manera sistemática permitirán especificar, identificar, evaluar, seleccionar y sintetizar los datos de las fuentes primarias recopiladas.

2.1. Preguntas de investigación

En la Tabla 1 se presenta el esquema PICOC, el cual permite contextualizar y delimitar adecuadamente la formulación del problema de investigación.

Tabla 1. PICOC para contextualizar la RSL

Elemento	Descripción del componentes	Detalle
P – Población	¿A quién o a qué se dirige el estudio?	Estudios, herramientas o proyectos en el ámbito de la ingeniería de software enfocados en la accesibilidad digital de documentos electrónicos.
I – Intervención	¿Qué técnica, método o herramienta se analiza?	Integración de los estándares de accesibilidad digital WCAG 2.1 y PDF/UA en el diseño y desarrollo de software o prototipos de conversión de documentos PDF a formatos accesibles (HTML o EPUB).
C – Comparación	¿Con qué alternativa se compara?	Enfoques o metodologías que no aplican o solo integran parcialmente los estándares de accesibilidad digital.
O – Resultados	¿Qué se espera obtener o medir?	Mejoras en la accesibilidad, conformidad técnica y usabilidad de los documentos digitales resultantes, verificadas con métricas de cumplimiento (WCAG A/AA, PDF/UA).
C - Contexto	¿Dónde o en qué ámbito se desarrolla?	En el contexto del desarrollo de software accesible orientado a la transformación digital inclusiva y sostenible.

En la Tabla 2 se detallan las preguntas de investigación que orientan el proceso de esta revisión sistemática de literatura.

Tabla 2. Preguntas de investigación para la RSL

N.º	Pregunta de investigación	Relación con objetivos del estudio	Nivel de prioridad (Alta / Media / Baja)
1	¿Qué enfoques, metodologías y herramientas de ingeniería de software se han empleado para integrar los estándares de accesibilidad digital WCAG 2.1 y PDF/UA en el diseño y desarrollo de sistemas o prototipos destinados a convertir documentos PDF en formatos accesibles (HTML o EPUB)?	Identificar y analizar la evidencia científica sobre la aplicación de estándares de accesibilidad en el desarrollo de software de conversión PDF–HTML/EPUB, y evaluar sus impactos en accesibilidad y conformidad técnica.	Alta
2	¿Qué métricas, validadores y criterios de evaluación se utilizan en la literatura científica para medir el cumplimiento de los estándares WCAG 2.1 y PDF/UA en sistemas o prototipos de conversión de documentos PDF a formatos accesibles?	Analizar los métodos y herramientas utilizados para evaluar la conformidad técnica y la accesibilidad en software de conversión de documentos.	Media

3	¿Cuáles son las principales limitaciones, desafíos y buenas prácticas reportadas en la integración de estándares de accesibilidad digital dentro del proceso de desarrollo de software inclusivo?	Sintetizar los retos técnicos y las estrategias de mejora documentadas en la implementación de accesibilidad digital.	Media
---	---	---	-------

2.2. Estrategia de Búsqueda

a. Definir las palabras claves:

Se identificaron las palabras claves en el idioma inglés, algunas de ellas fueron extraídas del Tesauro de la IEEE¹, otras del Tesauro de la UNESCO² y otros del ACM Computing Classification System, que servirán para definir las cadenas de búsqueda.

Tabla 3. Términos clave para la RSL

Término principal	Sinónimos / variantes	Traducción al inglés	Fuente de validación (IEEE / UNESCO / ACM)
Enfoques de ingeniería de software	Métodos de ingeniería de software, procesos de desarrollo de software	Software engineering methods, Software development process	IEEE Thesaurus
Accesibilidad digital	Inclusión digital, diseño universal	Digital inclusion, Universal design	UNESCO Thesaurus
Conversión de PDF	Transformación de PDF, extracción de contenido PDF	PDF conversion, Document transformation	Término de uso común en la literatura especializada
Formatos accesibles	HTML accesible, EPUB accesible	Accessible HTML, EPUB accessibility	Término de uso común en la literatura especializada
Estándares de accesibilidad	WCAG, PDF/UA	WCAG 2.1, PDF/UA	Término de uso común en la literatura especializada

¹ Tesauro de la IEEE, <https://www.ieee.org/publications/services/thesaurus-access-page.html> [12/03/2021]

² Tesauro de la UNESCO, <https://skos.um.es/unescothes/COL160/html, http://vocabularies.unesco.org/browser/thesaurus/es/>, [15/01/2021]

Metodologías de ingeniería de software	Desarrollo de software ágil	Software engineering methodologies, Agile software development	ACM Thesaurus
Herramientas de ingeniería de software	Ingeniería de software asistida por computadora, entornos de programación	Computer aided software engineering, Programming environments	IEEE Thesaurus
Validadores de accesibilidad	Evaluadores automáticos	Accessibility validators	Término de uso común en la literatura especializada
Métricas de conformidad	Índice de cumplimiento	Conformance metrics	Término de uso común en la literatura especializada
Pruebas de accesibilidad	Evaluación experta, test manual de accesibilidad	Accessibility testing	ACM / IEEE Thesaurus
Barreras técnicas de accesibilidad	Limitaciones técnicas, restricciones de implementación	Accessibility technical barriers	Término de uso común en la literatura especializada
Buenas prácticas de accesibilidad	Recomendaciones, pautas de implementación	Accessibility best practices	Término de uso común en la literatura especializada
Retos en la conversión de PDFs	Dificultades en etiquetado, problemas de estructura	PDF conversion challenges	Término de uso común en la literatura especializada

b. Definir las bases de datos

De acuerdo con los lineamientos de Kitchenham y de la guía práctica para desarrollar una RSL, en la Tabla 5 se presentan las bases de datos científicas multidisciplinarias y especializadas de gran impacto mundial que se utilizaron en la investigación.

Tabla 4. Bases de datos científicas

ID	Base de datos científica	Dirección Web
BD-1	Scopus	https://www.scopus.com/home.uri
BD-2	ACM Digital Library	https://dl.acm.org/

c. Definir y comprobar los scripts de búsqueda

Una vez definidas las bases de datos e identificadas las palabras clave, se realizaron consultas básicas de prueba utilizando los operadores lógicos AND y OR para verificar la pertinencia de los términos seleccionados. A partir de estas pruebas iniciales se generaron los scripts o cadenas de búsqueda finales, los cuales se muestran en la Tabla 5. Cada script de búsqueda presenta particularidades propias según la base de datos utilizada, ya que la sintaxis de consulta es diferente.

Tabla 5. Scripts de búsqueda por cada base de datos

Base de Datos	Cadena	Fecha de consulta	N.º de resultados obtenidos	N.º de pregunta
ACM	("accessible PDF" OR "PDF/UA" OR "tagged PDF" OR "PDF remediation") AND ("scientific documents" OR "academic documents" OR ETD OR "PDF to HTML" OR "document conversion" OR "semantic structure")	29/10/2025	27	1
SCOPUS	("accessible PDF" OR "PDF/UA" OR "tagged PDF" OR "PDF remediation") AND ("scientific documents" OR "academic documents" OR ETD OR "PDF to HTML" OR "document conversion" OR "semantic structure")	29/10/2025	46	1
ACM	("accessible PDF" OR "PDF/UA") AND ("WCAG 2.1" OR "accessibility standard") AND ("evaluation" OR "assessment" OR "testing" OR "compliance")	29/10/2025	18	2

SCOPUS	("accessible PDF" OR "PDF/UA") AND ("WCAG 2.1" OR "accessibility standard") AND ("evaluation" OR "assessment" OR "testing" OR "compliance")	29/10/2025	23	2
ACM	("accessible PDF" OR "PDF/UA") AND ("WCAG 2.1" OR "accessibility standard") AND ("challenges" OR "limitations" OR "barriers" OR "best practices" OR "guidelines")	29/10/2025	20	3
SCOPUS	("accessible PDF" OR "PDF/UA") AND ("WCAG 2.1" OR "accessibility standard") AND ("challenges" OR "limitations" OR "barriers" OR "best practices" OR "guidelines")	29/10/2025	26	3

2.3. Ejecutar scripts de búsqueda

Al ejecutar los scripts descritos en la Tabla 5, se presenta el resumen y total de fuentes primarias encontradas (Tabla 7).

Tabla 6. Ejecución de cada script de búsqueda por cada base de datos

Base de datos	Script aplicado	Fecha de ejecución	N.º de resultados	Tipo de documento (Artículo / Conf. / Revisión)	Intervalo temporal (años)	Relevancia (Alta / Media / Baja)
Primera Pregunta						
ACM	("accessible PDF" OR "PDF/UA" OR "tagged PDF" OR "PDF remediation") AND ("scientific documents" OR "academic documents" OR ETD OR "PDF to HTML" OR "document conversion" OR "semantic structure")	10/11/2025	20	Proceedings	2020–2025	Alta

ACM	("accessible PDF" OR "PDF/UA" OR "tagged PDF" OR "PDF remediation") AND ("scientific documents" OR "academic documents" OR ETD OR "PDF to HTML" OR "document conversion" OR "semantic structure")	10/11/2025	6	Research Article	2020– 2025	Alta
Scopus	("accessible PDF" OR "PDF/UA" OR "tagged PDF" OR "PDF remediation") AND ("scientific documents" OR "academic documents" OR ETD OR "PDF to HTML" OR "document conversion" OR "semantic structure")	10/11/2025	20	Artículo	2020– 2025	Alta
Scopus	("accessible PDF" OR "PDF/UA" OR "tagged PDF" OR "PDF remediation") AND ("scientific documents" OR "academic documents" OR ETD OR "PDF to HTML" OR "document conversion" OR "semantic structure")	10/11/2025	11	Conferencia	2020– 2025	Alta

