



# PROCESAMIENTO DEL LENGUAJE NATURAL

*Aplicaciones prácticas de PLN en investigación y  
en el mercado final*

## Objetivos

Analizar diferentes enfoques de generación automática de resúmenes mediante una recopilación bibliográfica sistematizada, desarrollando la capacidad de aplicar conocimientos adquiridos para resolver problemas en entornos multidisciplinarios del procesamiento del lenguaje natural. Evaluar las tendencias en el estado del arte y extraer conclusiones sobre los principales retos y avances en la generación de resúmenes, fomentando el trabajo en equipo y la autonomía en la investigación.

Jose Manuel Pinillos Rubio  
José Hernández Balboa  
Alejandro Benito Trujillo

Asignatura	Datos del alumno	Fecha
Procesamiento del Lenguaje Natural	Apellidos: Pinillos Rubio	07 de febrero de 2025
	Nombre: Jose Manuel	

# ÍNDICE

<b>ÍNDICE.....</b>	<b>2</b>
<b>1. INTRODUCCIÓN .....</b>	<b>3</b>
<b>2. ANÁLISIS DE ESTUDIOS SOBRE LA GENERACIÓN AUTOMÁTICA DE RESÚMENES .....</b>	<b>4</b>
2.1 Generador automático de resúmenes científicos en investigación turística .....	4
2.2 Generación automática de resúmenes: retos, propuestas y experimentos .....	5
2.3 <i>A comprehensive review of automatic text summarization techniques: method, data, evaluation and coding</i> .....	7
<b>3. ANÁLISIS DE SOLUCIONES COMERCIALES PARA EL AUTOCOMPLETADO Y LA GENERACIÓN AUTOMÁTICA DE RESÚMENES .....</b>	<b>9</b>
3.1 OpenAI Codex .....	9
3.2 Amazon Prime Video - X-Ray Recaps .....	10
<b>4. REFLEXIÓN SOBRE LA EVOLUCIÓN TECNOLÓGICA Y SU IMPACTO EN EL MERCADO .....</b>	<b>12</b>
<b>BIBLIOGRAFÍA .....</b>	<b>14</b>

Asignatura	Datos del alumno	Fecha
Procesamiento del Lenguaje Natural	Apellidos: Pinillos Rubio	07 de febrero de 2025
	Nombre: Jose Manuel	

# 1. INTRODUCCIÓN

El *Procesamiento de Lenguaje Natural (PLN)* ha evolucionado significativamente, permitiendo la implementación de modelos avanzados para tareas como el **autocompletado de texto** y la **generación automática de resúmenes**. En sus primeras etapas, los enfoques basados en reglas y métodos estadísticos, como *TextRank*, se empleaban para la síntesis de texto, pero presentaban limitaciones en la capacidad de reformulación y contextualización de la información. Con el desarrollo del *aprendizaje profundo*, surgieron modelos más sofisticados, capaces de predecir secuencias de texto con mayor precisión y generar resúmenes más fluidos y estructurados.

Este trabajo analiza el estado actual de estas dos aplicaciones del PLN, explorando su desarrollo en la literatura científica y su implementación en soluciones comerciales. Para ello, se revisan diferentes estudios que comparan los enfoques **extractivos y abstractivos** en la generación de resúmenes, destacando el impacto de modelos *transformer* como **BART, PEGASUS y T5**. En el ámbito del autocompletado, se examina cómo las arquitecturas basadas en *deep learning* han permitido la predicción contextualizada de texto, optimizando la escritura asistida y la generación de código.

Además de la revisión de estudios académicos, se investigan aplicaciones comerciales que han integrado estas tecnologías en sus productos. Se analizan los casos de **OpenAI Codex**, un modelo diseñado para el autocompletado de código, y **Amazon Prime Video - X-Ray Recaps**, que aplica PLN para generar resúmenes de contenido audiovisual. A través de esta comparación, se busca identificar la relación entre los avances en investigación y su implementación en el mercado, evaluando los desafíos y oportunidades en la integración de estos modelos en productos de uso masivo.

Finalmente, se reflexiona sobre la evolución de estas tecnologías y los retos pendientes en su aplicación. Se abordan cuestiones como la fidelidad del contenido generado, la eficiencia computacional y la reducción de sesgos en los modelos. Este análisis permite comprender el impacto del PLN en la mejora de la accesibilidad a la información y su potencial para optimizar la interacción con el lenguaje en distintos ámbitos.

