

Chatbots en la gestión de servicios de TI: mapeo sistemático sobre service desk, gestión de incidentes y resiliencia operativa

Chatbots in IT Service Management: A Systematic Mapping on Service Desk, Incident Management and Operational Resilience

ANTHONY RAFAEL SÁNCHEZ VASQUEZ^{1*}  | JESÚS DANIEL ORDOÑEZ UGARTE²  | ALBERTO CARLOS MENDOZA DE LOS SANTOS³ 

Afiliación:

^{1,2,3}Escuela de Ingeniería de Sistemas, Universidad Nacional de Trujillo, La Libertad, Perú

Información del artículo:

Recibido: 19/05/2026
Aceptado: 09/07/2026
Publicado: 22/07/2026

Autor de correspondencia: E-mail: arsanchezva@unitru.edu.pe

Resumen

La aplicación de inteligencia artificial en la gestión de servicios de tecnologías de la información (ITSM) constituye una de las principales líneas de innovación para mejorar la eficiencia operativa y la continuidad de los servicios. En este contexto, el estudio tuvo como objetivo caracterizar el estado actual de la investigación sobre chatbots aplicados a ITSM mediante un mapeo sistemático de la literatura, siguiendo las directrices metodológicas de Petersen y la declaración PRISMA 2020. La búsqueda se realizó en Scopus, Google Scholar, ACM Digital Library y SpringerLink, recuperándose 668 registros, de los cuales 18 estudios cumplieron los criterios de inclusión. Los resultados evidencian un predominio de modelos basados en transformadores y modelos de lenguaje de gran escala en las soluciones propuestas. Asimismo, la resiliencia operativa se evalúa principalmente mediante indicadores como el tiempo medio de resolución (MTTR) y el cumplimiento de los acuerdos de nivel de servicio (SLA). Además, se identifican como principales vacíos la ausencia de evaluaciones longitudinales y de métricas estandarizadas para valorar su impacto organizacional. Los chatbots constituyen una tecnología estratégica para la automatización de ITSM; no obstante, aún requieren avances en su evaluación, validación e integración con los marcos de gestión de servicios.

Palabras clave: chatbots; gestión de incidentes; inteligencia artificial.

Abstract

The application of artificial intelligence in Information Technology Service Management (ITSM) has emerged as one of the main drivers of innovation for improving operational efficiency and service continuity. In this context, this study aimed to characterize the current state of research on chatbots applied to ITSM through a systematic mapping study, following the methodological guidelines proposed by Petersen and the PRISMA 2020 statement. The literature search was conducted in Scopus, Google Scholar, ACM Digital Library, and SpringerLink, retrieving 668 records, of which 18 studies met the inclusion criteria. The results indicate a predominance of transformer-based models and large language models (LLMs) in the proposed solutions. Furthermore, operational resilience is primarily assessed using indicators such as Mean Time to Resolution (MTTR) and Service Level Agreement (SLA) compliance. In addition, the main research gaps identified include the lack of longitudinal evaluations and standardized metrics to assess organizational impact. Chatbots represent a strategic technology for ITSM automation; however, further advances are still required in their evaluation, validation, and integration with IT service management frameworks.

Keywords: chatbots; incident management; artificial intelligence.



Esta obra está bajo licencia internacional
Creative Commons Reconocimiento 4.0



Facultad de Ingeniería
Publicación Oficial

1. Introducción

La gestión de servicios de tecnologías de la información (ITSM) comprende el conjunto de procesos y prácticas que permiten a las organizaciones diseñar, entregar, operar y mejorar sus servicios tecnológicos, apoyándose en marcos de referencia como ITIL para estandarizar funciones esenciales, entre ellas la gestión de incidentes, la atención de solicitudes y el soporte al usuario mediante el service desk (Shazzad, 2023). En este contexto, la inteligencia artificial (IA) desempeña un papel fundamental en la transformación de la gestión de servicios, al favorecer la automatización de procesos, agilizar la toma de decisiones y fortalecer la continuidad operativa.

Entre las tecnologías basadas en IA, los chatbots, sustentados en técnicas de procesamiento del lenguaje natural (*Natural Language Processing*, NLP) y aprendizaje automático (*Machine Learning*, ML), permiten atender solicitudes de manera autónoma, clasificar y escalar incidentes, además de proporcionar soporte de primer nivel de forma ininterrumpida (Caldarini et al., 2022). Su incorporación al *service desk* contribuye a reducir la carga operativa del personal de soporte, disminuir los tiempos de resolución y mejorar la experiencia del usuario final, fortaleciendo con ello la resiliencia operativa de los servicios de TI (Agnihotri y Bhattacharya, 2024).

La incorporación de estas tecnologías responde a un escenario en el que la gestión de servicios de TI enfrenta exigencias cada vez mayores asociadas con la disponibilidad continua de los servicios, el incremento del volumen de solicitudes y la necesidad de ofrecer respuestas oportunas a los usuarios. En este sentido, la automatización ha dejado de centrarse únicamente en la ejecución de tareas repetitivas para orientarse hacia procesos que requieren comprender consultas en lenguaje natural, acceder a información distribuida y asistir en la toma de decisiones durante la operación del servicio. Esta evolución ha impulsado el desarrollo de agentes conversacionales con capacidades cada vez más sofisticadas, sustentadas en técnicas de procesamiento del lenguaje natural, aprendizaje automático y, más

recientemente, modelos de lenguaje de gran escala (*Large Language Models*, LLMs), cuya incorporación ha ampliado el alcance de estas herramientas en diferentes escenarios organizacionales (Caldarini et al., 2022; Peddinti et al., 2023).

Como consecuencia de esta evolución tecnológica, la investigación sobre chatbots aplicados a ITSM ha diversificado progresivamente sus líneas de trabajo. Inicialmente, los estudios se concentraron en automatizar la atención de consultas frecuentes y facilitar la interacción entre usuarios y centros de soporte. Posteriormente, el interés se extendió hacia la clasificación automática de tickets, la priorización de incidentes, la recuperación de conocimiento organizacional y la integración con plataformas de gestión de servicios, incorporando algoritmos cada vez más especializados para optimizar estos procesos. Más recientemente, la disponibilidad de LLMs ha favorecido el desarrollo de soluciones capaces de interpretar consultas más complejas, generar respuestas contextualizadas e interactuar con diferentes fuentes de información durante la ejecución de los procesos de soporte (Bruni et al., 2023; Peddinti et al., 2023).