Segunda Pregunta

ACM	("accessible PDF" OR "PDF/UA") AND ("WCAG 2.1" OR "accessibility standard") AND ("evaluation" OR "assessment" OR "testing" OR "compliance")	10/11/2025	16	Proceedings	2020– 2025	Alta
ACM	("accessible PDF" OR "PDF/UA") AND ("WCAG 2.1" OR "accessibility standard") AND ("evaluation" OR	10/11/2025	3	Research Article	2020– 2025	Alta

	"assessment" OR "testing" OR "compliance")					
Scopus	("accessible PDF" OR "PDF/UA") AND ("WCAG 2.1" OR "accessibility standard") AND ("evaluation" OR "assessment" OR "testing" OR "compliance")	10/11/2025	6	Conferencia	2020– 2025	Alta
Scopus	("accessible PDF" OR "PDF/UA") AND ("WCAG 2.1" OR "accessibility standard") AND ("evaluation" OR "assessment" OR "testing" OR "compliance")	10/11/2025	5	Artículo	2020– 2025	Alta
Tercera Pregunta						
ACM	("accessible PDF" OR "PDF/UA") AND ("WCAG 2.1" OR "accessibility standard") AND ("challenges" OR "limitations" OR "barriers" OR "best practices" OR "guidelines")	10/11/2025	3	Research Article	2020– 2025	Alta
ACM	("accessible PDF" OR "PDF/UA") AND ("WCAG 2.1" OR "accessibility standard") AND ("challenges" OR "limitations" OR "barriers" OR "best practices" OR "guidelines")	10/11/2025	18	Proceedings	2020– 2025	Alta
Scopus	("accessible PDF" OR "PDF/UA") AND ("WCAG 2.1" OR "accessibility standard") AND ("challenges" OR "limitations" OR "barriers" OR "best practices" OR "guidelines")	10/11/2025	7	Coferencias	2020– 2025	Alta
Scopus	("accessible PDF" OR "PDF/UA") AND ("WCAG 2.1" OR "accessibility standard") AND	11/11/2025	6	Artículo	2009– 2025	Alta

	("challenges" OR "limitations" OR "barriers" OR "best practices" OR "guidelines")				
--	---	--	--	--	--

Tabla 7. Scripts de búsqueda por cada base de datos

Base de datos	N.º de resultados obtenidos
Socpus	55
ACM Digital Library	66
Total	111

2.4. Descargar y almacenar los resultados de los scripts de búsqueda

Tabla 8. Descarga de los resultados obtenidos de cada base de datos

N.º	Base de datos	Formato de exportación (CSV / RIS / BibTeX)	Nombre del archivo	Fecha de descarga
2	ACM Digital Library	BibTeX	https://drive.google.com/drive/folders/16CNolQIUMSyerv03LC_Zt8_zw527qwWv?usp=sharing	12 / 11 / 2025
3	Scopus	BibTeX	https://drive.google.com/drive/folders/1wSOhA6P_M8mA_ZPJSYU07VkTYwWr6e4z-?usp=sharing	12 / 11 / 2025

De acuerdo con los resultados obtenidos en cada una de las bases de datos (Tabla 8), se exportó en el formato BibTeX (.bib) los metadatos “Citation information” y “Abstract & keywords”, que fueron importados en el gestor bibliográfico Mendeley (Fig. 1).

Figura 1. Resultado de la ejecución de la búsqueda

2.5 Criterios de selección

En la Tabla 10 y 11 se establecen los criterios de inclusión y exclusión.

Tabla 9. Parámetros

Campo	Valor
Período	2020–2025
Idiomas	Español / Inglés
Tipo de documento	Journal / Conference (revisado por pares)
Dominio	Ciencias de la Computación
Acceso	Título y resumen disponibles; DOI/URL verificable
Coherencia con PI	El estudio responde a al menos una PI

En la Tabla 8 y 9 se establecen los criterios de inclusión y exclusión.

Tabla 10. Criterios de inclusión

Código	Criterio de inclusión	Ejemplo (✓)
I1	Alineación con las Preguntas de Investigación (PI): El estudio responde al menos a una PI abordando: (a) un método/herramienta de conversión de PDF a formato accesible (PI1), (b) métricas o validadores aplicados al proceso de conversión o accesibilidad (PI2), o (c) desafíos, limitaciones o buenas prácticas en la conversión o accesibilidad documental (PI3).	“Un prototipo basado en IA para la conversión PDF–HTML accesible.” (PI1) “Evaluación automática de accesibilidad en documentos PDF utilizando PAC3.” (PI2) “Buenas prácticas para mejorar la accesibilidad en la generación de EPUB.” (PI3)

I2	Tipo de Estudio: El artículo es un estudio primario (experimento, prueba, estudio de caso, prototipo evaluado) o un estudio secundario (RSL, Mapeo Sistemático, revisión integradora) con método explícito.	“Estudio de caso sobre accesibilidad en la conversión de documentos educativos.” “Revisión Sistémica sobre herramientas de validación WCAG.”
I3	Alineación contextual: El estudio se centra específicamente en: (a) la conversión de PDF a formatos accesibles (HTML, EPUB, DAISY, etc.), o (b) la integración/cumplimiento de estándares de accesibilidad (WCAG, PDF/UA, ARIA) durante el proceso de conversión o en el producto final.	“Metodología para incorporar PDF/UA en un conversor PDF-HTML.” “Comparación del cumplimiento WCAG en distintos motores de conversión.”
I4	Reporte de evidencia evaluable: El estudio presenta datos o hallazgos procesables, ya sean cuantitativos (métricas, porcentajes, errores, desempeño, validaciones) o cualitativos (desafíos, recomendaciones, buenas prácticas fundamentadas).	“El conversor reduce 35% de errores PDF/UA.” “Resultados de pruebas de usabilidad con usuarios con discapacidad visual.” “Lista de buenas prácticas para mejorar la legibilidad tras la conversión.”

Tabla 11. Criterios de exclusión

Código	Criterio de exclusión	Ejemplo (X)
E1	Fuera del dominio o contexto: Se excluyen estudios que no abordan conversión de PDF ni accesibilidad documental, tales como: (a) accesibilidad web o móvil general sin relación a PDF; (b) herramientas o procesos de PDF sin enfoque en accesibilidad (OCR, firma digital, cifrado, compresión, edición general); (c) artículos puramente legales, normativos o filosóficos sin aplicación técnica, prototipo o proceso de ingeniería.	“Cómo crear un sitio web accesible con React.” “El mejor conversor de PDF a Word.” “La importancia legal de WCAG 2.1” (sin componente técnico).
E2	Fuentes sin revisión por pares o sin evidencia empírica: Se excluyen blogs, artículos de opinión, demos promocionales, whitepapers sin metodología, literatura gris no evaluada por pares, y publicaciones sin métodos verificables.	Post de blog: “Las 5 mejores herramientas de accesibilidad.”
E3	Datos insuficientes: Se excluyen documentos con información incompleta que impida evaluar su pertinencia, tales como: registros duplicados, artículos sin resumen, sin año, sin DOI/ISSN, o documentos cuyo acceso no permite verificar criterios de inclusión.	Registro sin resumen que no permite evaluar si cumple I1–I4.
E4	Tipo de publicación no relevante: Se excluyen documentos no considerados estudios formales: editoriales, prefacios, cartas al editor, abstracts sin artículo completo, posters sin paper asociado o presentaciones sin documento formal.	Editorial de una revista sobre software inclusivo.

2.6 Seleccionar los artículos 1: análisis de título y resumen

Tabla 12. Fase de Selección: Análisis por Título y Resumen

ID	Título (recortado)	Resumen (clave)	Decisión	Criterio aplicado	Observación breve
SCO001	Gaming, wellness, and technology...	El estudio se centra en la intersección de gaming y tecnología para la fuerza laboral global, la motivación y el bienestar, cubriendo temas como IA y desarrollo profesional.	X	E1 (Fuera del Dominio)	El contexto está completamente fuera del ámbito de accesibilidad de documentos, conversión de formatos (PDF, HTML) o estándares documentales.
SCO002	What Every Engineer Should Know About Digital Accessibility	Aborda la accesibilidad como una calidad central de la ingeniería de software. Prepara a los ingenieros para integrar principios rectores y prácticas en el ciclo de vida del desarrollo tecnológico en general.	X	E1 / I3 (Contexto no Alineado)	Se enfoca en la accesibilidad digital general de la ingeniería de software y no en la conversión específica de PDF a formatos accesibles o la integración de estándares de documentos.
SCO003	Enhancing accessibility... in PDF tables	Propone un enfoque heurístico (PI1) para revelar la semántica oculta de las tablas PDF a usuarios con discapacidad visual. Es evaluado cuantitativa y cualitativamente contra WCAG 2.1 y Sección 508 (PI2, I4).	✓	Cumple I3	Aborda directamente un problema de accesibilidad de documentos PDF y evalúa la solución contra estándares relevantes (WCAG 2.1).
SCO004	Mathematical Content Browsing...	Presenta un enfoque de navegación (PI1) que utiliza mundos virtuales y sustitución sensorial audiovisual para hacer accesible el contenido matemático en PDFs sin etiquetar. Fue evaluado en un ensayo de usuario (I4).	✓	Cumple I3	Se centra en un enfoque de accesibilidad y conversión/navegación para contenido específico (matemáticas) dentro de documentos PDF.

