

# Uso del procesamiento del lenguaje natural para mejorar la eficiencia de los servicios digitales: Una Revisión Sistemática de la Literatura

## Using Natural Language Processing to Enhance the Efficiency of Digital Services: A Systematic Literature Review

**Thalia Díaz Tunjar**

<https://orcid.org/0000-0002-2402-7940>

[thalia.diaz@unmsm.edu.pe](mailto:thalia.diaz@unmsm.edu.pe)

Universidad Nacional Mayor de San Marcos, Lima, Perú

RECIBIDO: 31/10/2023 - ACEPTADO: 25/11/2023 - PUBLICADO: 30/12/2023

### RESUMEN

El estudio se centra en el uso del Procesamiento de Lenguaje Natural (NLP) para la mejora de la eficiencia en los servicios digitales. El objetivo principal es evaluar y analizar los estudios publicados en un periodo de cinco años sobre el NLP y su influencia en los servicios digitales, incluyendo los medios de publicación predominantes, las revistas de alto impacto, las palabras clave recurrentes y la contribución geográfica en este campo. Para llevar a cabo este estudio, se realizó siguiendo las directrices de Kitchenham y Charters. La búsqueda se enfocó en bases de datos académicas reconocidas, utilizando ecuaciones de búsqueda avanzada con palabras clave relevantes. Los artículos seleccionados fueron evaluados según criterios específicos y se extrajo la información relevante de cada artículo utilizando la herramienta Mendeley Desktop. Los resultados de la investigación revelaron una tendencia creciente en la publicación de estudios sobre NLP en revistas de alto impacto, con una predominancia de los Estados Unidos, China, India y el Reino Unido en la contribución a este campo. Las palabras clave recurrentes destacaron la interconexión entre el NLP, Machine Learning y Deep Learning. Asimismo, se evidenció que el NLP está adquiriendo relevancia en la mejora de los servicios digitales, lo que sugiere un futuro prometedor para su aplicación en la transformación digital.

**Palabras clave:** Procesamiento de lenguaje natural, NLP, machine learning, servicios digitales.

### ABSTRACT

The study focuses on the use of Natural Language Processing (NLP) to enhance the efficiency of digital services. The primary goal is to evaluate and analyze research published over a five-year span regarding NLP and its impact on digital services. This includes the predominant publishing mediums, high-impact journals, recurring keywords, and the geographical contribution in this domain. The research was conducted following the guidelines set by Kitchenham and Charters. The search was centered on well-established academic databases, using advanced search queries with pertinent keywords. Selected articles were assessed based on specific criteria, and key information from each article was extracted using the Mendeley Desktop tool. Research outcomes highlighted an upward trend in publishing NLP studies in high-impact journals, with a notable contribution from the United States, China, India, and the United Kingdom. The recurring keywords underscored the link between NLP, Machine Learning, and Deep Learning. Additionally, the rising significance of NLP in improving digital services was observed, indicating a bright future for its role in digital transformation.

**Keywords:** Natural Language Processing, NLP, machine learning, digital services.

## I. INTRODUCCIÓN

En un mundo digitalmente conectado y en constante evolución, la eficiencia de los servicios digitales se ha vuelto crucial para la satisfacción de los usuarios y el éxito de las organizaciones en todos los sectores. Es fundamental comprender y utilizar de manera efectiva el flujo de información en línea para mejorar la calidad de los servicios digitales y brindar experiencias excepcionales a los usuarios. En este contexto, el procesamiento del lenguaje natural emerge como un campo clave de la inteligencia artificial con el potencial de transformar la eficiencia de los servicios digitales.

La revisión sistemática de la literatura tiene por objetivo evaluar y analizar la influencia del procesamiento del lenguaje natural en la mejora de la eficiencia de los servicios digitales. Para contextualizar esta revisión, es esencial considerar las investigaciones y desarrollos previos relacionados con NLP que han contribuido a la comprensión actual del tema.

Uno de los antecedentes clave es el trabajo de Akther et al. (2022) que se centra en el desarrollo de un corpus monolingüe de Bangla llamado "KUMono". Este modelo ha sido analizado en detalle y se ha demostrado su utilidad en la categorización de artículos, superando a los modelos existentes. Además, se discuten las futuras aplicaciones de este recurso, lo que demuestra cómo la investigación en NLP puede generar recursos valiosos para mejorar los servicios digitales.

Otro antecedente importante es el modelo propuesto por Alawad et al. (2021), que utiliza el aprendizaje profundo (DL) para extraer características del cáncer sin comprometer la privacidad de los registros. Este modelo ofrece un rendimiento comparable a los métodos tradicionales y permite la colaboración entre diferentes registros de cáncer. Esto ilustra cómo el NLP puede contribuir a la mejora de los servicios de atención médica y la investigación clínica.

El trabajo de Arias-Barahona et al. (2023) propone un modelo de clasificación de incidentes que utiliza el aprendizaje automático (ML) y el NLP en el área de soporte técnico de una empresa de desarrollo de software. El modelo logra una precisión impresionante del 98.97%, lo que destaca cómo la aplicación de NLP puede agilizar el tiempo de respuesta y la resolución de problemas de los clientes, mejorando así la eficiencia de los servicios de atención al cliente.

Estos antecedentes, junto con otros trabajos notables como el modelo generativo de Bhardwaj et al. (2020) que aborda la ambigüedad y la diversidad lingüística en NLP, y la revisión de representaciones semánticas en VerbNet de Brown et al. (2022), resaltan la relevancia y diversidad de las aplicaciones de NLP en una variedad de campos y sectores.

En resumen, esta revisión sistemática de la literatura tiene como propósito consolidar y sintetizar el conocimiento actual sobre el uso del procesamiento del lenguaje natural para mejorar la eficiencia de los servicios digitales. Con una base sólida en antecedentes bibliográficos, se explorarán los avances recientes y las tendencias emergentes en este emocionante campo científico, proporcionando una visión completa de su impacto y potencial.