Este panorama ha dado lugar a una producción científica cada vez más amplia y diversa, caracterizada por la incorporación de distintas tecnologías, metodologías y contextos de aplicación. Las investigaciones difieren tanto en los procesos de ITSM analizados como en las técnicas empleadas para desarrollar los agentes conversacionales y en los criterios utilizados para evaluar su desempeño. Mientras algunos estudios centran su atención en aspectos relacionados con la precisión de los modelos conversacionales o la clasificación automática de solicitudes, otros analizan indicadores asociados con la eficiencia operativa, la experiencia del usuario o la continuidad de los servicios tecnológicos (Agnihotri y Bhattacharya, 2024).

En este contexto, resulta pertinente realizar un mapeo sistemático de la literatura científica que permita identificar la evolución del conocimiento, las tecnologías empleadas y los enfoques utilizados para evaluar la aplicación de chatbots en la gestión de servicios de TI. En este sentido, la investigación se orienta a responder la



siguiente pregunta: ¿Cómo ha abordado la literatura científica la aplicación de chatbots en la gestión de servicios de TI, particularmente en el *service desk* y la gestión de incidentes, y cuál es su relación con la resiliencia operativa de los servicios?

2. Metodología

2.1. Diseño del estudio

El estudio se desarrolló mediante un Mapeo Sistemático de la Literatura (*Systematic Mapping Study*, SMS), fundamentado en la propuesta metodológica de Petersen et al. (2008) y en las directrices actualizadas de Petersen et al. (2015), las cuales establecen los procedimientos para planificar, ejecutar y sintetizar la evidencia en estudios de mapeo sistemático dentro del ámbito de la ingeniería de software. El proceso de identificación, selección e inclusión de los estudios se estructuró siguiendo la declaración PRISMA 2020 (Page et al., 2021), mientras que los principios de calidad metodológica y revisión basada en evidencia se apoyaron en las recomendaciones de Kitchenham y Charters (2007).

2.2. Enfoque metodológico

El estudio adoptó un enfoque de mapeo sistemático orientado a identificar, clasificar y analizar la producción científica relacionada con la aplicación de chatbots en la gestión de servicios de tecnologías de la información. El procedimiento metodológico se desarrolló en tres etapas: i) planificación, que comprendió la definición de la pregunta de investigación, los criterios de elegibilidad y la estrategia de búsqueda; ii) ejecución, que incluyó la identificación, selección y revisión de los estudios recuperados; y iii) síntesis de resultados, orientada a la clasificación temática y al análisis del corpus documental seleccionado.

2.3. Criterios de inclusión y exclusión

Se definieron criterios de inclusión y exclusión con el propósito de asegurar la pertinencia temática y la calidad metodológica de los estudios incorporados al análisis. Los criterios de

inclusión fueron: estudios publicados entre 2020 y 2026 (IC1); redactados en inglés o español (IC2); artículos de revista o trabajos de conferencia revisados por pares (IC3); investigaciones que abordaran chatbots o agentes conversacionales basados en NLP o ML aplicados al ámbito de las tecnologías de la información (IC4); estudios desarrollados en contextos de ITSM, *service desk*, gestión de incidentes, gestión de solicitudes o procesos relacionados (IC5); investigaciones con contribuciones conceptuales, técnicas, metodológicas, empíricas o de síntesis, incluyendo revisiones sistemáticas y mapeos previos sobre chatbots o IA aplicada a ITSM (IC6); y disponibilidad del texto completo (IC7).

Los criterios de exclusión comprendieron estudios duplicados entre bases de datos (EC1); publicaciones fuera del periodo 2020–2026 (EC2); investigaciones sobre chatbots sin aplicación al *service desk* o a la gestión de servicios de TI (EC3); estudios sobre inteligencia artificial en TI sin componente conversacional o de atención al usuario (EC4); literatura gris, como blogs, white papers, informes comerciales o páginas web sin respaldo académico verificable (EC5); estudios exclusivamente teóricos sin relación con procesos o resultados de ITSM (EC6); e investigaciones sobre resiliencia organizacional sin aplicación tecnológica en servicios de TI (EC7).

La aplicación de estos criterios se realizó en tres fases sucesivas. Inicialmente se revisaron los títulos de todos los registros recuperados; posteriormente se evaluaron los resúmenes de los estudios potencialmente elegibles; finalmente, se efectuó la lectura del texto completo para confirmar su inclusión definitiva. Las discrepancias entre los dos revisores fueron resueltas mediante consenso, sin requerir la participación de un tercer evaluador.

2.4. Estrategias de búsqueda

La búsqueda bibliográfica se realizó en Scopus, Google Scholar, ACM Digital Library y SpringerLink, seleccionadas por su cobertura en ingeniería de software, ciencias de la computación y tecnologías de la información. Las cadenas de búsqueda combinaron términos relacionados con chatbots, agentes

conversacionales y procesamiento del lenguaje natural con conceptos propios de ITSM, incluyendo service desk, help desk y gestión de incidentes, mediante los operadores booleanos AND y OR para maximizar la recuperación de estudios relevantes, tal como se muestra en la Tabla 1.

Las búsquedas se ejecutaron durante mayo de 2026, aplicando en todas las bases de datos filtros correspondientes al periodo 2020–2026 y al tipo de documento (artículos de revista y trabajos de conferencia revisados por pares). El término resiliencia operativa no se incorporó de manera independiente en las cadenas de búsqueda debido a su carácter transversal dentro de ITSM. En su lugar, este constructo fue analizado mediante variables reportadas en los estudios primarios, particularmente el tiempo medio de resolución (MTTR) y el cumplimiento de los acuerdos de nivel de servicio (SLA), considerados indicadores representativos de la resiliencia operativa en los procesos de soporte.

2.5. Recolección y extracción de la información

Los registros recuperados fueron exportados y consolidados en una hoja de cálculo para su depuración y organización. Tras la eliminación de

registros duplicados, los estudios seleccionados fueron sometidos al proceso de revisión descrito previamente. De cada publicación incluida se extrajeron variables relacionadas con autor(es), año de publicación, título, fuente, base de datos, tipo de estudio, proceso de ITSM abordado, técnica de inteligencia artificial utilizada, método de evaluación y evidencia relacionada con la resiliencia operativa. El proceso completo de identificación, selección e inclusión de los estudios se presenta mediante el diagrama de flujo PRISMA (Figura 1).