Asignatura	Datos del alumno	Fecha
Procesamiento del Lenguaje Natural	Apellidos: Pinillos Rubio	07 de febrero de 2025
	Nombre: Jose Manuel	

## 2. ANÁLISIS DE ESTUDIOS SOBRE LA GENERACIÓN AUTOMÁTICA DE RESÚMENES

El desarrollo de modelos para la generación automática de resúmenes ha evolucionado considerablemente, pasando de enfoques estadísticos tradicionales a sofisticadas arquitecturas basadas en *deep learning*. En este contexto, las investigaciones han explorado la eficacia de los modelos *transformer* en la síntesis de información, permitiendo la generación de resúmenes más precisos y coherentes. Este apartado analiza un estudio que compara diferentes enfoques de PLN para la generación de resúmenes extractivos y abstractivos, evaluando su rendimiento mediante métricas estándar y determinando su aplicabilidad en distintos contextos.

### 2.1 Generador automático de resúmenes científicos en investigación turística

El estudio examina la aplicación del PLN en la automatización de la generación de resúmenes científicos en el ámbito del turismo. Para ello, se basa en técnicas de *aprendizaje profundo* (*Deep Learning*), integrando **redes neuronales recurrentes (RNNs)** con el fin de mejorar la coherencia y contextualización de los resúmenes generados. Además, el sistema implementado incorpora arquitecturas basadas en *transformers*, específicamente **BERT y T5**, que optimizan la representación semántica del texto y permiten obtener resúmenes más estructurados y comprensibles. La combinación de estos modelos avanzados permite no solo extraer la información clave de los documentos científicos, sino también mejorar la calidad lingüística y la fluidez de los textos generados, asegurando que el contenido resumido conserve la esencia del original (Carmona et al., 2022).

Para validar la efectividad del modelo, se llevó a cabo un experimento en el que se implementó un sistema de generación automática de resúmenes aplicado a artículos científicos especializados en turismo. El procedimiento metodológico inició con la recopilación de un **corpus compuesto por publicaciones científicas** obtenidas de bases de datos indexadas, asegurando que los documentos analizados fueran relevantes y de alta calidad. Antes de ser introducidos en el modelo basado en *transformers*, los textos fueron sometidos a un proceso de **preprocesamiento** que incluyó **tokenización y normalización**, lo que garantizó una estructura adecuada para su procesamiento computacional. Posteriormente, la capacidad del sistema para

Asignatura	Datos del alumno	Fecha
Procesamiento del Lenguaje Natural	Apellidos: Pinillos Rubio	07 de febrero de 2025
	Nombre: Jose Manuel	

generar resúmenes de calidad fue medida mediante métricas como **ROUGE**, las cuales permiten evaluar la similitud entre los resúmenes generados automáticamente y los redactados por los autores de los artículos, asegurando un análisis cuantitativo del desempeño del modelo (Carmona et al., 2022).

Los resultados obtenidos en la investigación reflejan un desempeño positivo del sistema en términos de precisión y calidad del resumen generado. Se lograron puntuaciones satisfactorias en **ROUGE-1, ROUGE-2 y ROUGE-L**, lo que indica que el modelo es capaz de captar la información esencial del texto original con un alto grado de fidelidad. Se observó, además, que los modelos **abstractivos** basados en *transformers* lograron generar contenido más estructurado y natural en comparación con los enfoques **extractivos**, los cuales tienden a limitarse a la selección de fragmentos específicos del texto. No obstante, el estudio también identificó que los modelos abstractivos presentan ciertas dificultades en la selección precisa de información clave, lo que puede afectar la fidelidad del resumen en algunos casos. Estos hallazgos son especialmente relevantes en el contexto de la investigación turística, ya que el sistema desarrollado optimiza el acceso a información clave dentro de un gran volumen de documentos, facilitando la síntesis de conocimiento en el área y mejorando la eficiencia en la consulta de artículos científicos (Carmona et al., 2022).

## 2.2 Generación automática de resúmenes: retos, propuestas y experimentos

El análisis presentado en esta investigación profundiza en la evolución de las técnicas empleadas en la generación automática de resúmenes, comparando el desempeño de los enfoques extractivos y abstractivos. Mientras que los modelos extractivos seleccionan oraciones clave sin alterar la estructura original del texto, los abstractivos reformulan la información mediante redes neuronales avanzadas, lo que permite generar resúmenes más naturales y adaptados al contexto. La validación de estos métodos se realiza a través de métricas estandarizadas como **ROUGE**, que miden la similitud entre los resúmenes generados automáticamente y aquellos elaborados manualmente por expertos, proporcionando una evaluación cuantitativa de la calidad del contenido sintetizado (Nikolaevna et al., n.d.).