## II. MATERIALES Y MÉTODOS

### A. Materiales

#### Procesamiento de lenguaje natural

El procesamiento de lenguaje natural se define como el campo de estudio que se ocupa de la interacción entre las computadoras y el lenguaje humano. Según Jurafsky y Martin (2013) en su libro "Speech and Language Processing" (3ª edición), el procesamiento de lenguaje natural abarca la comprensión y generación de lenguaje humano, así como la extracción de información y el análisis de sentimientos a partir de texto o voz.

La importancia del NLP radica en su capacidad para abordar la variabilidad, la ambigüedad y la complejidad inherentes al lenguaje humano. La investigación en este campo, como el estudio realizado por Jager (2023), que se enfoca en el concepto de "ruido semántico" y argumenta que este debe ser reconocido y aceptado como una parte esencial del lenguaje, tiene influencia en el diseño y propósito de los modelos de lenguaje. Además, el trabajo de Giannopoulou et al. (2020) destaca cómo las redes neuronales profundas pueden ser utilizadas para mejorar la clasificación de contenido temático, lo que demuestra la capacidad del NLP para adaptarse y evolucionar.

El NLP también ha demostrado su relevancia en áreas como la salud, donde Houssein et al. (2021) analizan su uso en registros de salud electrónicos para extraer información clínica valiosa. Además, su aplicación en la administración pública, como se muestra en el estudio de Kowalski et al. (2020),

revela el potencial de estas técnicas para mejorar la calidad de los servicios públicos.

### Servicios Digitales

Se refiere a cualquier tipo de servicio o contenido que se ofrece en formato digital a través de plataformas en línea. Los servicios digitales abarcan una amplia gama de áreas, desde el comercio electrónico y la banca en línea hasta la comunicación y el entretenimiento en línea, la educación en línea, la atención médica virtual y mucho más. Estos servicios utilizan tecnologías avanzadas, como el aprendizaje automático y el procesamiento de lenguaje natural Kaczmarek et al. (2022) destaca la mejora en la experiencia del usuario y ofrecer servicios más personalizados y eficientes.

El crecimiento de los servicios digitales ha sido impulsado en parte por la investigación y la innovación tecnológica, como se destaca en estudios como el de Nam et al. (2024), que examina los avances en el aprendizaje multimodal y los modelos de aprendizaje profundo en la traducción bidireccional entre imágenes y procesamiento de lenguaje natural.

Además, los servicios digitales enfrentan desafíos en términos de idiomas y características específicas de las regiones, como se señala en la investigación

de Larabi et al. (2019), que aborda los desafíos del procesamiento de lenguaje natural en árabe y la necesidad de herramientas y conjuntos de datos especializados.

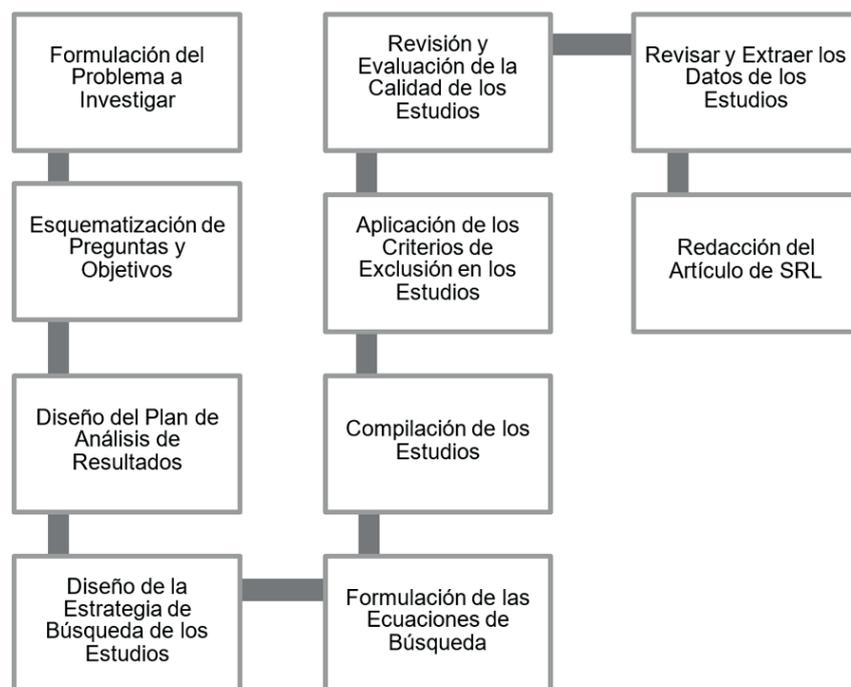
### Herramientas Usadas

En cuanto a la investigación, es importante destacar el empleo de Mendely Desktop como la herramienta para la organización de artículos, así como también el uso de Inteligencia Artificial Generativas para aprender las características y patrones presentes en los datos de los estudios, además permite analizar la información utilizando diferentes parámetros y criterios establecidos en la investigación.

### B. Método de la Revisión Sistemática

El método de revisión utilizado en este estudio se basa en las directrices de Kitchenham y Charters (2009) para llevar a cabo una revisión sistemática de la literatura (SRL). Este enfoque integral aborda de manera efectiva los desafíos relacionados con la investigación, las fuentes de búsqueda, los criterios de selección de los estudios, la elección de estudios, la evaluación de calidad de los estudios, las estrategias de extracción de datos de los estudios y compilación de resultados. La figura 1 muestra claramente el flujo de la SRL.

**Figura 1.**  
*Proceso de la Revisión Sistemática.*



Fuente: Elaboración Propia

### Formulación del Problema a Investigar

Durante la revisión sistemática de la literatura, la pregunta de investigación es esencial para definir la formulación de la estrategia de búsqueda, extracción y análisis de datos. Es decir, el investigador necesita contar con una comprensión clara de la motivación y los objetivos del estudio (Tabla 1).