Como complemento al mapeo sistemático, se realizó un análisis bibliométrico descriptivo sobre los documentos recuperados mediante las ecuaciones de búsqueda en Scopus y Springer Nature, previo a la aplicación de los criterios de inclusión y exclusión. Este análisis tuvo como finalidad caracterizar el panorama general de la producción científica mediante indicadores relacionados con las áreas temáticas, palabras clave, distribución geográfica, disciplinas de investigación, principales temáticas y Objetivos de Desarrollo Sostenible (ODS). Los resultados obtenidos se presentan mediante estadísticas descriptivas expresadas en frecuencias y porcentajes.

Tabla 1

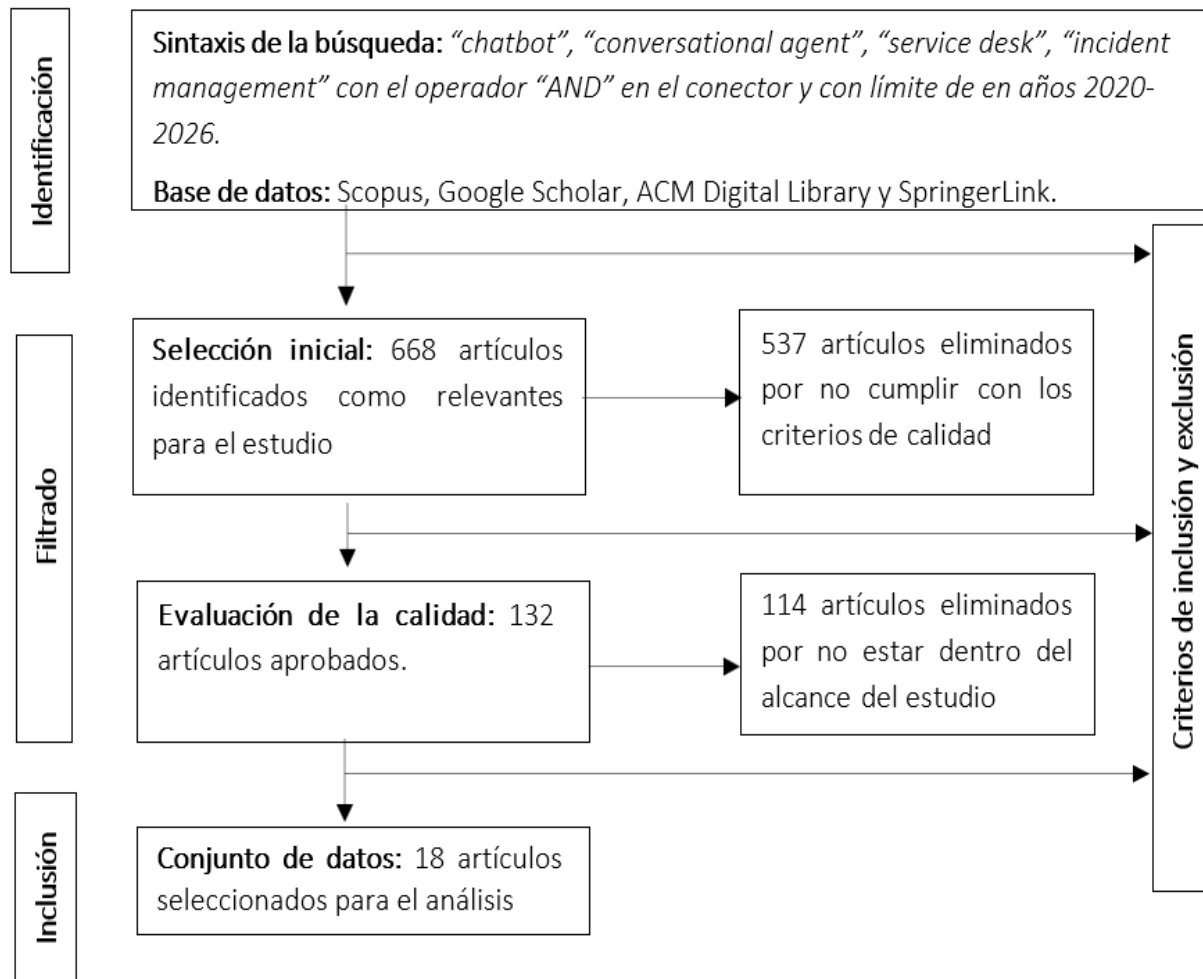
Estrategias de búsqueda empleadas en las bases de datos.

Base de datos	Términos de Búsqueda	Resultados	Seleccionados
Scopus	("chatbot*" OR "conversational agent" OR "virtual assistant" OR "NLP") AND ("IT Service Management" OR "ITSM" OR "service desk" OR "help desk" OR "incident management")	104	9
Google Scholar	allintitle: chatbot "service desk" OR "help desk" OR ITSM	16	3
ACM Digital Library	(Title:("chatbot" OR "conversational agent" OR "virtual assistant" OR "NLP") AND Title:("IT Service Management" OR "ITSM" OR "service desk" OR "help desk" OR "incident management"))	7	2
SpringerLink	("chatbot*" OR "conversational agent" OR "virtual assistant" OR "dialogue system" OR "conversational AI") AND ("IT Service Management" OR "ITSM" OR "service desk" OR "help desk" OR "incident management" OR "IT support" OR "ticketing system")	541	4



Figura 1

Diagrama de flujo del proceso de selección de estudios según la declaración PRISMA 2020