SCO005	On PDF/A Conformance and Font Usage...	Reporta un estudio sobre la conformidad PDF/A-1 (PI3) y el uso de fuentes en documentos PDF del sector público sueco. Aborda el desafío del archivo a largo plazo.	<input checked="" type="checkbox"/>	Cumple I3	El foco está en un estándar de accesibilidad/mantenimiento a largo plazo (PDF/A) aplicado a documentos PDF (I3).
SCO006	Authoring tools are critical to the accessibility...	Propone 15 recomendaciones (PI3) para herramientas de autoría (procesadores de texto). Se centra en la responsabilidad de la herramienta para producir formatos accesibles (ODF, PDF/UA o HTML) e incorpora ATAG.	<input checked="" type="checkbox"/>	Cumple I3	Aborda buenas prácticas (PI3) y la integración de estándares (PDF/UA, ATAG) en las herramientas que generan documentos accesibles.
SCO007	E-learning and disability in higher education	Evalúa la práctica y provisión para hacer el e-learning accesible a estudiantes con discapacidades. Trata la legislación, guías y estándares aplicables a la educación a distancia.	<input checked="" type="checkbox"/>	E1 / I3 (Fuera del Dominio)	El contexto se enfoca en la accesibilidad del e-learning y la educación superior y no cumple con el requisito específico de conversión de PDF o integración de estándares de documentos (I3).
ACM001	MathML and other XML Technologies for Accessible PDF from LATEX	Describe el enfoque para usar MathML dentro de PDF Etiquetado (Tagged PDF) para documentos matemáticos. Introduce formalizaciones para interpretar y validar árboles de estructura PDF como XML.	<input checked="" type="checkbox"/>	Cumple I3	Aborda directamente la formalización e integración de estándares (PDF/UA) para la accesibilidad de documentos PDF de contenido STEM.
ACM002	Automatically producing accessible and reusable PDFs with LATEX	Describe los objetivos y el estado actual del proyecto "LATEX Tagged PDF" para crear PDFs accesibles y reutilizables.	<input checked="" type="checkbox"/>	Cumple I3	Se centra en el proceso de ingeniería (Tagging) para la producción automatizada de PDFs accesibles que cumplen con PDF/UA.

ACM003	Towards More Accessible Scientific PDFs...	Presenta un proceso y evalúa el prototipo PAVE 2.0 (PI1) para mejorar la precisión del etiquetado en PDFs científicos. Un estudio de usuarios (I2) mostró una mejora en las puntuaciones de etiquetado (42.0% a 80.1%) (PI2, I4).	<input checked="" type="checkbox"/>	Cumple I3	Aborda la remediación de PDF para mejorar el etiquetado y la conformidad con PDF/UA, evaluando con métricas cuantitativas.
ACM004	Uncovering the New Accessibility Crisis in Scholarly PDFs...	Análisis a gran escala (20K PDFs) para evaluar el cumplimiento de accesibilidad (Tagging, Alt-Text, etc.). Muestra que menos del 3.2% cumple todos los criterios y observa una caída preocupante desde 2019 (I4, PI2, PI3).	<input checked="" type="checkbox"/>	Cumple I3	Aborda el desafío (PI3) de la accesibilidad de PDFs académicos utilizando métricas empíricas (I4).
ACM005	FormA11y—Research and Development of a Tool for Remediating PDF Forms...	Se centra en el desarrollo de FormA11y (PI1), una herramienta para la remediación de formularios PDF para accesibilidad. Un estudio de usuarios (I2) mostró que los participantes remediaron formularios 2.8 veces más rápido (PI2, I4).	<input checked="" type="checkbox"/>	Cumple I3	Aborda el desarrollo de una herramienta (PI1) para la remediación de un tipo específico de documento PDF (formularios) y menciona PDF U/A.
ACM006	VoiceMath: Enhancing Accessibility in STEM Education through Formulae Transcription in Video Lectures	Presenta el proyecto VoiceMath (PI1), una aplicación web para la transcripción de vídeos de conferencias universitarias con contenido STEM a formatos accesibles.	<input checked="" type="checkbox"/>	E1 / I3 (Contexto no Alineado)	El foco es la accesibilidad de vídeos de conferencias y la transcripción de fórmulas en ese contexto, no la accesibilidad, conversión o remediación de documentos PDF.

SCO008	Towards More Accessible Scientific PDFs... Step-by-Step PDF Remediation...	Presenta el prototipo PAVE 2.0 (PI1) para la remediación de PDF científicos. Un estudio de usuarios (I2) mostró una mejora en la precisión del etiquetado del 42.0% al 80.1% (PI2, I4).	<input checked="" type="checkbox"/>	Cumple I3	Aborda directamente una herramienta de remediación y la mejora del etiquetado en documentos PDF.
SCO009	Accessible PDFs: Applying Artificial Intelligence for Automated Remediation of STEM PDFs	Propone aplicar Inteligencia Artificial para automatizar la remediación de PDFs (PI1/PI3), centrándose en el etiquetado y la accesibilidad de fórmulas matemáticas. Menciona PDF/UA (I3).	<input checked="" type="checkbox"/>	Cumple I3	Describe una metodología de ingeniería para reducir barreras en PDFs científicos, mencionando PDF/UA.
SCO010	FormA11y—Research and Development of a Tool for Remediating PDF Forms...	Desarrolla la herramienta FormA11y (PI1) para la remediación de formularios PDF (I3). Un estudio de usuarios (I2) demostró que los participantes remediaron formularios 2.8 veces más rápido (I4, PI2) y con más precisión.	<input checked="" type="checkbox"/>	Cumple I3	Se centra en el desarrollo y evaluación de una herramienta (PI1) para un tipo específico de documento PDF (formularios).
SCO011	Uncovering the New Accessibility Crisis in Scholarly PDFs...	Ánalisis a gran escala (20K PDFs) para evaluar el cumplimiento de accesibilidad (I2). Revela que menos del 3.2% satisface todos los criterios y observa una caída preocupante desde 2019 (I4, PI3).	<input checked="" type="checkbox"/>	Cumple I3	Aborda el desafío (PI3) de la accesibilidad de PDFs académicos utilizando métricas empíricas (I4) y criterios como Tagged PDF.
SCO012	Mobility and Utility in Robot Mediated Interaction: An Interactive Workshop...	Taller interactivo para la identificación de casos de uso de robots de telepresencia móviles y evaluación de nuevas affordances en la colaboración remota.	<input checked="" type="checkbox"/>	E1 (Fuera del Dominio)	El foco es la robótica de telepresencia y la interacción humano-computadora móvil, lo cual está fuera del dominio de accesibilidad de documentos PDF

SCO013	Gaming, wellness, and technology...	Se centra en el gaming, el bienestar, la IA y la tecnología en la fuerza laboral global.		E1 (Fuera del Dominio)	El contexto es la tecnología laboral y el bienestar, no la accesibilidad de documentos PDF o la conversión de formatos.
SCO014	Help-Seeking Situations Related to Visual Interactions on Mobile Platforms...	Investiga barreras de accesibilidad en plataformas móviles que acceden a bibliotecas digitales, enfocándose en interacciones visuales.		E1 / I3 (Contexto no Alineado)	El foco es la accesibilidad de plataformas móviles y la interacción visual, no la conversión o remediación de documentos.
SCO015	“Malicious” Pictorials: How Alt Text Matters to Screen Reader Users...	Identifica nueve patrones de "alt text malicioso" (PI3) en medios densos en imágenes. Propone patrones accesibles para guiar a los autores.		Cumple I3	Aborda una buena práctica (PI3) fundamental para la accesibilidad de documentos que contienen imágenes.
SCO016	Chart4Blind: An Intelligent Interface for Chart Accessibility Conversion	Propone Chart4Blind (PI1), una interfaz que convierte gráficos bitmap en formatos accesibles (SVG, CSV, Alt Text) que cumplen con estándares (I3).		Cumple I3	Aborda una herramienta de conversión (PI1) y la generación de accesibilidad para gráficos, elementos críticos en documentos.
SCO017	What Every Engineer Should Know About Digital Accessibility	Discute la accesibilidad como una cualidad central de la ingeniería de software en general, aplicando principios a diversas tecnologías.		E1 / I3 (Contexto no Alineado)	Se centra en la accesibilidad digital general y la ingeniería de software, no en la conversión de PDF o estándares de documentos.
SCO018	The Accessibility Paradox: Can Research Articles Inspecting Accessibility Be Inaccessible?	Evalúa cómo las versiones PDF de artículos académicos cumplen con estándares como PDF/UA y WCAG (PI2, I3). Ninguno fue "problem-free".		Cumple I3	Aborda la evaluación del cumplimiento de estándares (PDF/UA) en documentos PDF académicos.
SCO019	Enhancing accessibility... Presenting semantic information in PDF tables	Propone un enfoque heurístico (PI1) para revelar la semántica oculta de las tablas PDF. Evaluado contra WCAG 2.1 (PI2, I4).		Cumple I3	Aborda una solución técnica (PI1) para mejorar la accesibilidad de elementos complejos en documentos PDF.