### Fuentes de Búsqueda y Estrategias de Búsqueda

Se utilizaron diversas bibliotecas, como: Web of Science, IEEE Xplore, Google Scholar, ScienceDirect, Scopus y Taylor & Francis, para búsqueda de artículos y/o revistas científicas de investigación.

El proceso de búsqueda fue efectuado de acuerdo a las ecuaciones de búsqueda específicas del estudio, detalladas en la Tabla 3.

La estrategia de búsqueda se fundamentó en el empleo de palabras clave relevantes. "natural language processing" se consideró como variable independiente, mientras que "digital services" se trató como variable dependiente, como se detalla en la Tabla 2. Estas palabras clave se emplearon para optimizar y reestructurar el contenido.

La Tabla 3 muestra las ecuaciones de búsqueda avanzada y las fuentes correspondientes, haciendo uso de los operadores booleanos "AND" y "OR".

**Tabla 1.**  
*Preguntas y objetivos de la investigación.*

Pregunta de Investigación	Objetivo
<b>RQ1:</b> ¿Cuántos estudios se publicaron en un periodo de cinco años de antigüedad sobre procesamiento de lenguaje natural y su impacto en la mejora de la eficiencia de los servicios digitales?	Encontrar la cantidad de estudios que fueron publicados en los últimos cinco años para medir el uso del procesamiento de lenguaje natural y su impacto en la mejora de la eficiencia de los servicios digitales.
<b>RQ2:</b> ¿Cuáles son los principales medios de publicación (Journal o Conference) de las investigaciones sobre procesamiento de lenguaje natural y su impacto en la mejora de la eficiencia de los servicios digitales?	Investigar los medios de publicación (Journal o Conference) utilizados para dar a conocer las investigaciones sobre Procesamiento de lenguaje natural y su impacto en mejora en la eficiencia de los servicios digitales.
<b>RQ3:</b> ¿Cuáles son los niveles de cuartil de las revistas donde se han publicado las investigaciones sobre el uso del procesamiento de lenguaje natural y su impacto en la mejora de la eficiencia de los servicios digitales?	Determinar en qué nivel de cuartil se encuentran las revistas académicas donde se han publicado investigaciones sobre el uso del procesamiento de lenguaje natural en la mejora de la eficiencia de los servicios digitales.
<b>RQ4:</b> ¿Cuáles son las palabras clave más frecuentes en las investigaciones publicadas sobre el uso del procesamiento de lenguaje natural y su impacto en la mejora de la eficiencia de los servicios digitales?	Identificar las palabras clave utilizadas con mayor frecuencia en las investigaciones públicas sobre el uso del procesamiento de lenguaje natural y su impacto en la mejora en la eficiencia de los servicios digitales.
<b>RQ5:</b> ¿Cuáles son los países con mayor frecuencia en la publicación de investigaciones sobre el uso del procesamiento de lenguaje natural y su impacto en la mejora de la eficiencia de los servicios digitales?	Estudiar los países que han sido más activos en la investigación y aplicación del procesamiento de lenguaje natural en la mejora de la eficiencia de los servicios digitales.

Fuente: Elaboración propia.

**Tabla 2.**  
*Palabras clave relevantes.*

Descriptor	Descripción	Sinónimos
natural language processing	Variable Independiente (A)	NLP/machine learning/ expert system
digital services	Variable Dependiente (B)	virtual services /public service

Fuente: Elaboración propia.

**Tabla 3.**  
*Ecuaciones de búsqueda avanzada con las fuentes correspondientes.*

Fuente	Ecuación de Búsqueda Avanzada
<b>Web of Science</b>	(ALL=("natural language processing" ) OR ALL=("NLP") OR ALL=("machine learning") OR ALL=("expert system") AND (ALL=("digital services" ) OR ALL=("virtual services") OR ALL=("public service")))
<b>IEEE Xplore</b>	("Full Text & Metadata": "natural language processing" OR "Full Text & Metadata": "NLP" OR "Full Text & Metadata": "machine learning" OR "Full Text & Metadata": "expert system") AND ("Full Text & Metadata": "digital services" OR "Full Text & Metadata": "virtual services" OR "Full Text & Metadata": "public service"))
<b>ScienceDirect</b>	"natural language processing" OR "NLP" OR "machine learning" OR "expert system" AND ("digital services" OR "virtual services" OR "public service")
<b>Scopus</b>	TITLE-ABS-KEY ("natural language processing" OR "NLP" OR "machine learning" OR "expert system") AND ("digital services" OR "virtual services" OR "public service")
<b>Taylor &amp; Francis</b>	"natural language processing" OR "NLP" OR "machine learning" OR [All: "expert system"] AND [All: "digital services"] OR [All: "virtual services"] OR [All: "public service"]]

### Criterios de Elección de los Estudios

Con el propósito de garantizar que los artículos seleccionados estén de acorde con la motivación y los objetivos de la investigación se definen los siguientes criterios de exclusión:

CE1: Los artículos cuentan con más de 5 años de antigüedad O

CE2: Los artículos son de Revisión Sistemática de Literatura o Revisión Bibliométrica O