3. Resultados

3.1. Caracterización bibliométrica de la producción científica

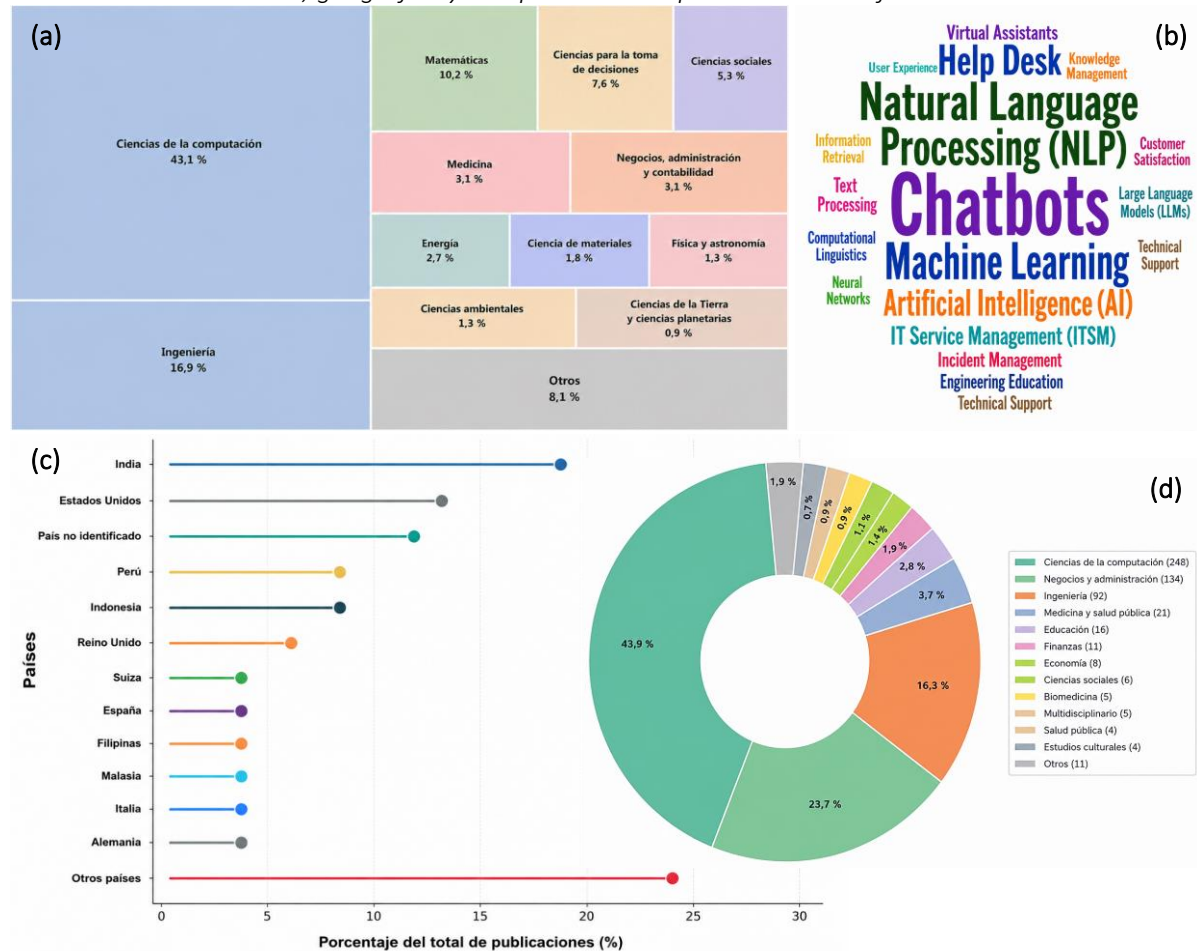
La Figura 2 presenta la caracterización bibliométrica de la producción científica recuperada mediante la estrategia de búsqueda en Scopus y Springer Nature. En Scopus (Figura 2a), las publicaciones se concentran principalmente en Ciencias de la Computación, seguidas por Ingeniería, Matemáticas y Ciencias de la Decisión, lo que evidencia el predominio de disciplinas relacionadas con la informática y el desarrollo tecnológico. Asimismo, las palabras clave con mayor frecuencia (Figura 2b) corresponden a Chatbot, *Natural Language Processing Systems*, *Help Desk*, *Natural Language Processing* y *Artificial Intelligence*, reflejando el enfoque de la investigación hacia la

automatización de procesos de ITSM mediante técnicas de inteligencia artificial y procesamiento del lenguaje natural.

En cuanto a la distribución geográfica (Figura 2c), India, Estados Unidos, Perú e Indonesia concentran el mayor número de publicaciones, evidenciando una participación internacional con mayor presencia de países que impulsan el desarrollo de soluciones basadas en inteligencia artificial. Por su parte, los resultados obtenidos en Springer Nature (Figura 2d) muestran que las investigaciones se desarrollan principalmente en las disciplinas de Ciencias de la Computación, Negocios y Administración e Ingeniería, mientras que las temáticas predominantes corresponden a Inteligencia Artificial, Inteligencia Computacional, Procesamiento del Lenguaje Natural e Industria 4.0.

Figura 2

Caracterización temática, geográfica y disciplinaria de la producción científica



3.2. Estudios incluidos en el mapeo sistemático

Los estudios analizados muestran una mayor orientación hacia la gestión de incidentes y triaje (55 %), seguida por el *service desk* y *help desk* (28 %) y la gestión del conocimiento y la seguridad (17 %) (Figura 3b). En cuanto a las tecnologías empleadas, predominan los modelos de lenguaje de gran escala (LLMs), seguidos por los modelos basados en transformadores y los motores de diálogo tradicionales (Figura 3c). Asimismo, la producción científica presenta un incremento a partir de 2023, con su mayor concentración en 2025 (Figura 3d). Finalmente, la relación con los Objetivos de Desarrollo Sostenible obtenida de Springer Nature del total de la ecuación de búsqueda evidencia un mayor número de publicaciones vinculadas con Educación de calidad, Buena salud y bienestar e Industria, innovación e infraestructura (Figura 3a).

El análisis de los estudios seleccionados (Tabla 2) evidencia una evolución en el desarrollo

y aplicación de chatbots dentro de la gestión de servicios de tecnologías de la información. Las investigaciones se orientan principalmente al diseño e implementación de agentes conversacionales para automatizar actividades de soporte y optimizar la atención de incidentes. En contraste, los trabajos más recientes incorporan modelos basados en transformadores, modelos de lenguaje de gran escala y arquitecturas híbridas, ampliando su aplicación hacia tareas de mayor complejidad, como la clasificación inteligente de tickets, la recuperación de conocimiento, la priorización de incidentes y la asistencia a la toma de decisiones. Asimismo, se observa una mayor diversidad de enfoques metodológicos, que abarcan estudios experimentales, desarrollos tecnológicos, casos de estudio, modelamientos matemáticos y revisiones sistemáticas, evidenciando la consolidación de esta línea de investigación.