SCO020	On PDF/A Conformance and Font Usage in PDF Documents...	Investiga la conformidad PDF/A (PI3) y el uso de fuentes en documentos PDF del sector público, motivado por el requisito de PDF/A-1.	<input checked="" type="checkbox"/>	Cumple I3	Se centra en el cumplimiento de estándares (PDF/A) en documentos PDF.
SCO021	Author Reflections on Creating Accessible Academic Papers	Clasifica elementos de contenido y entrevista a autores para entender los desafíos (PI3) en la creación de contenido accesible en documentos académicos.	<input checked="" type="checkbox"/>	Cumple I3	Aborda los desafíos y las prácticas (PI3) en la creación de documentos académicos.
SCO022	Authoring tools are critical to the accessibility of the documents...	Propone 15 recomendaciones (PI3) para herramientas de autoría, enfocándose en la producción de formatos accesibles como PDF/UA, ODF o HTML.	<input checked="" type="checkbox"/>	Cumple I3	Se centra en las buenas prácticas e integración de estándares (PDF/UA) en las herramientas que generan documentos accesibles.
SCO023	Towards Making University PDF's Conform to Universal Accessibility Standard	Reporta las reglas más comúnmente invalidadas del estándar PDF/UA (PI3) en PDFs universitarios y propone recomendaciones.	<input checked="" type="checkbox"/>	Cumple I3	Analiza la conformidad con el estándar PDF/UA en documentos PDF.
SCO024	An Investigation of E-Government Web Accessibility...	Investiga la accesibilidad de sitios web de E-Government en Namibia utilizando herramientas de evaluación web (Lighthouse, WAVE).	<input checked="" type="checkbox"/>	E1 / I3 (Fuera del Dominio)	El foco es la accesibilidad de portales web de E-Government, no la metodología de conversión o remediación de documentos PDF.
SCO025	The Portable Document Format: An Analysis of PDF Accessibility	Analiza la accesibilidad de una muestra de documentos PDF, relacionando técnicas con WCAG 2.1 (I3). Utiliza el PDF	<input checked="" type="checkbox"/>	Cumple I3	Ánálisis de barreras de accesibilidad y uso de métricas en documentos PDF.

		Accessibility Checker (PI2).	<input checked="" type="checkbox"/>		
ACM007	Towards More Accessible Scientific PDFs... Step-by-Step PDF Remediation...	Presenta PAVE 2.0 (PI1) para la remediación de PDF, incluyendo un enfoque para el texto alternativo de fórmulas matemáticas. Un estudio de usuarios (I2) mostró una mejora en la precisión del etiquetado del 42.0% al 80.1% [1, PI2, I4].	<input checked="" type="checkbox"/>	Cumple I3	Aborda directamente la remediación de PDFs y la mejora de la precisión del etiquetado (Tagged PDF).
ACM008	Automatically producing accessible and reusable PDFs with LATEX	Describe los objetivos del proyecto "LATEX Tagged PDF" (PI3) para crear PDFs accesibles y reutilizables. Menciona el cumplimiento con PDF/UA y el etiquetado (Tagging).	<input checked="" type="checkbox"/>	Cumple I3	Se centra en el proceso de producción automatizada de PDFs accesibles (conversión/creación) desde LaTeX.
ACM009	Accessible PDFs: Applying Artificial Intelligence for Automated Remediation of STEM PDFs	Propone aplicar Inteligencia Artificial (PI1) para automatizar la remediación de PDFs, especialmente fórmulas matemáticas, con el objetivo de facilitar la creación de PDFs científicos accesibles. Menciona PDF/UA.	<input checked="" type="checkbox"/>	Cumple I3	Describe una metodología de ingeniería (PI1) para la accesibilidad de PDFs complejos (STEM).
ACM010	FormA11y—Research and Development of a Tool for Remediating PDF Forms...	Desarrolla la herramienta FormA11y (PI1) para la remediación de formularios PDF, un tipo distinto de documento. Un estudio de usuarios (I2) demostró que los usuarios remediaron	<input checked="" type="checkbox"/>	Cumple I3	Se centra en una herramienta para la remediación de un tipo específico de documento PDF (formularios), mencionando PDF U/A.

		formularios 2.8 veces más rápido [7, I4, PI2].			
ACM011	Uncovering the New Accessibility Crisis in Scholarly PDFs...	Análisis a gran escala (20K PDFs) para evaluar el cumplimiento de accesibilidad (I2) basado en criterios como Tagged PDF y Alt-Text. Halla que menos del 3.2% cumple y observa una caída preocupante en accesibilidad desde 2019 (I4, PI2, PI3).	<input checked="" type="checkbox"/>	Cumple I3	Aborda el desafío (PI3) de la accesibilidad de PDFs académicos mediante métricas empíricas (I4).
ACM012	Disability Intimacy in HCI: Defining a Community-Driven Research Agenda	Taller enfocado en expandir la investigación sobre la Intimidad en la Discapacidad dentro de la Investigación de Accesibilidad HCI, incluyendo metodologías y éticas.	<input checked="" type="checkbox"/>	E1 (Fuera del Dominio)	El tema es social y ético en el contexto de la Interacción Humano-Computadora (HCI), no la accesibilidad, conversión o remediación de documentos PDF.
ACM013	Mobility and Utility in Robot Mediated Interaction: An Interactive Workshop...	Taller centrado en re-imaginar los robots de telepresencia móviles (ej. "iPad-on-a-stick") para identificar casos de uso en colaboración remota.	<input checked="" type="checkbox"/>	E1 (Fuera del Dominio)	El foco es la robótica de telepresencia y la interacción móvil, lo cual está fuera del dominio de accesibilidad de documentos PDF
SCO026	Digital accessibility in the professional workplace	Define la accesibilidad digital y su relevancia en el lugar de trabajo, enfocándose en requisitos legales y desafíos generales.	<input checked="" type="checkbox"/>	E1 / I3 (Contexto no Alineado)	Se centra en la accesibilidad digital general en el ámbito laboral, no en la conversión o remediación de documentos.
SCO027	Gaming, wellness, and technology...	Aborda la intersección de gaming y tecnología para el bienestar y la	<input checked="" type="checkbox"/>	E1 (Fuera del Dominio)	El contexto está fuera del ámbito de accesibilidad de documentos, conversión o estándares.

		motivación en la fuerza laboral global.			
SCO028	Help-Seeking Situations Related to Visual Interactions on Mobile Platforms...	Investiga las barreras de accesibilidad de usuarios BVI en plataformas móviles que acceden a bibliotecas digitales.	<input checked="" type="checkbox"/>	E1 / I3 (Contexto no Alineado)	El foco es la accesibilidad de plataformas móviles y la interacción visual, no la estructura o remediación de documentos.
SCO029	“Malicious” Pictorials: How Alt Text Matters to Screen Reader Users...	Identifica patrones de "alt text malicioso" (PI3) y propone patrones accesibles para guiar a los autores de medios densos en imágenes.	<input checked="" type="checkbox"/>	Cumple I3	Aborda una buena práctica (PI3) fundamental para la accesibilidad de imágenes en documentos.
SCO030	Chart4Blind: An Intelligent Interface for Chart Accessibility Conversion	Propone Chart4Blind (PI1), una interfaz que convierte gráficos bitmap en formatos accesibles (SVG, CSV, Alt Text) que cumplen con estándares.	<input checked="" type="checkbox"/>	Cumple I3	Aborda una herramienta de conversión (PI1) de elementos documentales (gráficos) a formatos accesibles.
SCO031	What Every Engineer Should Know About Digital Accessibility	Prepara a los ingenieros para integrar la accesibilidad como una calidad central de la ingeniería de software en general.	<input checked="" type="checkbox"/>	E1 / I3 (Contexto no Alineado)	Se centra en principios de accesibilidad digital general en la ingeniería de software.
SCO032	The Accessibility Paradox: Can Research Articles Inspecting Accessibility Be Inaccessible?	Evalúa cómo las versiones PDF de artículos académicos cumplen con estándares como PDF/UA y WCAG (PI2, I3).	<input checked="" type="checkbox"/>	Cumple I3	Aborda la evaluación del cumplimiento de estándares (PDF/UA) en documentos PDF académicos.
SCO033	Enhancing accessibility... Presenting semantic information in PDF tables	Propone un enfoque heurístico (PI1) para revelar la semántica oculta de las tablas PDF. Evaluado contra WCAG 2.1 (PI2, I4).	<input checked="" type="checkbox"/>	Cumple I3	Aborda una solución técnica (PI1) para mejorar la accesibilidad de elementos complejos en documentos PDF.

SCO034	Mathematical Content Browsing...	Presenta un enfoque de navegación (PI1) para hacer accesible el contenido matemático en PDFs sin etiquetar (I3).	<input checked="" type="checkbox"/>	Cumple I3	Aborda la accesibilidad de contenido complejo en PDFs mediante sustitución sensorial audiovisual.
SCO035	On PDF/A Conformance and Font Usage in PDF Documents...	Investiga la conformidad PDF/A (PI3) y el uso de fuentes en documentos PDF del sector público.	<input checked="" type="checkbox"/>	Cumple I3	Se centra en el cumplimiento de estándares (PDF/A) en documentos PDF.
SCO036	Key Issues Affecting the Inclusion of Alt Text in Scholarly PDF Publications	Explora los desafíos (PI3) que enfrentan los editores al incluir Alt Text en publicaciones PDF académicas.	<input checked="" type="checkbox"/>	Cumple I3	Aborda directamente el desafío de accesibilidad (Alt Text) en documentos PDF.
SCO037	Author Reflections on Creating Accessible Academic Papers	Entrevista a autores para entender los desafíos (PI3) en la creación de contenido accesible (tablas, imágenes) en documentos académicos.	<input checked="" type="checkbox"/>	Cumple I3	Aborda los desafíos y las prácticas (PI3) en la creación de documentos académicos accesibles.
SCO038	Authoring tools are critical to the accessibility of the documents...	Propone 15 recomendaciones (PI3) para herramientas de autoría, enfocándose en la producción de formatos accesibles como PDF/UA, ODF o HTML.	<input checked="" type="checkbox"/>	Cumple I3	Se centra en las buenas prácticas e integración de estándares (PDF/UA) en las herramientas que generan documentos accesibles.
ACM014	Towards Making University PDFs Conform to Universal Accessibility Standard	Reporta las reglas más comúnmente invalidadas de PDF/UA (PI3) en PDFs universitarios.	<input checked="" type="checkbox"/>	Cumple I3	Analiza la conformidad con el estándar PDF/UA en documentos PDF.
ACM015	An Investigation of E-Government Web Accessibility...	Investiga la accesibilidad de sitios web de E-Government en Namibia utilizando herramientas de evaluación web.	<input checked="" type="checkbox"/>	E1 / I3 (Fuera del Dominio)	El foco es la accesibilidad de portales web, no la remediación o conversión de documentos PDF.