Asignatura	Datos del alumno	Fecha
Procesamiento del Lenguaje Natural	Apellidos: Pinillos Rubio	07 de febrero de 2025
	Nombre: Jose Manuel	

Para analizar el desempeño de los modelos, la investigación implementa un experimento que evalúa distintos enfoques de resumen automático. Se trabaja con un **corpus compuesto por documentos académicos y noticias periodísticas**, lo que permite examinar la capacidad de los modelos para ajustarse a distintos estilos de escritura y niveles de formalidad. La metodología experimental inicia con una fase de **preprocesamiento de los textos**, en la que se aplican técnicas de tokenización y limpieza antes de alimentar los modelos. A continuación, se implementan métodos extractivos como **TextRank**, así como modelos abstractivos basados en *deep learning*, con el fin de comparar los diferentes enfoques. La evaluación de los resúmenes generados se realiza mediante métricas cuantitativas como **ROUGE**, analizando aspectos como la precisión, coherencia y cobertura de la información clave. Además, el estudio incorpora una evaluación cualitativa, en la que especialistas en lingüística y PLN valoran la legibilidad y utilidad de los resúmenes obtenidos, permitiendo una visión más detallada de su aplicabilidad en distintos contextos (Nikolaevna et al., n.d.).

Los resultados obtenidos muestran diferencias significativas entre los dos enfoques analizados. Los modelos extractivos, como **TextRank**, presentan un rendimiento favorable en métricas como **ROUGE-1 y ROUGE-L**, destacando por su capacidad para preservar la estructura y la información clave del texto original. Sin embargo, los modelos abstractivos basados en *deep learning* generan resúmenes más comprensibles y con mejor fluidez, aunque en algunos casos introducen alteraciones en el contenido original, lo que puede comprometer la fidelidad del resumen. A pesar de estas limitaciones, se concluye que los modelos abstractivos representan un avance significativo en términos de legibilidad y coherencia textual, aunque aún requieren mejoras en la conservación de la información esencial. Estos resultados son especialmente relevantes para aplicaciones donde se prioriza la comprensión general del contenido sobre la reproducción exacta de las frases originales, como en la **síntesis de artículos científicos y la generación de resúmenes informativos en medios digitales** (Nikolaevna et al., n.d.).

Asignatura	Datos del alumno	Fecha
Procesamiento del Lenguaje Natural	Apellidos: Pinillos Rubio	07 de febrero de 2025
	Nombre: Jose Manuel	

## 2.3 A comprehensive review of automatic text summarization techniques: method, data, evaluation and coding

Este trabajo examina en profundidad las estrategias utilizadas en la generación automática de resúmenes, abarcando desde métodos estadísticos tradicionales hasta la aplicación de modelos avanzados de *deep learning*. En particular, se estudian diversas arquitecturas basadas en *transformers*, evaluando su rendimiento en distintos conjuntos de datos y comparando su eficacia en la producción de resúmenes extractivos y abstractivos. Para determinar la calidad y coherencia de los textos generados, se emplean métricas de evaluación como **ROUGE**, **BLEU** y **METEOR**, que permiten cuantificar la similitud entre los resúmenes automáticos y los de referencia. Además, la investigación analiza modelos como **BART**, **GPT**, **T5** y **PEGASUS**, examinando su capacidad para sintetizar información en distintos escenarios y estableciendo la evolución de estas tecnologías dentro del PLN (Cajueiro et al., 2023).

Desde el punto de vista metodológico, el estudio combina una **revisión sistemática** con un **análisis experimental**, lo que permite evaluar el desempeño de diversas arquitecturas de *transformers* aplicadas a la generación de resúmenes. Para ello, se utilizan bases de datos ampliamente reconocidas en la comunidad de PLN, como **CNN/DailyMail Corpus** y **XSum dataset**, que contienen artículos y noticias acompañados de sus respectivos resúmenes de referencia. El proceso experimental incluye el **entrenamiento y prueba de modelos como BART, GPT, T5 y PEGASUS**, evaluando su capacidad para abstraer información y generar resúmenes de alta calidad. La validación de los modelos se lleva a cabo mediante métricas como **ROUGE**, **BLEU** y **METEOR**, que permiten analizar la precisión y relevancia del contenido sintetizado, proporcionando un marco cuantitativo para comparar el desempeño de los modelos analizados (Cajueiro et al., 2023).