Figura 3
Síntesis de los principales resultados del mapeo sistemático

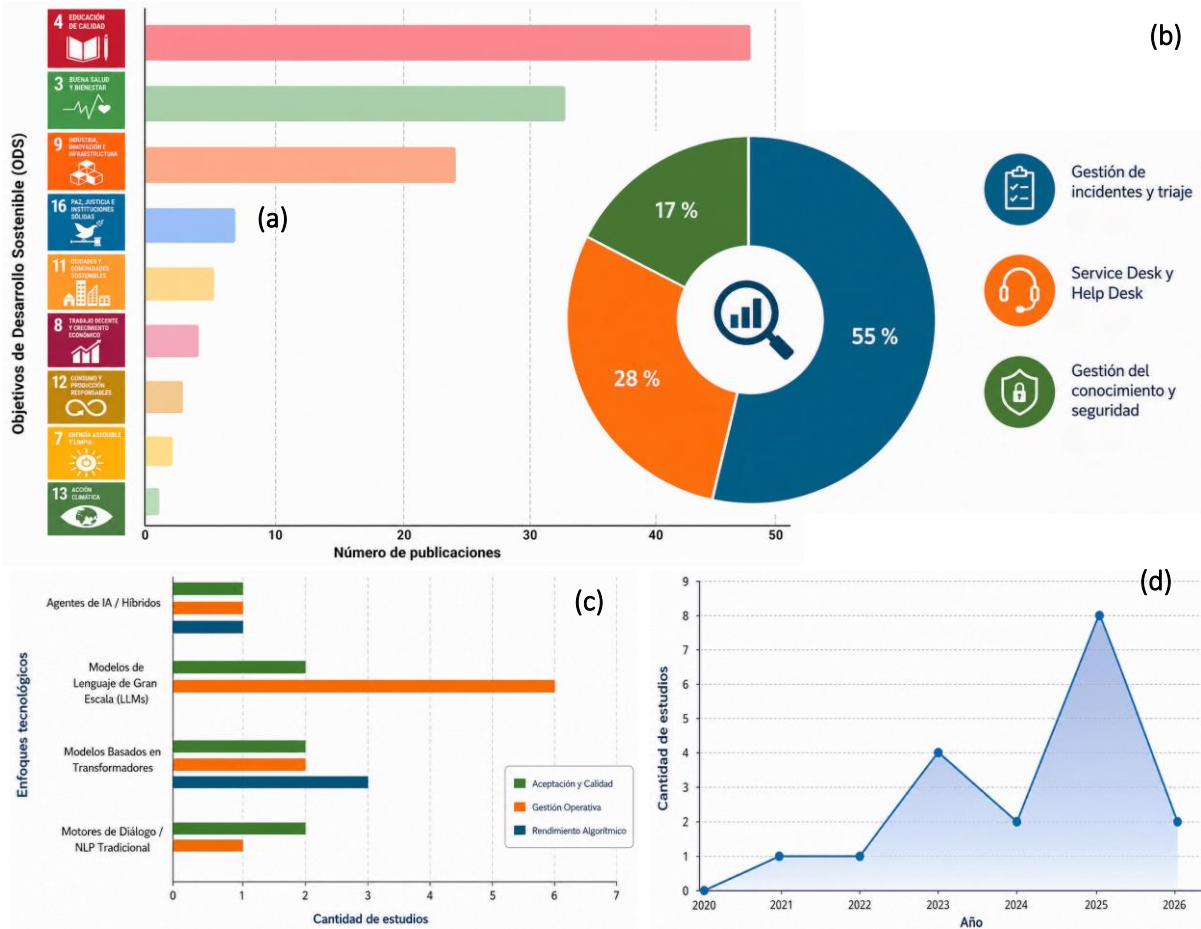


Tabla 2
Matriz de mapeo sistemático de los estudios seleccionados

Autor (Año)	Tipo de estudio	Técnica de IA / Tecnología	Proceso ITSM abordado	Hallazgos principales
Stütz y Riener (2021)	Experimental empírico	Chatbot textual basado en flujos e intenciones	Service Desk (UX y soporte interno)	La aceptación de los bots corporativos es equivalente al canal tradicional y supera a los asistentes de uso privado.
Malla (2025)	Investigación aplicada	IA Generativa e integración vía API (Zendesk)	Operaciones de Helpdesk y enrutamiento	Automatización de tareas repetitivas y optimización del direccionamiento de tickets de soporte técnico.
Saha y Hoi (2022)	Desarrollo tecnológico (AIOps)	Procesamiento de Lenguaje Natural Neuronal (Neural NLP / ICA)	Gestión de Incidentes (Análisis de Causa Raíz - RCA)	Extracción automatizada de conocimiento estructurado a partir de bitácoras históricas y registros de expertos (PRB).
Zangari et al. (2023)	Técnico comparativo	Arquitecturas basadas en BERT y Word Embeddings	Clasificación multicanal de soporte técnico	Los transformadores reducen drásticamente el error en la categorización jerárquica de solicitudes masivas.

Tabla 2 (Continuación/1)

Autor (Año)	Tipo de estudio	Técnica de IA / Tecnología	Proceso ITSM abordado	Hallazgos principales
Iparraguirre-Villanueva et al. (2023a)	Pre experimental	Agente Inteligente estructurado en Dialogflow	Gestión de Incidencias de TI	Reducción del porcentaje de incidentes sin resolver y optimización del tiempo de respuesta en entornos corporativos.
Ramirez y Andrade-Arenas (2023)	Caso de estudio y prototipo	Sistema multiplataforma con chatbot integrado	Procesos de Help Desk institucional	Validación experta aprueba la arquitectura híbrida (asistencia automatizada y escalado humano directo).
Iparraguirre-Villanueva et al. (2023b)	Revisión sistemática	Metodología PRISMA (Análisis de 47 artículos)	Gestión de Incidentes corporativos	Mapeo de la literatura científica (2019-2022) que vincula el uso de agentes conversacionales con la productividad.
Ordóñez-Camacho et al. (2024)	Desarrollo e implementación	Chatbot fundamentado en LLM corporativo	Help Desk de atención institucional	Capacidad avanzada para interpretar consultas ambiguas de usuarios mediante el procesamiento semántico.
Setyowibowo et al. (2025)	Caso de estudio cualitativo	Chatbot alineado al marco ITIL v4 y Scrum	Gestión de Incidentes en logística	Elevación del cumplimiento de los niveles de servicio (SLA) de respuesta desde un 64 % histórico hasta un 26 % de mejora directa.
Sagbay et al. (2025)	Técnico experimental	Ajuste fino (Fine-tuning) sobre modelo BERT	Service Desk (Asignación automática de casos)	Clasificación precisa de tickets técnicos utilizando un dataset mixto de 110,000 registros estructurados.
Pfño Alccahuamani et al. (2026)	Arquitectura híbrida web	API de OpenRouter (IA sin estado) y Telegram Bot	Gestión de Incidentes en el sector público	Reducción del tiempo medio de resolución técnica de 120 minutos a un promedio óptimo de 31 min.
Peralta et al. (2025)	Modelamiento matemático	NLP combinado con modelos matemáticos borrosos	Gestión de Incidentes empresariales	Optimización matemática de la asignación de recursos y soporte de primer nivel basado en conocimiento.
Razma y Jurkus (2026)	Empírico comparativo	Multilingual BERT y XLM-RoBERTa frente a ML	Clasificación de solicitudes en ESM/ITSM	Los transformadores superan a los modelos estadísticos tradicionales al procesar clases minoritarias desbalanceadas.