ACM016	Automated Generation of Accessible PDF	Presenta AGAP (PI1) que automatiza la generación de PDFs accesibles desde LaTeX. Evalúa la herramienta mostrando un número mucho menor de violaciones (I4, PI2).	<input checked="" type="checkbox"/>	Cumple I3	Aborda la conversión automatizada (PI1) para generar PDFs accesibles a partir de LaTeX.
ACM017	The Portable Document Format: An Analysis of PDF Accessibility	Analiza la accesibilidad de una muestra de documentos PDF, utilizando el PDF Accessibility Checker (PI2) y relacionando técnicas con WCAG 2.1.	<input checked="" type="checkbox"/>	Cumple I3	Análisis de barreras de accesibilidad y uso de métricas en documentos PDF.

2.7 Seleccionar los artículos 2: análisis de introducción, metodología y conclusiones

Tabla 13. Fase de Selección: Análisis por Introducción, Metodología y Conclusiones.

ID	Intro alineada a PI (S/N)	Metodología válida (S/N)	Conclusiones trazables (S/N)	Riesgos tipificados	Decisión	Nota breve
SCO003	S	S	S	SI	<input checked="" type="checkbox"/>	Propone un prototipo heurístico para extraer semántica oculta en tablas PDF para BVI
SCO004	S	S	S	NO	<input checked="" type="checkbox"/>	Aborda la accesibilidad de contenido matemático (ecuaciones) que está gráficamente representado en PDFs inaccesibles

SCO005	S	N	N	Si	X	Se centra en los desafíos técnicos y legales de la incrustación de fuentes y la conformidad con PDF/A para el archivo a largo plazo.
SCO006	S	S	S	SI	<input checked="" type="checkbox"/>	Se enfoca en mejorar las herramientas de autoría (procesadores de texto) para facilitar la creación de documentos accesibles
ACM001	S	S	S	NO	<input checked="" type="checkbox"/>	Se centra en mejorar la accesibilidad de las fórmulas matemáticas generadas por LaTeX, siguiendo los estándares PDF 2.0/PDF/UA-2
ACM002	S	N	S	Si	<input type="triangle-down"/>	Describe el proyecto del Consorcio LATEX para integrar el etiquetado semántico en el core del sistema, eliminando la necesidad de post-procesamiento manual
ACM003	S	S	S	NO	<input checked="" type="checkbox"/>	Presenta PAVE 2.0, un proceso semi-automático de 8 pasos para la remediación

					de PDFs científicos
ACM004	S	S	S	NO	<input checked="" type="checkbox"/>
ACM005	S	S	S	NO	<input checked="" type="checkbox"/>
SCO009	S	N	N	SI	<input checked="" type="checkbox"/>
SCO015	S	S	S	NO	<input checked="" type="checkbox"/>

					"maliciosos" que causan problemas estructurales y de usabilidad, un hallazgo cualitativo clave (PI3, I4).
SCO016	S	S	S	NO	<input checked="" type="checkbox"/> Describe el desarrollo de Chart4Blind, una interfaz que convierte gráficos a formatos accesibles (SVG, impresiones táctiles), y reporta una evaluación exploratoria
SCO018	S	S	S	NO	<input checked="" type="checkbox"/> Estudio empírico que investiga la "paradoja de la accesibilidad" examinando la conformidad de artículos académicos sobre accesibilidad. Mide el cumplimiento (PI2) y encuentra un bajo porcentaje de artículos con contenido etiquetado (3.1%)

SCO020	N	S	S	SI	<input checked="" type="checkbox"/>	El foco principal es el cumplimiento de PDF/A (para archivo y preservación a largo plazo) y los problemas de licencia de fuentes. No aborda directamente la conversión o remediación para la accesibilidad de lectores de pantalla (E1b).
SCO021	S	S	S	NO	<input checked="" type="checkbox"/>	Estudio cualitativo (reflexiones de autores) sobre los desafíos del flujo de trabajo de publicación, incluyendo el riesgo de que el proceso editorial elimine o invalide el etiquetado de accesibilidad.
ACM016	S	S	N	SI	<input checked="" type="checkbox"/>	Propone una solución para la Generación Automatizada de PDF Accesibles desde el sistema de composición LaTeX, abordando la necesidad de metadatos de accesibilidad en este formato.

ACM017	S	S	S	SI	<input checked="" type="checkbox"/>	Estudio de caso que evalúa documentos PDF con PDF Accessibility Checker 3.0. Identifica los fallos más comunes en Contenido y Lenguaje Natural (PI2, I4)
--------	---	---	---	----	-------------------------------------	--

2.8 Seleccionar los artículos 3: lectura crítica de artículos y control de calidad

Tabla 14. Fase de Selección: lectura crítica de artículos y control de calidad.

ID	C1 Claridad	C2 Validez	C3 Rigor	C4 Fuente	C5 Relevancia	Total	Nivel	Decisión	Evidencia (pág./fig.)
SCO003	2	2	2	2	2	10	Alta	<input checked="" type="checkbox"/> Incluir	Metodología heurística: Sec. 3; Evaluación: Cuantitativa (Precisión, Recall, F-score) en datasets estándar
SCO004	2	2	2	2	2	10	Alta	<input checked="" type="checkbox"/> Incluir	Problema: Contenido matemático inaccesible en PDFs no etiquetados, Método: Exploración por mundo virtual y sustitución sensorial audio-visual,
SCO006	2	1	2	2	2	9	Alta	<input checked="" type="checkbox"/> Incluir	15 recomendaciones de diseño de interacción para herramientas de autoría (PI3)
ACM001	2	1	1	2	2	8	Media	<input checked="" type="checkbox"/> Incluir	Propuesta para usar MathML en PDF/UA-2 desde LaTeX, abordando

						6			fórmulas como imágenes (PI1, PI3)
ACM002	1	2	1	2	0	6	Media	<input checked="" type="checkbox"/> Excluir	Discusión teórica/técnica sobre PDF 2.0 y reutilización de contenido (WTPDF), no accesibilidad activa.
ACM003	2	2	2	2	2	10	Alta	<input checked="" type="checkbox"/> Incluir	Diseño: Proceso de remediación PAVE 2.0 (8 pasos); Métricas: Comparación de Tag Accuracy (Acrobat vs. PAVE 2.0); Estudio de usuario (Novatos vs. Experimentados)
ACM004	2	2	2	2	2	10	Alta	<input checked="" type="checkbox"/> Incluir	Muestra: Análisis cuantitativo a gran escala (20K PDFs); Conformidad: Solo el 3.2% cumple los 6 criterios; Correlación negativa fuerte con LaTeX ($r=-0.74$)
ACM005	2	2	2	2	2	10	Alta	<input checked="" type="checkbox"/> Incluir	Comparativa: Herramienta semi-automática FormA11y vs. Acrobat Pro para formularios PDF; Resultados: Significativamente más rápido y preciso en la remediación.
SCO015	2	1	2	2	2	9	Alta	<input checked="" type="checkbox"/> Incluir	Analiza deficiencias de Alt Text en documentos científicos; Problema clave: El Alt Text es el criterio menos cumplido
SCO016	2	1	2	2	2	9	Alta	<input checked="" type="checkbox"/> Incluir	Interfaz para conversión de gráficos a formatos accesibles (PI1).

SCO018	2	2	2	2	2	10	Alta	<input checked="" type="checkbox"/> Incluir	Estudio empírico que investiga la baja accesibilidad en artículos científicos sobre accesibilidad; Resultados: Solo el 3.9% contiene el título en la barra de título, y 84.4% falló en no usar texto alternativo cuando es innecesario
SCO021	2	1	2	2	2	9	Alta	<input checked="" type="checkbox"/> Incluir	Metodología: Estudio cualitativo (Entrevistas/Reflexiones); Rigor: Codificación detallada; Hallazgos: Impacto negativo del flujo de trabajo editorial; Recomendaciones: Fig. 6
ACM016	2	1	1	2	2	8	Media	<input checked="" type="checkbox"/> Incluir	Propuesta AGAP desde LaTeX para solventar la falta de accesibilidad inherente (PI1, PI3)
ACM017	2	1	2	2	2	9	Alta	<input checked="" type="checkbox"/> Incluir	Estudio de caso, uso de PAC 3.0 para validar conformidad (PI2); fallos en Contenido y Lenguaje Natural

2.8 Extraer las respuestas a las preguntas de investigación

Tabla 15. Extracción las respuestas a las preguntas de investigación.