CE3: Los títulos y las palabras clave de los artículos no son los adecuados O

CE4: Los artículos carecen de relación con el tema propuesto O

CE5: Los artículos carecen de una metodología y/o modelo definidos O

CE6: El resumen de los artículos no proporciona información significativa

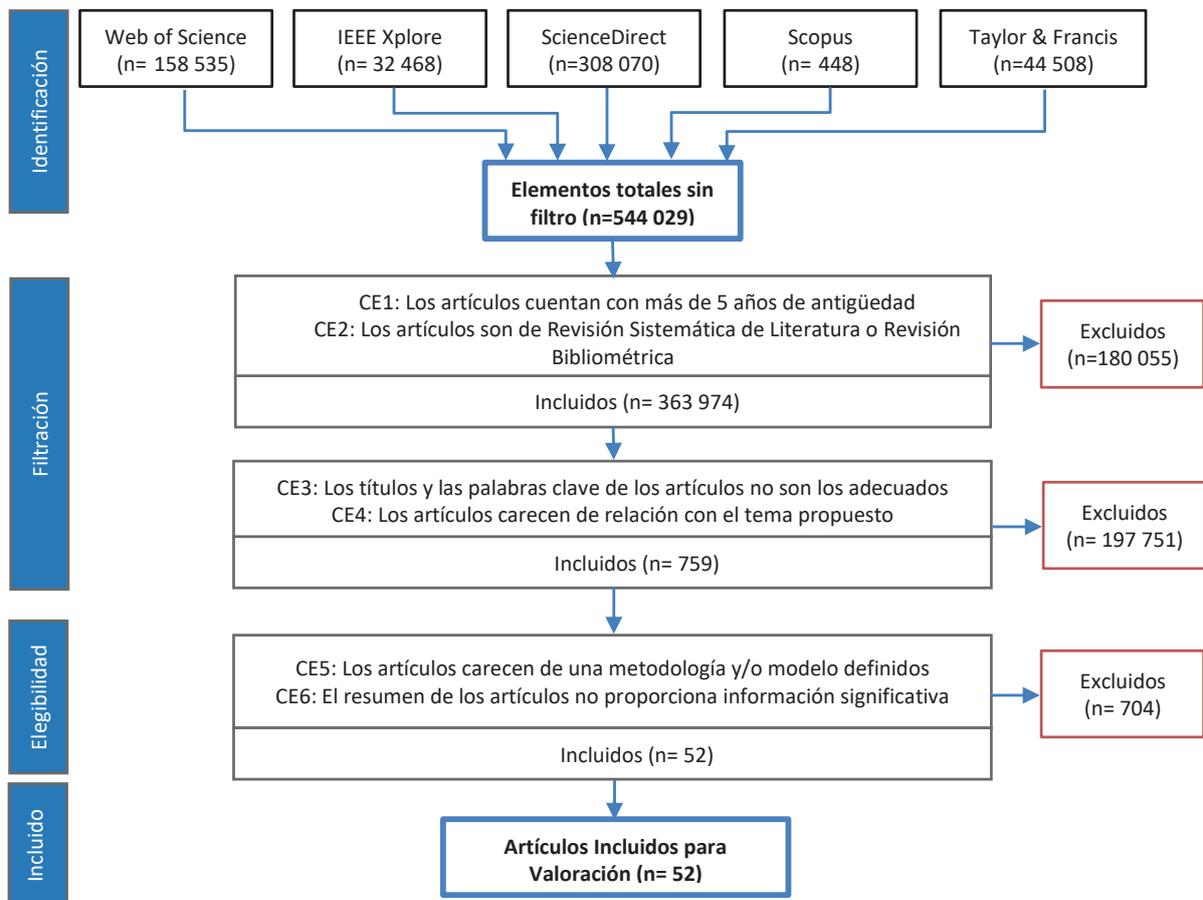
### Elección de Estudios

Aplicando las ecuaciones búsquedas específicas se obtuvieron 544 029 artículos utilizando las palabras clave. Posteriormente se llevaron a cabo varios pasos para identificar, filtrar y seleccionar los artículos. El resultado fue de 52 artículos incluidos, tal como se visualiza en la Figura 2.

### Evaluación de la Calidad de los Estudios

En este estudio, se establecieron y aplicaron criterios de calidad (QA) para determinar de los artículos que forman parte de esta revisión final. Estos QA se emplearon para evaluar la pertinencia de los artículos de investigación en relación con las preguntas de investigación específicas planteadas.

**Figura 2.**  
Gráfico PRISMA de artículos seleccionados.



Fuente: Elaboración Propia

Se emplearon un total de seis QA con el fin de asegurar la calidad de los artículos seleccionados.

QA1: ¿La estructura del documento es adecuada?

QA2: ¿Los objetivos de investigación se presentan de manera clara en el documento?

QA3: ¿Se proporciona suficiente contexto histórico en el documento?

QA4: ¿La definición del área temática específica es precisa y clara?

QA5: ¿Se incluyen experimentos prácticos en el documento?

QA6: ¿La metodología experimental es apropiada y estándar?

Estos criterios se aplicaron para asegurar la calidad de los artículos seleccionados y garantizar que cumplieran con los estándares requeridos para la investigación.

### Estrategias de Extracción de Datos de los Estudios

De acuerdo a los objetivos planteados de la investigación, los artículos seleccionados sirvieron para

abordar las preguntas de formuladas. De cada estudio se recopiló información esencial como: Id, nombre de la revista, título, autores, fuente, país, URL, número de páginas, idioma, año, resumen, palabras clave y tipo de publicación. La herramienta utilizada para gestionar los artículos fue Mendeley, tal como se representa en la Figura 3.