Tabla 2 (Continuación/2)

Autor (Año)	Tipo de estudio	Técnica de IA / Tecnología	Proceso ITSM abordado	Hallazgos principales
Bashar y Nayak (2025)	Desarrollo	Híbrido: BERT-QAM + ChatGPT-QAM	Gestión del Conocimiento	La integración extractiva/generativa reduce alucinaciones en reportes científicos.
Balasubramanian et al. (2025)	Revisión Sistemática	GenAI (LLMs, GANs, VAEs)	Cyber Threat Intelligence	Taxonomía de modelos para detectar malware y phishing de forma proactiva.
Patel et al. (2025)	Revisión Sistemática	LLMs (GPT-4, Claude)	Soporte TI y Operaciones Cloud	El sistema OASIS genera resúmenes 100 veces más rápido que un ingeniero humano.
Murdock (2025)	Crítico	Infraestructura de IA	Economía Política Digital	La IA opera bajo una lógica de extracción de datos y dominancia de grandes corporaciones.
Jain et al. (2024)	Desarrollo tecnológico y evaluación empírica	LLMs, DistilBERT, embeddings semánticos (Hugging Face BGE), clustering jerárquico adaptativo	Gestión de incidentes (clustering, priorización y resolución de tickets)	La integración de factores espaciales y temporales en algoritmos de similitud semántica.

La clasificación de las tecnologías de inteligencia artificial empleadas en los estudios seleccionados (Tabla 3) pone de manifiesto una evolución progresiva desde motores de diálogo basados en reglas hacia modelos fundamentados en transformadores, modelos de lenguaje de gran escala y arquitecturas híbridas. Mientras las primeras soluciones se orientaron principalmente a la automatización de tareas básicas del service desk, las tecnologías más recientes amplían su aplicación hacia la clasificación inteligente de tickets, la gestión del conocimiento y el soporte contextual al usuario. Dentro de esta evolución, los modelos de lenguaje de gran escala constituyen la categoría con mayor representación, seguidos por los modelos basados en transformadores, lo que evidencia el cambio de la investigación hacia enfoques con mayores capacidades de comprensión e inferencia del lenguaje.

Los estudios analizados emplean diversas métricas para evaluar el desempeño de los chatbots en entornos de ITSM (Tabla 4). La

evaluación se fundamenta principalmente en indicadores de rendimiento algorítmico, como Accuracy, Precision, Recall y F1-Score, utilizados para validar la clasificación automática de tickets y la detección de anomalías. Asimismo, los estudios basados en modelos de lenguaje de gran escala incorporan métricas de calidad textual, como ROUGE y BLEU, con el propósito de valorar la coherencia y fidelidad de las respuestas generadas. Desde una perspectiva operativa, destacan indicadores como el tiempo medio de resolución (MTTR), el cumplimiento de los acuerdos de nivel de servicio (SLA) y la resolución en el primer contacto, orientados a cuantificar el impacto de estas soluciones sobre el desempeño de los procesos de ITSM. De forma complementaria, varios estudios incluyen métricas relacionadas con la aceptación del usuario, la confiabilidad de las respuestas y la eficiencia computacional, reflejando una tendencia hacia evaluaciones que integran aspectos técnicos, operacionales y de experiencia de uso.

Tabla 3

Clasificación de las tecnologías de inteligencia artificial aplicadas a ITSM

Categoría tecnológica	Algoritmos y herramientas clave	Ventajas clave en el dominio ITSM	Estudios asociados y frecuencia (n)
Motores de Diálogo y NLP Tradicional	Dialogflow, reglas, flujos guiados.	Bajo costo, predictibilidad en service desk.	Stütz y Riener (2021); Ramirez y Andrade-Arenas (2023); Iparraguirre-Villanueva et al. (2023a). (n = 3)
Modelos Basados en Transformadores (Deep Learning)	BERT, Multilingual BERT, XLM-ROBERTa, embeddings.	Comprensión técnica profunda, manejo de tickets desbalanceados.	Zangari et al. (2023); Sagbay et al. (2025); Razma y Jurkus (2026); Jain et al. (2024); Saha y Hoi (2022). (n = 5)
Modelos de Lenguaje de Gran Escala (LLMs)	Arquitecturas GPT/Llama, integración RAG, APIs.	Procesamiento de consultas ambiguas, fluidez y adaptación.	Malla (2025); Ordóñez-Camacho et al. (2024); Setyowibowo et al. (2025); Pfuño et al. (2026); Peralta et al. (2025); Bashar y Nayak (2025); Patel et al. (2025). (n = 7)
Agentic AI / Híbridos	Agentes autónomos, BabyAGI, AutoGPT, RAG avanzado.	Autonomía total, reducción de alucinaciones.	Balasubramanian et al. (2025); Murdock (2025). (n = 2)

Tabla 4

Métricas y métodos de evaluación empleados sobre chatbots aplicados a ITSM

Categoría	Métricas clave	Rango / Valor referencial	Propósito de la evaluación	Estudios representativos
Rendimiento algorítmico	Accuracy, Precision, Recall, F1-Score, AUC (ROC)	F1-Score: 85–98 % / AUC: 0,92	Validar la clasificación de tickets y detección de anomalías ante datos desbalanceados	Razma y Jurkus (2026); Sagbay et al. (2025); Zangari et al. (2023)
Calidad textual	ROUGE-1, BLEU, Exact Match, Partial Match	ROUGE-1: 0,78–0,91 / BLEU: variable según corpus	Evaluar la fidelidad y coherencia de resúmenes y respuestas generadas por LLMs	Bashar y Nayak (2025); Patel et al. (2025)
Gestión operativa	MTTR, Tasa de cumplimiento SLA, Tasa de resolución en primer contacto, Porcentaje de incidentes sin resolver	MTTR: de 120 min a 31 min / SLA: mejora del 26 %	Medir el impacto real en tiempos de respuesta, continuidad del servicio y cumplimiento de acuerdos	Iparraguirre-Villanueva et al. (2023a); Pfuño Alccahuamani et al. (2026); Setyowibowo et al. (2025)
Aceptación humana	Modelo TAM, Escala SUS, PLS-SEM, Escala Likert, Net Promoter Score (NPS)	SUS: >70 puntos (aceptable) / TAM: correlación utilidad-adopción	Validar la intención de uso, usabilidad percibida y adopción real del sistema por el usuario	Ramirez y Andrade-Arenas (2023); Stütz y Riener (2021)