ID	PI-1 – Resumen	PI-1 – Evidencia (pág/fig/t ab)	F.E. SI	PI-2 – Resumen	PI-2 – Evidencia (pág/fig/tab)	F.E. IS	PI-3 – Resumen	PI-3 – Evidencia (pág/fig/tab)	F.E. CA
SCO003	Se propone un enfoque heurístico para extraer características explícitas e implícitas (metadata, funcional, estructural, contextual) de tablas PDF sin paso a	Heurístico ; Librerías: PyPDF2, Camelot, Pandas; Marco general [Fig. 2 p. 21]; Análisis		Evaluación cuantitativa de la extracción de estructura y encabezados mediante Precisión, Recall y F-score. Evaluación cualitativa realizada por expertos	Precisión, Recall, F-score; Evaluación cualitativa; Resultados Captions [Table 2 p. 34-35]; Resultados Headers [Table 3 p. 36-37]; Resultados Estructurales [Table 4 p. 38];		Desafío: El formato PDF se basa en señales visuales, lo que reta la comprensión de tablas por usuarios BVI. Buena Práctica: Proporcionar un resumen	Desafío del formato visual; Resumen laico; Sobrecarga cognitiva; Limitación de solo top-level headers.	Alta

	etiquetar. Se utilizan librerías de código abierto como PyPDF2, Camelot y Pandas para la extracción y el análisis de datos	paso [Fig. 3 p. 26].		dominio para la conformidad con WCAG 2.1 y la Sección 508.	Conformance WCAG 2.1 [Table 7 p. 48-53]. Al		laico (lay summary) antes de la navegación para dar una visión holística.			
SCO004	Utiliza la sustitución sensorial audio-visual (modificación del algoritmo vOICe) para representar elementos gráficos (líneas de fracción, raíces). Implementa navegación interactiva (Modos Texto y Gráfico) basada en el modelo de "mundo virtual". Extrae elementos textuales y ubicaciones usando Poppler y PdfMinor	Sustitución sensorial; vOICe; Mundo virtual; Modos de navegación (Texto y Gráfico); Poppler y PdfMinor.	Alta	Métrica Binaria: Identificación completamente correcta (78.1% de acierto). Métrica de Exactitud: Puntuación de Correctitud (basada en la identificación y colocación geométrica correcta de símbolos y elementos gráficos).	Identificación completamente correcta; Puntuación de Correctitud; Resultados por ecuación [Tabla 4 p. 179] y [Tabla 5 p. 180].	Identificación completamente correcta; Puntuación de Correctitud; Resultados por ecuación [Tabla 4 p. 179] y [Tabla 5 p. 180].	Alta	Desafío: Las ecuaciones en PDF no etiquetados pierden la estructura bidimensional (fracciones, exponentes), volviéndose ambiguas en texto plano. Buena Práctica: La combinación de modos (texto y gráfico) en la navegación mejora la comprensión.	Ecuaciones inaccesibles; Pérdida de estructura; Mejora significativa en la Etapa 2. Errores comunes (omisión de símbolos/mal emplazamiento de corchetes)	Alta

SCO006	<p>Propone 15 recomendaciones para el diseño inclusivo de herramientas de autoría (procesadores de texto), que deben proveer marcado semántico para fórmulas (LaTeX, MathML) y otros elementos. El flujo de trabajo debe seguir las ATAG.</p>	<p>15 recomendaciones; Diseño inclusivo; LaTeX y MathML [192, Rec. 5]; Seguir ATAG</p>	<p>Los validadores deben verificar el contraste de color, la presencia de texto alternativo y la consistencia del orden de lectura estructural y visual. Se basa en la clasificación de barreras de WCAG 2.1.</p>	<p>Verificar contraste [187, Rec. 2]; Verificar Alt Text [193, Rec. 7]; Orden de lectura [196, Rec. 12]; Basado en WCAG 2.1 y WAI; Barreras Cognitivas.</p>	<p>Alta</p>	<p>Buena Práctica: Las herramientas de autoría deben facilitar la creación de documentos accesibles por defecto, ofreciendo plantillas y flujos de trabajo guiados (ATAG). Desafío: Considerar barreras que afectan a la accesibilidad cognitiva.</p>	<p>Herramientas accesibles por defecto; Accesibilidad cognitiva; Reducir la dependencia del conocimiento experto</p>	<p>Alta</p>
ACM001	<p>Enfoque para generar PDF accesibles desde LaTeX incrustando MathML (como Elementos de Estructura PDF) y otras Tecnologías XML. Utiliza la propiedad ActualText para ocultar contenido gráfico y reemplazarlo con caracteres estándar para tecnologías asistivas (ej. para paréntesis extensibles)</p>	<p>Uso de MathML como Structure Elements; Uso de tecnología XML; Uso de luamml (LuaTEX); Uso de propiedad ActualTex t para delimitad ores grandes.</p>	<p>Se desarrollan formalizaciones para interpretar y validar el Árbol de Estructura PDF como XML. Se menciona el uso de VeraPDF para la validación de conformidad.</p>	<p>Formalizaciones para PDF Structure Trees como XML; Herramienta show_pdf_tags (script Lua); Esquemas Relax NG y MathML Core; Uso de VeraPDF [340, Ref. 7].</p>	<p>Alta</p>	<p>Buena Práctica: El uso del atributo intent en MathML 4 permite desambiguar el significado de las fórmulas (ej. valor absoluto vs. cardinalidad) y facilitar lecturas localizadas por lectores de pantalla. Desafío: Los estándares anteriores trataban las fórmulas como imágenes simples.</p>	<p>Atributo intent en MathML 4; Lecturas localizadas; Desafío de fórmulas como imágenes.</p>	<p>Alta</p>

ACM003	PAVE 2.0 implementa un proceso semi-automático guiado de 8 pasos. Utiliza Inteligencia Artificial (AI) para la detección y generación automática de texto alternativo para fórmulas matemáticas, siguiendo reglas MathSpeak.	Proceso de 8 pasos. AI-based method. Math Editor [382, Fig. 4 (385)]. MathSpeak.	Score de Precisión de Etiquetado (<i>Tag Accuracy Score</i>) con trece criterios.	13 criterios. Matterhorn Protocol. Calidad del Alt Text no evaluada. Resultados de precisión.	Alta	Buena Práctica: El proceso guiado mejora la precisión del etiquetado significativa mente (hasta un 90% más alto que Adobe). Desafío: Adobe Acrobat Pro carece de un flujo de trabajo claro. Desafío: El Alt Text no es la mejor solución para fórmulas complejas; se sugiere un enfoque interactivo.	Aumento de precisión (novatos 39.2% a 75.2%, expertos 42.0% a 80.1%). Adobe no tiene un flujo claro. Necesidad de un estándar PDF/UA-2 para MathML.
ACM004	Análisis a gran escala (20K PDFs). Uso de PyPDF2 para extraer metadatos de creación (/Creator, /Producer).	Ánalisis de 20K PDFs. Uso de PyPDF2.	Uso del Adobe Acrobat Pro PDF accessibility checker para <i>bulk processing</i> . Criterios Adobe-6 (Default Language, Tagged PDF, Tab Order, Appropriate Nesting, Alt-text, Table headers). Evaluación comparativa con PAC3.	Adobe Acrobat Pro. Criterios Adobe-6. Normalized Total Compliance. Comparación con PAC3.	Alta	Crisis: Solo el 3.2% de los PDFs cumplen con los 6 criterios; el 74.9% no cumplen ninguno [236, 261, Table 2 (259)]. Desafío: El uso de LaTeX se correlaciona fuertemente con la baja conformidad ($\rho=-0.74$). Buena Práctica: Microsoft Word produce los PDFs más accesibles	3.2% cumple, 74.9% falla [236, 261, Table 2 (259)]. LaTeX $\rho=-0.74$. MS Word $\rho=0.66$. Fallo en calidad de Alt Text.

					($\rho=0.66$). Desafío: Los checkers automáticos fallan en evaluar la calidad del Alt Text y la lógica del orden de tabulación.			
ACM005	FormA11y es una herramienta semi-automática y guiada. Implementa un flujo de trabajo en capas/pasos (Definir Campos, Agrupar, Tooltips) [292, 305, 307, Fig. 4 p. 10]. Utiliza Machine Learning para generar una línea base de remediación, revisada por el humano. Es una extensión de A11y.	Herramienta FormA11 y. Flujo de trabajo por pasos [305, 307, Fig. 4 p. 10]. ML para línea base. Extensión de A11y.	Alta	Mide Time Performance y Task Performance. Uso de métricas de error (Precisión, Recall, FP, FN, Total Error Remaining). Evaluación de usabilidad mediante System Usability Scale (SUS).	Time Performance y Task Performance. Métricas de error. SUS.	Desafío: La remediación de formularios es no-intuitiva, repetitiva y abrumadora. Buena Práctica: FormA11y es 2.8 veces más rápido que Adobe Acrobat Pro. Buena Práctica: El proceso guiado reduce la carga cognitiva. Desafío Legal: El cumplimiento con ADA Title II y Sección 508 aumenta la demanda de tooling.	Proceso no-intuitivo/repetitivo. FormA11y 2.8 veces más rápido. Flujo guiado. Requisitos legales (Section 508, ADA).	Alta

SCO015	La metodología incluye Walkthroughs de Expertos con lectores de pantalla (JAWS, NVDA). Uso de herramientas de creación como Adobe InDesign.	Walkthroughs con JAWS y NVDA. Uso de InDesign.	Evaluación basada en WCAG 2.2. Se identifican nueve patrones de Alt Text mal formado/malicioso (Missing, Repetitive, Verbose, etc.).	WCAG 2.2. Nueve patrones maliciosos.	Alta	Desafío: El PDF es un formato que crea barreras significativas. Desafío: Las herramientas de creación (InDesign) requieren horas de tutoriales para manejar la accesibilidad. Buena Práctica: Recolección de Alt Text directamente de los autores en la fase de envío.	PDF es inaccesible. InDesign difícil de usar. Recolección de Alt Text.	Alta
SCO016	enciona herramientas de extracción de datos de gráficos como ChartDetective y LineEX (que utiliza Instance Segmentation).	ChartDetective; LineEX.	Se refiere al uso de modelos de segmentación de instancias para la extracción de datos.	Uso de modelos de segmentación.	Baja	Desafío: Los gráficos son típicamente raster graphics sin metadatos textuales necesarios para la accesibilidad.	Gráficos son raster graphics.	Media
SCO018	Evaluación dual utilizando el Adobe Accessibility Checker (ACC) y una evaluación manual para una inspección exhaustiva.	Tool-based evaluation (ACC) y Manual evaluation .	Se evaluaron los siete aspectos del ACC. Complementado con evaluación manual de orden de lectura lógico y contraste de color. Se basó en WCAG y PDF/UA.	Siete aspectos del ACC. WCAG y PDF/UA. Orden de lectura y contraste manual.	Media	Paradoja/Desafío: El 100% de los artículos evaluados (enfocados en accesibilidad) fallaron en cumplir los requisitos. Limitación: Solo el 12.5% de los artículos tenían etiquetas estructurales	100% de los artículos fallaron. 12.5% con etiquetas adecuadas. Deficiencias de MS Word/LaTeX.	Alta