### Compilación de Resultados

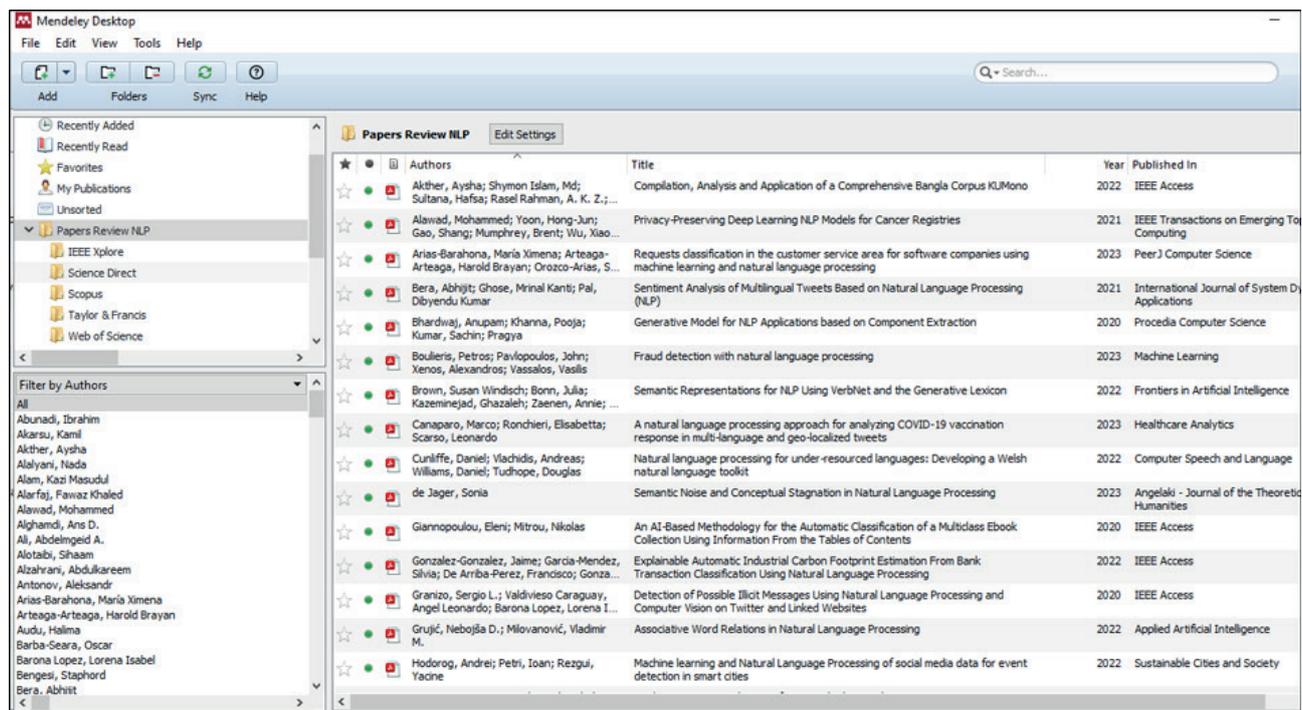
La información recopilada fue organizada y estructurada con el propósito de resolver las preguntas de la RQ1 a la RQ5 planteadas en la investigación, con el fin de obtener datos cuantitativos. Posteriormente, los datos se sometieron a análisis para identificar tendencias en el uso del procesamiento del lenguaje natural para mejorar la eficiencia de los servicios digitales.

## III. RESULTADO

### A. Descripción General de los Estudios

Las fuentes más importantes para las investigaciones sobre procesamiento del lenguaje natural y su impacto para mejorar los servicios digitales son los artículos de revistas científicas y los trabajos presentados en conferencias especializadas. Estas

Figura 3. Reporte de Mendeley Desktop.



Fuente: Elaboración Propia

fuentes académicas proporcionan información rigurosa y actualizada sobre los avances y hallazgos en el campo del procesamiento del lenguaje natural y su aplicación en los servicios digitales, como se puede observar en la Figura 4.

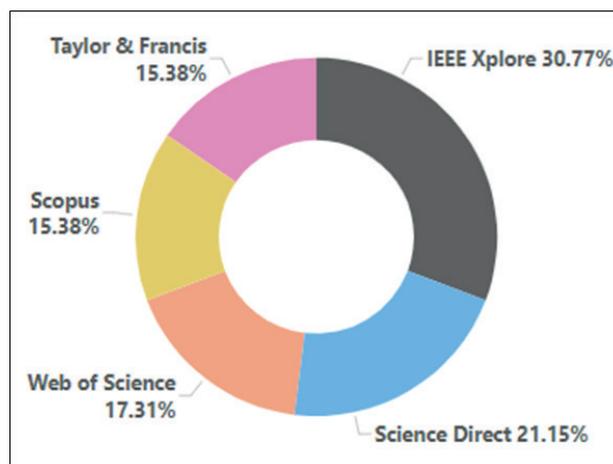
Según las fuentes de búsqueda empleadas en este estudio resalta que IEEE Xplore fue la fuente más prolífica en esta revisión sistemática, con un total de 16 artículos, lo que equivale al 30.77% del conjunto total. Le siguió Science Direct con un total de 11 artículos, lo que representa el 21.15%. Web of Science contribuyó con 9 artículos, equivalente al 17.31% de los documentos revisados. Por otra parte, Scopus y Taylor&Francis enriqueció el estudio con 8 artículos cada uno, representando el 15.38% del total por cada uno.

## B. Preguntas y Respuestas de la Investigación

**RQ1: ¿Cuántos estudios se publicaron en un periodo de cinco años de antigüedad sobre procesamiento de lenguaje natural y su impacto en la mejora de la eficiencia de los servicios digitales?**

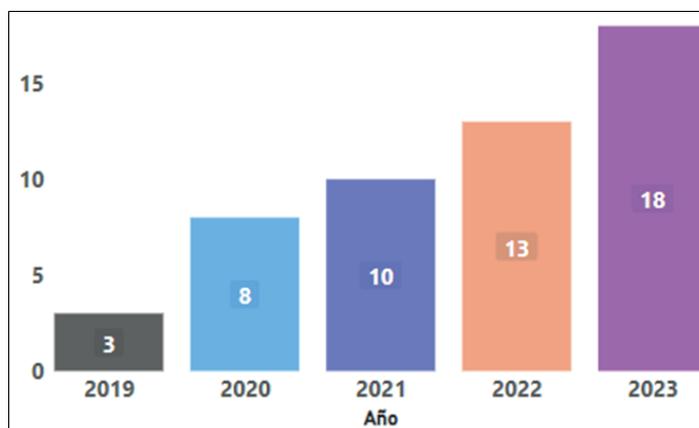
En los últimos cinco años, se han publicado artículos que exploran el impacto del procesamiento de lenguaje natural en diversas industrias y sectores. Dichos estudios examinan las aplicaciones de NLP en áreas como la salud, la educación, la administración pública, la tecnología y más. Estos estudios reflejan el creciente interés en el potencial del NLP para mejorar la calidad y eficiencia de los servicios digitales en estas áreas. Esto demuestra que el NLP es una tecnología versátil capaz de

Figura 4.  
Fuentes de Investigación.