Tabla 4 (Continuación/1)

Categoría	Métricas clave	Rango / Valor referencial	Propósito de la evaluación	Estudios representativos
Factores críticos de calidad	Tasa de alucinación, Interpretabilidad XAI, Tasa de falsos positivos, Cobertura semántica	Tasa de alucinación: reducida mediante CFM / Falsos positivos: <10 % en modelos híbridos	Asegurar transparencia, confiabilidad y control de errores en decisiones automatizadas en entornos reales	Bashar y Nayak (2025)
Eficiencia computacional	Tiempo de inferencia, Latencia de respuesta, Costo de entrenamiento (GPU/h)	Latencia: <50 ms en tiempo real / Entrenamiento: ~120 s en modelos ajustados	Evaluar la viabilidad técnica y escalabilidad del despliegue en producción	Sagbay et al. (2025); Pfuño Alccahuamani et al. (2026)

4. Discusión

Los resultados del estudio indican que la evolución de los chatbots en la gestión de servicios de tecnologías de la información responde a un proceso de maduración tecnológica más que a una simple incorporación de nuevas herramientas de inteligencia artificial. Mientras las primeras aplicaciones se orientaban a automatizar consultas frecuentes y actividades propias del *service desk*, la literatura muestra un desplazamiento hacia soluciones con mayor capacidad para comprender consultas complejas, recuperar conocimiento y asistir procesos de decisión dentro de ITSM. Esta tendencia coincide con la evolución descrita por Caldarini et al. (2022), quienes señalan que el desarrollo del procesamiento del lenguaje natural ha ampliado progresivamente el alcance funcional de los agentes conversacionales, así como con la revisión realizada por Aldhafeeri et al. (2025), que identifica a los modelos de lenguaje de gran escala como el principal motor de la nueva generación de chatbots, destacando el uso de estrategias como Retrieval-Augmented Generation (RAG), el ajuste fino específico del dominio y la interacción contextual para mejorar su desempeño.

Los hallazgos también muestran que la incorporación de modelos generativos no implica el reemplazo de las tecnologías previamente utilizadas. Por el contrario, la coexistencia de

motores conversacionales tradicionales, modelos basados en transformadores y modelos de lenguaje de gran escala indica que cada enfoque responde a necesidades operativas distintas dentro de ITSM. Esta interpretación resulta consistente con la revisión de Mauricio Saavedra (2025), quien concluye que la efectividad de los chatbots generativos depende de su integración con los procesos de gestión de servicios y de la utilización de métricas que evalúen simultáneamente el desempeño técnico, la eficiencia operativa y la experiencia del usuario, en lugar de considerar únicamente indicadores de precisión del modelo.

Otro aspecto que merece atención es que la evolución tecnológica identificada en este estudio coincide con una ampliación del alcance funcional de los chatbots. La literatura reciente señala que los modelos generativos ya no se limitan a responder consultas, sino que participan en actividades como la recuperación inteligente de información, la asistencia contextual y la automatización de procesos complejos, siempre que se incorporen mecanismos de adaptación al dominio, supervisión humana y estrategias para reducir errores de generación. Estas observaciones respaldan la tendencia identificada en el presente mapeo sistemático y refuerzan la necesidad de considerar a los chatbots como componentes integrados dentro del ecosistema

de ITSM, más que como herramientas aisladas de automatización.

Asimismo, los resultados amplían la revisión sistemática de Iparraguirre-Villanueva et al. (2023b), centrada en estudios publicados entre 2019 y 2022, al evidenciar que la investigación reciente ha desplazado su interés desde la automatización de tareas operativas hacia aplicaciones que incorporan inteligencia artificial generativa, modelos de lenguaje de gran escala e integración con plataformas empresariales. Esta evolución sugiere que el desarrollo de los chatbots en ITSM está avanzando hacia soluciones con mayor capacidad de adaptación al contexto, aunque todavía persisten desafíos relacionados con la estandarización de las evaluaciones, la validación en entornos reales de operación y la disponibilidad de evidencia longitudinal que permita confirmar su impacto sostenido sobre los procesos de gestión de servicios.

No obstante, se presenta algunas limitaciones que deben considerarse al interpretar los resultados. En primer lugar, la estrategia de búsqueda dependió de los términos definidos en las ecuaciones utilizadas, por lo que estudios que emplearon una terminología diferente podrían no haber sido recuperados. En segundo lugar, aunque la búsqueda se realizó en Scopus, Google Scholar, ACM Digital Library y SpringerLink, y se restringió a publicaciones revisadas por pares en inglés y español entre 2020 y 2026, es posible que investigaciones relevantes indexadas exclusivamente en otras fuentes o publicadas en diferentes idiomas no hayan sido consideradas.

Las futuras investigaciones deberían orientarse hacia la evaluación de los chatbots en entornos organizacionales reales mediante estudios longitudinales, el desarrollo de marcos estandarizados para la evaluación de su desempeño y la integración de modelos de lenguaje de gran escala con marcos de gestión de servicios como ITIL, COBIT e ISO/IEC 20000. Asimismo, resulta pertinente profundizar en aspectos relacionados con la explicabilidad, la gobernanza y la seguridad de los sistemas basados en inteligencia artificial, así como ampliar la evidencia en sectores aún poco

explorados, como la administración pública, la salud y las infraestructuras.

5. Conclusiones

La aplicación de chatbots en ITSM ha evolucionado desde soluciones orientadas a la automatización del *service desk* hacia arquitecturas basadas en transformadores y modelos de lenguaje de gran escala, ampliando sus capacidades para la gestión de incidentes, la clasificación inteligente de tickets y la asistencia contextual al usuario.

La contribución de estas tecnologías a la resiliencia operativa depende de su integración con los procesos de gestión de servicios y no únicamente de sus capacidades conversacionales. En este sentido, las arquitecturas híbridas que combinan automatización con intervención humana constituyen el enfoque más consistente para mejorar la eficiencia operativa y la continuidad del servicio.

Persisten desafíos relacionados con la estandarización de los métodos de evaluación y la validación de estas soluciones en entornos organizacionales reales. En consecuencia, futuras investigaciones deberían priorizar estudios longitudinales, incorporar indicadores operativos y económicos, y fortalecer la integración de los modelos de inteligencia artificial con marcos de gestión de servicios como ITIL v4.