						adecuadas. Desafío: Las herramientas de autoría (MS Word, LaTeX) tienen deficiencias en la generación de PDF accesibles.	
SCO021	Metodología de escritura que integra la accesibilidad iterativamente (ej. escribir Alt Text al añadir la imagen). Se utiliza LaTeX (Overleaf) y MS Word como herramientas de autoría.	Proceso iterativo para Alt Text. Uso de LaTeX, MS Word y LaTeX.	Alta	Se hace referencia a WCAG 2.1, PDF/UA (ISO 14289:1), y guías específicas (ACM, CHI, ASSETS).	WCAG 2.1. PDF/UA. Guías de ACM/CHI/ASSETs.	Media	Desafío: La accesibilidad es a menudo vista como el último paso (modelo <i>waterfall</i>). Desafío: Las herramientas de autoría hacen que el etiquetado sea inaccesible para un usuario ciego. Desafío: Necesidad de considerar barreras cognitivas. Buena Práctica: Requerir la publicación de documentos accesibles. Accesibilidad como último paso. Etiquetado inaccesible para ciegos. Barreras cognitivas. Modelo de responsabilidad es. Alta

ACM016	AGAP automatiza la generación de PDF accesibles desde LaTeX. Proporciona <i>feedback</i> en tiempo de compilación. Utiliza síntesis de voz (TTS) y atajos de teclado. Emplea el paquete "accessibility" para el etiquetado.	AGAP automatiza LaTeX. Feedback en tiempo de compilación. Uso de paquete "accessibility".	Evaluación usando un validador de accesibilidad estándar (Adobe Acrobat Accessibility Checker). Se cuantifica el número de violaciones.	Validador estándar.	Baja	Desafío: LaTeX por defecto no genera PDF estructurados/etiquetados, especialmente para contenido matemático. Desafío/Limitación: La herramienta de Adobe no verifica todos los criterios de accesibilidad.
ACM017	Aplica un Método de 5 Fases (Selección, Revisión con Checker, Registro, Análisis, Sugerencias) [467, Fig. 1 p. 5]. Se sugiere estructurar el documento fuente para generar marcadores (<i>bookmarks</i>) automáticamente.	Método de 5 Fases [467, Fig. 1 p. 5].	Validador principal: PDF Accessibility Checker (PAC 3.0), que valida la conformidad con ISO 32000-1 (PDF/UA-1) y WCAG 2.1. Se analizan 23 técnicas de PDF y de WCAG 2.1. Las métricas son el recuento de barreras de accesibilidad por categoría (Contenido, Lenguaje Natural, Descripciones Alternativas, etc.) [469, 470, Fig. 2 p. 6].	PAC 3.0. ISO 32000-1 (PDF/UA-1) y WCAG 2.1. 23 técnicas de PDF. Recuento de barreras [469, 470, Fig. 2 p. 6].	Alta	Desafío: La mayoría de los documentos no son compatibles con PDF/UA; el 60% contiene errores de Contenido. Desafío: El error más frecuente es la imposibilidad de identificar el Lenguaje Natural del contenido. Buena Práctica: Etiquetar las tablas correctamente con TABLE, TR, TH, y TD. No compatibles con PDF/UA. 60% Contenido, 29% Lenguaje Natural. Etiquetado de tablas.

3. Resultados

P1: ¿Qué enfoques, metodologías y herramientas de ingeniería de software se han empleado para integrar los estándares de accesibilidad digital WCAG 2.1 y PDF/UA en el diseño y desarrollo de sistemas o prototipos destinados a convertir documentos PDF en formatos accesibles (HTML o EPUB)?

Enfoques de Ingeniería de Software empleados

A partir de los estudios seleccionados, se identificaron tres enfoques principales:

1. Procesamiento heurístico y extracción semántica: Utilizado en X estudios, mediante bibliotecas como Camelot, PyPDF2, Pandas, combinado con heurísticas estructurales para segmentar tablas, regiones y bloques textuales. Este enfoque se usa especialmente cuando los PDFs carecen de etiquetas estructurales WCAG o PDF/UA[5].
2. Conversión directa asistida por reglas y plantillas: Herramientas como SciA11y transforman PDF a HTML accesible, aunque los estudios reportan errores en encabezados, tablas y orden de lectura[5].
3. Remediación guiada por pasos (User-Centered Design) Presentado en sistemas como PAVE 2.0 y Ally, que dividen la remediación en subtareas (regiones, lectura, encabezados, tablas, listas, fórmulas), disminuyendo la complejidad del árbol de estructura PDF[5].

Metodologías identificadas

- Descomposición de tareas desde HCI (Human–Computer Interaction)[6].
- Procesos iterativos centrados en el usuario (PAVE 2.0)[6].
- Workflows automáticos basados en ML para detección de elementos estructurales[6].
- Pipeline de remediación híbrido (automático + manual asistido)[7].

Herramientas de ingeniería de software utilizadas

Los estudios reportan herramientas en tres categorías principales:

1. Extracción y análisis estructural
 - Camelot, PyPDF2, PDFMiner, Pandas [5].
 - NLP con TF-IDF para seleccionar contexto relevante [5].
2. Detección automática mediante IA/ML
 - Modelos de detección de objetos: GFLv2 + ResNeXt101 (PAVE 2.0) [6].
 - Segmentación de instancias: Mask2Former (Chart4Blind) [8].
 - OCR: Tesseract para textos embebidos [8].

3. Representación accesible

- Uso de MathML (LuaMML, LuaTeX) integrado según PDF/UA-2 [9].
- Visualizaciones accesibles con D3.js y SVG[8].
- Generación automática de LaTeX desde imágenes de fórmulas mediante IA [5].

P2: ¿Qué métricas, validadores y criterios de evaluación se utilizan en la literatura científica para medir el cumplimiento de los estándares WCAG 2.1 y PDF/UA en sistemas o prototipos de conversión de documentos PDF a formatos accesibles?

A partir del análisis de los estudios incluidos en la revisión, se identificó que la evaluación del cumplimiento de los estándares WCAG 2.1 y PDF/UA en sistemas o prototipos de conversión de documentos PDF a formatos accesibles se realiza mediante tres componentes principales: validadores automatizados, métricas cuantitativas de rendimiento y criterios formales de evaluación. A continuación se describen de manera estructurada.

Validadores y herramientas automatizadas

La literatura científica emplea diversas herramientas de verificación automática con el fin de evaluar la conformidad técnica de los documentos generados:

1. **Adobe Accessibility Checker (ACC):** Utilizada ampliamente como herramienta estándar para la generación de informes de accesibilidad. Permite identificar errores relacionados con etiquetado, estructura, metadatos y elementos alternativos [10].
2. **PDF Accessibility Checker (PAC / PAC 3):** Es la herramienta más empleada para verificar el cumplimiento del estándar PDF/UA, proporcionando un análisis detallado del etiquetado, la semántica estructural y el orden lógico del documento [5], [6].
3. **VeraPDF:** Aunque está orientada principalmente a verificar el estándar PDF/A, algunos estudios la utilizan como complemento para evaluar aspectos estructurales y metadatos relevantes para la accesibilidad [9].
4. **Validadores especializados integrados en sistemas prototipo:** Por ejemplo, herramientas internas como los verificadores de accesibilidad de AGAP detectan violaciones a los requisitos durante la compilación de documentos, permitiendo un control temprano de errores[11].

Estas herramientas permiten automatizar parte del proceso de validación, aunque no sustituyen las revisiones manuales cuando se requieren interpretaciones semánticas.

Métricas cuantitativas de evaluación del rendimiento

Los estudios analizados incorporan métricas tradicionales de clasificación y detección para evaluar el desempeño de los sistemas de extracción, segmentación y etiquetado estructural, especialmente cuando se aplica aprendizaje automático:

1. Precisión: Mide el porcentaje de elementos correctamente identificados entre todos los elementos detectados por el sistema [5].
2. Exhaustividad (Recall): Evalúa la capacidad del sistema para recuperar todos los elementos relevantes presentes en el documento [5].
3. Puntuación F (F-Score): Métrica armónica que sintetiza precisión y exhaustividad, utilizada para evaluar modelos de detección y extracción de estructuras, tales como encabezados, pies de foto, celdas de tabla o regiones [5].
4. Puntuación de Conformidad: Empleada para medir la precisión del etiquetado de los elementos estructurales [6], [12]. Se expresa como:

$$\text{Puntuación} = \left(\frac{CT+WT}{CT} \right) \times 100$$

donde CT representa el número de etiquetas correctas y WT el número de etiquetas incorrectas en los criterios evaluados.

Estas métricas permiten cuantificar el grado de correspondencia entre la salida del sistema y la estructura accesible exigida por los estándares.