Fuente: Elaboración Propia

Figura 5.  
Artículos publicados por año.



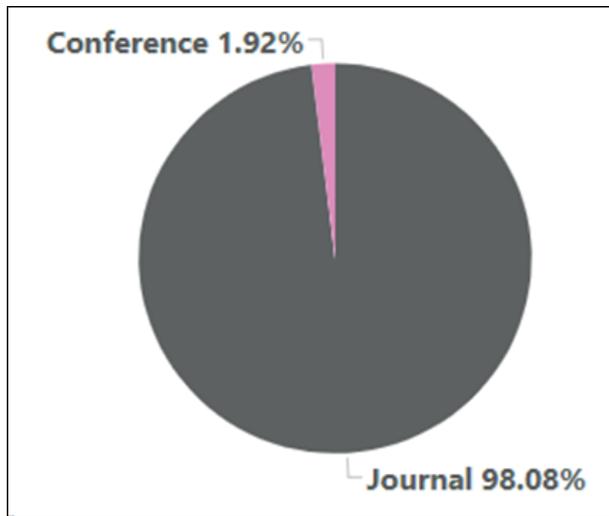
Fuente: Elaboración Propia

transformar múltiples sectores, mejorando la comunicación, la toma de decisiones y la experiencia del usuario en diversas industrias.

**RQ2: ¿Cuáles son los principales medios de publicación (Journal o Conference) de las investigaciones sobre procesamiento de lenguaje natural y su impacto en la mejora de la eficiencia de los servicios digitales?**

Al analizar los datos de los artículos sobre procesamiento de lenguaje natural y su impacto, se puede observar que la mayoría de los estudios se publican en revistas académicas de tipo Journal como lo muestra la figura 6. Esto indica que la comunidad académica está liderando el avance en este campo, centrándose en la investigación fundamental y teórica en NLP. Aunque las conferencias también son importantes, las revistas son la plataforma principal para la presentación y discusión exhaustiva de investigaciones en NLP. Esto demuestra la necesidad de un riguroso escrutinio académico en el desarrollo de técnicas de procesamiento de lenguaje natural y su aplicación en diferentes sectores e industrias.

Figura 6. Medios de publicación.



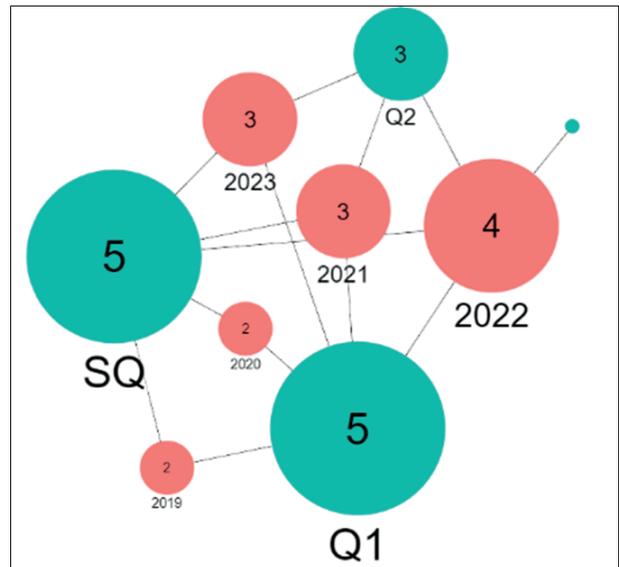
Fuente: Elaboración Propia

**RQ3: ¿Cuáles son los niveles de cuartil de las revistas donde se han publicado las investigaciones sobre el uso del procesamiento de lenguaje natural y su impacto en la mejora de la eficiencia de los servicios digitales?**

La tendencia muestra un incremento constante en las publicaciones en revistas de cuartil Q1 en los últimos años, lo que refleja el crecimiento y el reconocimiento de la importancia del NLP en

servicios digitales. Es evidente que el campo de NLP ha ganado tracción, y la comunidad académica está contribuyendo activamente a este dominio en revistas de alto impacto. Esta tendencia también resalta la relevancia y el potencial del NLP en la transformación y mejora de servicios digitales tal como se puede ver en la figura 7.

Figura 7. Cuartil por año de las revistas analizadas.



Fuente: Elaboración Propia

La mayoría de las publicaciones (un total de 31) se encuentra en el cuartil Q1, lo que representa un 59.6% de las referencias analizadas. Esto indica que una gran parte de las investigaciones sobre NLP se publican en revistas altamente reconocidas. El cuartil Q2, con 7 referencias, representa el 13.5%. Por último, las publicaciones en el cuartil SQ (sin cuartil especificado) suman 4, es decir, el 7.7%.

**RQ4: ¿Cuáles son las palabras clave más frecuentes en las investigaciones publicadas sobre el uso del procesamiento de lenguaje natural y su impacto en la mejora de la eficiencia de los servicios digitales?**

Las palabras clave más frecuentes en las investigaciones proporcionadas sobre el impacto del procesamiento de lenguaje natural en los servicios digitales son: natural language processing (30 ocurrencias), machine learning (9 ocurrencias), deep learning (6 ocurrencias), sentiment análisis (3 ocurrencias) y neural network (3 ocurrencias) y artificial intelligence (3 ocurrencias).

Estas palabras clave destacan los principales temas de investigación en el campo del NLP como se

muestra en la figura 8. El término "Natural Language Processing" es, con mucho, el más frecuente, lo que es esperado dado el enfoque de las investigaciones. Además, términos como "Machine Learning" y "Deep Learning" sugieren que estas metodologías son esenciales en la investigación actual de NLP.