Contribución de los autores

A. Sánchez: Conceptualización, investigación, metodología, administración del proyecto, recursos y redacción del borrador original. **J. Ordóñez:** Conceptualización, investigación, metodología, administración del proyecto, recursos, visualización y redacción del borrador original. **A. Mendoza:** Supervisión, validación y revisión y edición del manuscrito.

Conflictos de interés

Los autores manifiestan no tener conflictos de interés.



6. Referencias bibliográficas

- Agnihotri, A., y Bhattacharya, S. (2024). Chatbots' effectiveness in service recovery. *International Journal of Information Management*, 76(102679), 102679. <https://doi.org/10.1016/j.ijinfomgt.2023.102679>
- Aldhafeeri, L., Aljumah, F., Thabyan, F., Alabbad, M., AlShahrani, S., Alanazi, F., y Al-Nafjan, A. (2025). Generative AI chatbots across domains: A systematic review. *Applied Sciences (Basel, Switzerland)*, 15(20), 11220. <https://doi.org/10.3390/app152011220>
- Balasubramanian, P., Liyana, S., Sankaran, H., Sivaramakrishnan, S., Pusuluri, S., Pirttikangas, S., y Peltonen, E. (2025). Generative AI for cyber threat intelligence: applications, challenges, and analysis of real-world case studies. *Artificial Intelligence Review*, 58(11). <https://doi.org/10.1007/s10462-025-11338-z>
- Bashar, M. A., y Nayak, R. (2025). Chatting with organisational data: a generative AI approach applied to scientific reports for information seeking. *Knowledge and Information Systems*, 67(11), 10657–10690. <https://doi.org/10.1007/s10115-025-02551-x>
- Bruni, R., Bianchi, G., y Papa, P. (2023). Hyperparameter black-box optimization to improve the automatic classification of support tickets. *Algorithms*, 16(1), 46. <https://doi.org/10.3390/a16010046>
- Caldarini, G., Jaf, S., y McGarry, K. (2022). A literature survey of recent advances in chatbots. *Information (Basel)*, 13(1), 41. <https://doi.org/10.3390/info13010041>
- Iparraguirre-Villanueva, O., Obregon-Palomino, L., Pujay-Iglesias, W., y Cabanillas-Carbonell, M. (2023a). Agente inteligente para la gestión de incidencias. *RISTI - Revista Ibérica de Sistemas e Tecnologías de Informação*, (51), 99–115. <https://doi.org/10.17013/risti.51.99-115>
- Iparraguirre-Villanueva, O., Obregon-Palomino, L., Pujay-Iglesias, W., Sierra-Liñan, F., y Cabanillas-Carbonell, M. (2023b). Productivity of incident management with conversational bots-a review. *IAES International Journal of Artificial Intelligence (IJ-AI)*, 12(4), 1543. <https://doi.org/10.11591/ijai.v12.i4.pp1543-1556>
- Jain, S., Gupta, A., y Neha, K. (2024, del 8 al 11 de octubre). AI enhanced ticket management system for optimized support [conferencia]. *AIMLSystems '24: Proceedings of the 4th International Conference on AI-ML System*, Baton Rouge, Louisiana, Estados Unidos. <https://doi.org/10.1145/3703412.3703433>
- Kitchenham, B., y Charters, S. (2007). *Guidelines for performing Systematic Literature Reviews in Software Engineering* [Technical Report EBSE-2007-01, Keele University and University of Durham]. https://legacyfileshare.elsevier.com/promis_misc/525444systematicreviewsguide.pdf
- Malla, P. (2025). Improving IT support efficiency through an AI-powered ITSM chatbot. *Journal of Advances in Developmental Research*, 16(2). <https://doi.org/10.71097/ijaidr.v16.i2.1662>
- Mauricio Saavedra, K. M., Lulichac Ramos, G. M. F., y Mendoza de los Santos, A. C. (2025). Evaluación sistemática de la eficacia operativa y la experiencia del usuario de chatbots generativos en plataformas ITSM. *Ingeniería Investiga*, 7. <https://doi.org/10.47796/ing.v7i00.1193>
- Murdock, G. (2025). Artificial intelligence as primitive accumulation: enclosure, extraction, exploitation. *Communication and Change*, 1(1). <https://doi.org/10.1007/s44382-025-00004-1>

- (EASE). Station, Road Swindon, Inglaterra.
<https://doi.org/10.14236/ewic/EASE200>
8.8



- 30
 - 31
 - 32
 - 33
 - 34
 - 35
 - 36
 - 37
 - 38
 - 39
 - 40
 - 41
 - 42
 - 43
 - 44
 - 45
 - 46
 - 47
 - 48
 - 49
 - 50
 - 51
 - 52
 - 53
 - 54
 - 55
 - 56
 - 57
 - 58
 - 59
 - 60
 - 61
 - 62
 - 63
 - 64
 - 65
 - 66
 - 67
 - 68
 - 69
 - 70
 - 71
 - 72
 - 73
 - 74
 - 75
 - 76
 - 77
 - 78
 - 79
 - 80
 - 81
 - 82
 - 83
 - 84
 - 85
 - 86
 - 87
 - 88
 - 89
 - 90
 - 91
 - 92
 - 93
 - 94
 - 95
 - 96
 - 97
 - 98
 - 99
 - 100
- Pennsylvania, Estados Unidos.
<https://doi.org/10.1145/3510457.3513030>
- Setyowibowo, D., Raharjo, T., y Fitriani, A. N. (2025). Project management strategy for chatbot implementation based on ITIL v4: A logistics company case. *CommIT (Communication and Information Technology) Journal*, 19(1), 59–70.
<https://doi.org/10.21512/commit.v19i1.12106>
- Shazzad, M. (2023). Impact of ITSM Solutions and ITIL Framework on Small to Medium-Sized Organizations: An Empirical Study on Challenges & Benefits [Tesis de maestría, Hochschule Neu-Ulm University of Applied Sciences].
https://publications.hnu.de/3086/1/Maister's%20Thesis%20of%20Mohammad%20Shazzad%20Hossen_01.pdf
- Stütz, L., y Riener, A. (2021, del 5 al 8 de setiembre). Let's chat internal: User acceptance of an in-company service desk chatbot [conferencia]. *MuC '21: Proceedings of Mensch und Computer 2021*, Ingolstadt, Alemania.
<https://doi.org/10.1145/3473856.3474005>
- Zangari, A., Marcuzzo, M., Schiavinato, M., Gasparetto, A., y Albarelli, A. (2023). Ticket automation: An insight into current research with applications to multi-level classification scenarios [conferencia]. *Expert Systems With Applications*, 225(119984), 119984.
<https://doi.org/10.1016/j.eswa.2023.119984>