Criterios de evaluación basados en estándares y listas de verificación

Además de la evaluación automática, los estudios emplean criterios de conformidad derivados directamente de los estándares internacionales, así como análisis cualitativos guiados por expertos:

1. Protocolo Matterhorn: Constituye un conjunto de 136 posibles fallas relacionadas con PDF/UA. Es utilizado como referencia para identificar errores críticos de etiquetado, roles semánticos, orden de lectura y metadatos [7].
2. Criterios esenciales de accesibilidad técnica (Adobe-6): Empleados para verificar aspectos básicos del documento accesible, tales como [10]:
 - idioma predeterminado,
 - presencia de etiquetas (Tagged PDF),
 - corrección del orden de tabulación o enfoque,
 - anidamiento estructural adecuado,
 - presencia y calidad de texto alternativo,
 - definición correcta de encabezados de tabla.
3. Criterios específicos para elementos complejos
Se evalúan estructuras como tablas, gráficos y fórmulas mediante criterios relacionados con [6]:
 - identificación de títulos, filas y columnas,
 - etiquetado adecuado con <TR>, <TH>, <TD>,
 - presencia de texto alternativo o descripción semántica para fórmulas.
4. Evaluación cualitativa y heurística por expertos
Algunos estudios recurren a evaluadores expertos en accesibilidad para valorar aspectos no capturados por los validadores, como [6]:

- la relevancia semántica del texto alternativo,
- la coherencia del orden de lectura,
- la usabilidad del documento generado,
- la eficacia general del prototipo a partir de recorridos cognitivos.

Esta evaluación complementa la verificación automática, especialmente en tareas que requieren juicio humano para garantizar conformidad plena con WCAG 2.1.

P3: ¿Cuáles son las principales limitaciones, desafíos y buenas prácticas reportadas en la integración de estándares de accesibilidad digital dentro del proceso de desarrollo de software inclusivo?

El análisis de los estudios incluidos en la revisión permitió identificar un conjunto de limitaciones, desafíos y buenas prácticas relacionadas con la integración de los estándares de accesibilidad digital, en el proceso de desarrollo de software inclusivo. Estos elementos reflejan tanto barreras técnicas como organizacionales, así como estrategias recomendadas para mejorar la calidad de las soluciones orientadas a la accesibilidad.

Limitaciones y desafíos reportados

Los estudios revisados señalan diversas dificultades que afectan tanto la adopción de estándares como la implementación efectiva de funcionalidades accesibles. Entre los principales desafíos se destacan los siguientes:

1. Falta de conciencia y conocimiento especializado

La literatura señala que la causa más recurrente de inaccesibilidad es la escasa formación de autores y desarrolladores en temas de accesibilidad digital. La remediación de documentos, especialmente en PDF, exige conocimientos técnicos y experiencia que no siempre están presentes en los equipos de desarrollo[12] .

2. Herramientas poco eficientes, costosas o difíciles de usar

Las herramientas actuales de remediación suelen ser percibidas como complejas, lentas, propensas a errores y, en muchos casos, dependientes de software propietario de alto costo (por ejemplo, Adobe Acrobat Pro). Esto limita la adopción de procesos de accesibilidad en contextos con pocos recursos [6], [7].

3. Limitaciones inherentes al formato PDF

A nivel estructural, el PDF fue diseñado para la fidelidad visual y no para la accesibilidad, lo que complica la extracción semántica y la navegación mediante lectores de pantalla. Esta naturaleza del formato representa una barrera técnica fundamental para la accesibilidad [10].

4. Errores frecuentes en la automatización

Aunque los sistemas basados en inteligencia artificial han avanzado en autoetiquetado y reconocimiento de estructuras, los estudios coinciden en que estos

modelos aún presentan fallos significativos en elementos complejos como tablas, encabezados, listas o fórmulas matemáticas. Esto puede generar una “falsa impresión de accesibilidad” si no se revisa manualmente[6].

5. Complejidad semántica del contenido estructurado

La representación accesible de contenidos complejos, en particular, fórmulas matemáticas y tablas no estructuradas, continúa siendo uno de los mayores desafíos. Estos elementos requieren una interpretación semántica detallada que supera las capacidades de la mayoría de las herramientas automáticas actuales [9].

Buenas prácticas y recomendaciones identificadas

Los estudios también evidencian un conjunto de prácticas efectivas para abordar dichos desafíos y mejorar la integración de estándares de accesibilidad en el desarrollo de software.

1. Diseño centrado en la usabilidad (Human-Centred Design)

- Flujos de trabajo simplificados: Herramientas como PAVE 2.0 o FormA11y implementan procesos guiados paso a paso que reducen la carga cognitiva, mejoran la precisión y disminuyen la necesidad de conocimientos técnicos avanzados [6], [7].
- Interfaces intuitivas: Se recomienda evitar la interacción directa con la estructura interna del PDF y priorizar interacciones familiares para los usuarios, como arrastrar y soltar o dibujar líneas para segmentar tablas[6].

2. Integración estratégica de tecnologías de apoyo

- Aplicación de IA y ML: El uso de modelos de machine learning y OCR permite generar versiones preliminares parcialmente remediadas, reduciendo tareas repetitivas como la detección de formularios o la segmentación de gráficos [6], [8], [10].
- Soporte avanzado para contenido especializado: El empleo de tecnologías como MathML o LaTeX (vía herramientas como *luamml*) facilita la representación accesible y navegable de contenido matemático[9].

3. Prácticas de publicación inclusiva

- Publicación en formatos alternativos accesibles: Se recomienda ofrecer versiones adicionales del contenido en HTML, EPUB o SVG, formatos que favorecen la accesibilidad semántica [5], [10], [13].
- Responsabilidad compartida: Los autores deben crear documentos accesibles desde el origen (incluyendo texto alternativo significativo), mientras que las instituciones editoriales deben proporcionar herramientas, lineamientos y controles de calidad claros [6].
- Descripciones adecuadas para gráficos y contenido complejo: Para asegurar la comprensión, se aconseja ofrecer un texto alternativo breve acompañado de descripciones largas detalladas [6].

- Uso de señales visuales y estructuras progresivas: Estrategias como la codificación por colores o los menús progresivos ayudan a mejorar la interacción del usuario y reducen la sobrecarga cognitiva durante la remediación [7], [14].

Referencias

- [1] B. Kitchenham and C. Stuart, “Guidelines for performing Systematic Literature Reviews in Software Engineering,” Durham, 2007. doi: 10.1145/1134285.1134500.
- [2] S. Pizard, F. Acerenza, V. Casella, S. Moreno, and D. Vallespir, “Conceptos de Ingeniería de Software Basado en Evidencias,” Montevideo, 2015.
- [3] G. Urrútia and X. Bonfill, “Declaración PRISMA: una propuesta para mejorar la publicación de revisiones sistemáticas y metaanálisis,” *Med Clin (Barc)*, vol. 135, no. 11, pp. 507–511, Sep. 2010, doi: 10.1128/JCM.00582-08.
- [4] L. Chamba-Eras, “Ciencias de la Computación basada en evidencia: Guía práctica para desarrollar una Revisión Sistemática de Literatura,” Loja, 2021.
- [5] N. Fayyaz, S. Khusro, and null Imranuddin, “Enhancing accessibility for the blind and visually impaired: Presenting semantic information in PDF tables,” *Journal of King Saud University - Computer and Information Sciences*, vol. 35, no. 7, 2023, doi: 10.1016/j.jksuci.2023.101617.
- [6] F. M. Schmitt-Koopmann, E. M. Huang, H.-P. Hutter, and A. Darvishy, “Towards More Accessible Scientific PDFs for People with Visual Impairments: Step-by-Step PDF Remediation to Improve Tag Accuracy,” in *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, in CHI ’25. New York, NY, USA: Association for Computing Machinery, 2025. doi: 10.1145/3706598.3713084.
- [7] S. Paliwal *et al.*, “FormA11y—Research and Development of a Tool for Remediating PDF Forms for Accessibility,” *ACM Trans. Comput.-Hum. Interact.*, vol. 32, no. 1, Apr. 2025, doi: 10.1145/3702317.
- [8] O. Moured, M. Baumgarten-Egemole, K. E. Müller, A. Roitberg, T. Schwarz, and R. Stiefelhagen, “Chart4Blind: An Intelligent Interface for Chart Accessibility Conversion,” 2024, pp. 504–514. doi: 10.1145/3640543.3645175.
- [9] R. Kruger, F. De Wet, and T. R. Niesler, “Mathematical Content Browsing for Print-disabled Readers Based on Virtual-world Exploration and Audio-visual Sensory Substitution,” *ACM Trans Access Comput*, vol. 16, no. 2, 2023, doi: 10.1145/3584365.
- [10] A. Kumar and L. L. Wang, “Uncovering the New Accessibility Crisis in Scholarly PDFs: Publishing Model and Platform Changes Contribute to Declining Scholarly Document Accessibility in the Last Decade,” in *Proceedings of the 26th International ACM SIGACCESS Conference on Computers and Accessibility*, in

ASSETS '24. New York, NY, USA: Association for Computing Machinery, 2024.
doi: 10.1145/3663548.3675634.

- [11] S. Zulfiqar, S. Arooj, U. Hayat, S. Shahid, and A. A. Karim, “Automated Generation of Accessible PDF,” 2020, doi: 10.1145/3373625.3418045.
- [12] F. Mittelbach, U. Fischer, D. Carlisle, and J. Wright, “MathML and other XML Technologies for Accessible PDF from LATEX,” in *Proceedings of the 2025 ACM Symposium on Document Engineering*, in DocEng ’25. New York, NY, USA: Association for Computing Machinery, 2025. doi: 10.1145/3704268.3748669.
- [13] R. Menzies, G. W. Tigwell, and M. Crabb, “Author Reflections on Creating Accessible Academic Papers,” *ACM Trans. Access. Comput.*, vol. 15, no. 4, Oct. 2022, doi: 10.1145/3546195.
- [14] A. Yin *et al.*, “‘Malicious’ Pictorials: How Alt Text Matters to Screen Reader Users’ Experience of Image-Dense Media,” 2024, pp. 1262–1274. doi: 10.1145/3643834.3660747.

Anexos

[Colocar toda la evidencia en capturas o enlaces a repositorios, prompts de IA generativa, entre otros, que permita la trazabilidad y reproducibilidad de la RL]