**RQ5: ¿Cuáles son los países con mayor frecuencia en la publicación de investigaciones sobre el uso del procesamiento de lenguaje natural y su impacto en la mejora de la eficiencia de los servicios digitales?**

El país que presenta mayor coocurrencia es Estados Unidos debido a que lidera en la cantidad de investigaciones en NLP, seguido de China, India y el Reino Unido. Estos países son reconocidos por su fuerte presencia en el ámbito de la investigación tecnológica tal como lo muestra la figura 9. Resulta interesante destacar la diversidad geográfica, con la representación de países de casi todos los continentes, lo cual indica la importancia global del NLP.

El NLP es una disciplina crucial en la era de la información, y su impacto se evidencia en la cantidad y diversidad de investigaciones producidas en todo el mundo. Estos datos reflejan una clara tendencia hacia la colaboración internacional en este campo,

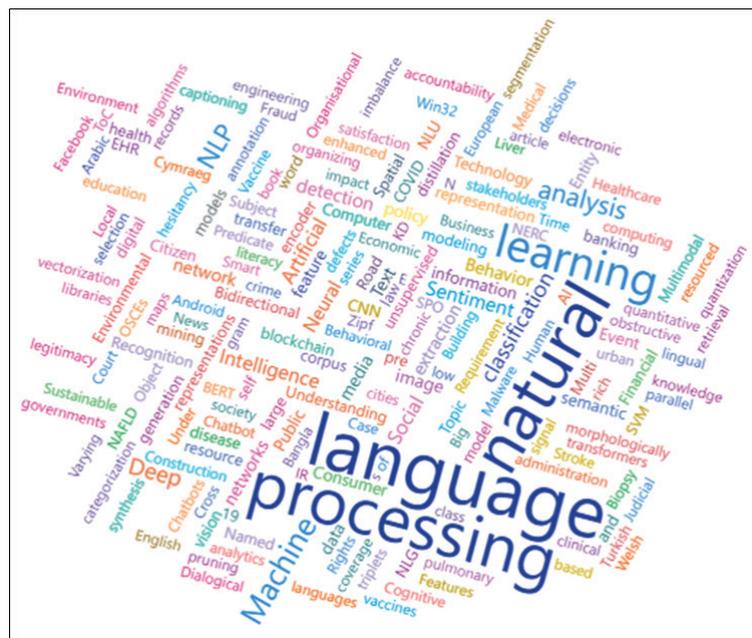
lo que puede conducir a avances más rápidos y soluciones más inclusivas en el futuro.

**IV. CONCLUSIONES**

La revisión sistemática de la literatura realizada en este estudio ha brindado una visión detallada de la investigación relacionada con el procesamiento del lenguaje natural y cómo este influye en la mejora de los servicios digitales en diversos sectores. Los resultados obtenidos de esta revisión resaltan la importancia del NLP como una tecnología versátil con el potencial de transformar múltiples industrias y mejorar la calidad de los servicios digitales.

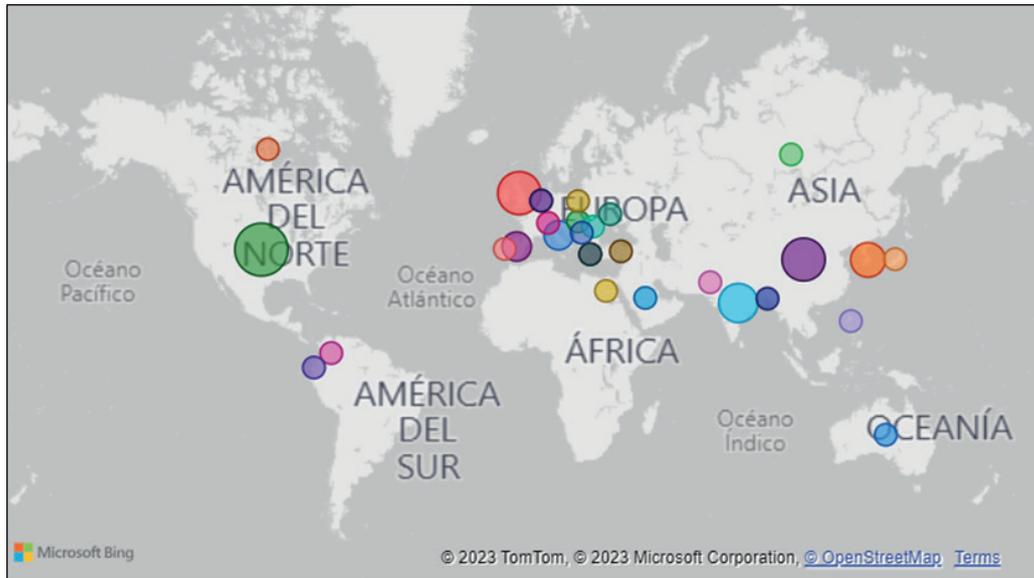
Se ha observado un incremento constante en la cantidad de estudios publicados en los últimos cinco años que exploran las aplicaciones del NLP en diferentes áreas, como la salud, la administración pública, la seguridad ferroviaria, la gestión de desastres y más (Kowalski et al., 2020; Nam et al., 2024; Olusegun et al., 2023; Qurashi et al., 2023; Shahbazi y Byun, 2022; Shahbazi y Byun, 2021; Shoostarian et al., 2023; Singh et al., 2021; Song et al., 2022; Wang et al., 2020). Estos estudios reflejan el creciente interés en aprovechar el poder del procesamiento de lenguaje natural para abordar desafíos específicos y mejorar la eficiencia de los servicios digitales.

Figura 8. Palabras clave recurrentes encontradas.



Fuente: Elaboración Propia

**Figura 9.**  
Palabras clave recurrentes encontradas.



Fuente: Elaboración Propia

La mayoría de las investigaciones se publican en revistas académicas de alto impacto, lo que subraya el papel fundamental de la comunidad académica en el avance de la investigación en NLP (Nam et al., 2024). Además, se ha observado un aumento constante en la publicación de investigaciones en revistas de Cuartil Q1, lo que demuestra el crecimiento y el reconocimiento de la relevancia del NLP en la mejora de los servicios digitales (Nam et al., 2024).

Las palabras clave más recurrentes en estos estudios incluyen "Procesamiento de lenguaje natural," "Aprendizaje automático," "Aprendizaje profundo," "Análisis de sentimientos" y "Redes neuronales", lo cual destaca la importancia de estas áreas en la investigación actual de NLP (Olusegun et al., 2023). Además, se ha identificado una creciente colaboración internacional en este campo, con una diversidad geográfica de países representados en la investigación (Olusegun et al., 2023).

## REFERENCIA

- [1] Akther, A., Shymon, M., Sultana, H., Rasel, AK., Saha, S., Alam, KM., Debnath, R. (2022). Compilation, Analysis and Application of a Comprehensive Bangla Corpus KUMono. <https://doi.org/10.1109/ACCESS.2022.3195236>
- [2] Alawad, M., Yoon, HJ., Gao, S., Mumphrey, B., Wu XC., Durbin EB., Jeong, JC., Hands I., Rust D., Coyle L., Penberthy L., Tourassi G. (2021). Privacy-Preserving Deep Learning NLP Models for Cancer Registries. <https://doi.org/10.1109/TETC.2020.2983404>
- [3] Arias-Barahona, MX., Arteaga-Arteaga, HB., Orozco-Arias, S., Flórez-Ruíz, JC., Valencia-Díaz, MA., Tabares-Soto, R. (2023). Requests classification in the customer service area for software companies using machine learning and natural language processing. <https://doi.org/10.7717/peerj-cs.1016>
- [4] Bhardwaj, A., Khanna, P., Kumar, S., Pragma. (2020). Generative Model for NLP Applications based on Component Extraction. <https://doi.org/10.1016/j.procs.2020.03.391>
- [5] Brown, SW., Bonn, J., Kazeminejad, G., Zaenen, A., Pustejovsky, J., Palmer, M. (2022). Semantic Representations for NLP Using VerbNet and the Generative Lexicon. <https://doi.org/10.3389/frai.2022.821697>
- [6] de Jager, S. (2023). Semantic Noise and Conceptual Stagnation in Natural Language Processing. <https://doi.org/10.1080/0969725X.2023.2216555>
- [7] Giannopoulou, E., Mitrou, N. (2020). An AI-Based Methodology for the Automatic

- Classification of a Multiclass Ebook Collection Using Information From the Tables of Contents. <https://doi.org/10.1109/ACCESS.2020.3041651>
- [8] Houssein, EH., Mohamed, RE., Ali, AA. (2021). Machine Learning Techniques for Biomedical Natural Language Processing: A Comprehensive Review. <https://doi.org/10.1109/ACCESS.2021.3119621>
- [9] Kaczmarek, I., Iwaniak, A., Świetlicka, A., Piwowarczyk, M., Nadolny, A. (2022). A machine learning approach for integration of spatial development plans based on natural language processing. <https://doi.org/10.1016/j.scs.2021.103479>
- [10] Kowalski, R., Esteve, M., Jankin, S. (2020). Improving public services by mining citizen feedback: An application of natural language processing. <https://doi.org/10.1111/padm.12656>
- [11] Larabi, S., Alalyani, N., Alotaibi, S., Ghouzali, S., Abunadi, I. (2019). Arabic Natural Language Processing and Machine Learning-Based Systems. <https://doi.org/10.1109/ACCESS.2018.2890076>
- [12] Nam, W., Jang, B. (2024). A survey on multimodal bidirectional machine learning translation of image and natural language processing. <https://doi.org/10.1016/j.eswa.2023.121168>
- [13] Olusegun, R., Oladunni, T., Audu, H., Houkpati, YAO., Bengesi, S. (2023). Text Mining and Emotion Classification on Monkeypox Twitter Dataset: A Deep Learning-Natural Language Processing (NLP) Approach. <https://doi.org/10.1109/ACCESS.2023.3277868>
- [14] Qurashi, AW., Farhat, ZA., Holmes, V., Johnson, AP. (2023). New Avenues for Automated Railway Safety Information Processing in Enterprise Architecture: An NLP Approach. <https://doi.org/10.1109/ACCESS.2023.3272610>
- [15] Shahbazi, Z., Byun, YC. (2022). Blockchain-Based Event Detection and Trust Verification Using Natural Language Processing and Machine Learning. <https://doi.org/10.1109/ACCESS.2021.3139586>
- [16] Shahbazi, Z., Byun, YC. (2021). Fake Media Detection Based on Natural Language Processing and Blockchain Approaches. <https://doi.org/10.1109/ACCESS.2021.3112607>
- [17] Shooshtarian, S., Gurmu, AT., Sadick, AM. (2023). Application of natural language processing in residential building defects analysis: Australian stakeholders' perceptions, causes and types. <https://doi.org/10.1016/j.engappai.2023.107178>
- [18] Singh, S., Mahmood, A. (2021). The NLP Cookbook: Modern Recipes for Transformer Based Deep Learning Architectures. <https://doi.org/10.1109/ACCESS.2021.3077350>
- [19] Song, D., Gao, S., He, B., Schilder, F. (2022). On the Effectiveness of Pre-Trained Language Models for Legal Natural Language Processing: An Empirical Study. <https://doi.org/10.1109/ACCESS.2022.3190408>
- [20] Wang, D., Su, J., Yu, H. (2020). Feature Extraction and Analysis of Natural Language Processing for Deep Learning English Language. <https://doi.org/10.1109/ACCESS.2020.2974101>
- [21] Kitchenham, B., Charters, S. (2009). Guidelines for performing Systematic Literature Reviews in Software Engineering. EBSE Technical Report EBSE-2007-01. School of Computer Science and Mathematics, Keele University".
- [22] Jurafsky, D., Martin, J. (2013): Speech and Language Processing, Third Edition, New Jersey, Prentice Hall. Retrieved from: <https://web.stanford.edu/~jurafsky/slp3/>.