Los resultados obtenidos muestran diferencias notables en el rendimiento de los modelos evaluados. Se destaca que **BART y PEGASUS** logran los mejores resultados en la generación de resúmenes abstractivos, debido a su capacidad para producir textos bien estructurados y comprensibles. En contraste, **T5 y GPT** presentan un desempeño sólido en tareas de resumen, aunque en ciertos casos generan información no presente en el texto original, lo que disminuye la fidelidad de los resultados. A partir de estos hallazgos, el estudio concluye que, si bien los modelos avanzados de PLN han mejorado notablemente la calidad de los resúmenes generados

Asignatura	Datos del alumno	Fecha
Procesamiento del Lenguaje Natural	Apellidos: Pinillos Rubio	07 de febrero de 2025
	Nombre: Jose Manuel	

automáticamente, aún persisten desafíos relacionados con la **fidelidad del contenido y la reducción de sesgos en la selección de información**. Estos resultados son de gran relevancia para la aplicación de estas tecnologías en distintos contextos, como la **automatización de resúmenes en noticias, artículos científicos y documentos empresariales**, proporcionando directrices claras sobre sus ventajas y limitaciones para optimizar su implementación en entornos reales (Cajueiro et al., 2023).



Asignatura	Datos del alumno	Fecha
Procesamiento del Lenguaje Natural	Apellidos: Pinillos Rubio	07 de febrero de 2025
	Nombre: Jose Manuel	

### 3. ANÁLISIS DE SOLUCIONES COMERCIALES PARA EL AUTOCOMPLETADO Y LA GENERACIÓN AUTOMÁTICA DE RESÚMENES

El PLN ha permitido el desarrollo de soluciones comerciales aplicadas a múltiples sectores, desde la asistencia en la escritura de código hasta la generación automática de resúmenes de contenido multimedia. En esta sección se analizarán dos productos que utilizan técnicas avanzadas de PLN: **OpenAI Codex**, enfocado en el autocompletado de código, y **Amazon Prime Video - X-Ray Recaps**, diseñado para la generación automática de resúmenes en el ámbito del entretenimiento.

#### 3.1 OpenAI Codex

OpenAI, empresa fundada en 2015, es la responsable del desarrollo de Codex, un modelo de inteligencia artificial diseñado para el autocompletado y generación de código. Además de Codex, OpenAI ha desarrollado otras soluciones avanzadas en el ámbito de la IA, como **GPT-3** y **GPT-4**, ampliamente utilizados en generación de texto, así como **DALL·E**, un modelo especializado en la generación de imágenes a partir de descripciones textuales (*OpenAI Codex / OpenAI, n.d.*).

Codex está dirigido principalmente a **desarrolladores y programadores** que buscan asistencia en la escritura de código mediante autocompletado inteligente. Su capacidad para interpretar instrucciones en **lenguaje natural** y convertirlas en código en distintos lenguajes de programación lo convierte en una herramienta útil tanto para profesionales experimentados como para quienes están en proceso de aprendizaje. Gracias a su integración con entornos de desarrollo, Codex optimiza el flujo de trabajo y reduce el tiempo necesario para escribir y depurar código (*OpenAI Codex / OpenAI, n.d.*).

El modelo de negocio de OpenAI para Codex se basó en el acceso a través de una **API privada**, lo que permitía a desarrolladores y empresas integrarlo en sus propias aplicaciones y servicios. En su lanzamiento, Codex estuvo disponible en una fase beta privada y gratuita, con planes de

Asignatura	Datos del alumno	Fecha
Procesamiento del Lenguaje Natural	Apellidos: Pinillos Rubio	07 de febrero de 2025
	Nombre: Jose Manuel	

escalado gradual para garantizar su estabilidad y seguridad en aplicaciones comerciales (*OpenAI Codex / OpenAI*, n.d.).

Desde el punto de vista tecnológico, Codex es una extensión de **GPT-3**, entrenado con un gran volumen de datos que incluyen tanto lenguaje natural como millones de líneas de código fuente provenientes de repositorios públicos, como los alojados en **GitHub**. Si bien su mayor precisión se encuentra en el lenguaje **Python**, también es compatible con otros lenguajes de programación como **JavaScript, Go, Perl, PHP, Ruby, Swift y TypeScript**, lo que lo convierte en una herramienta versátil para distintas necesidades de desarrollo (*OpenAI Codex / OpenAI*, n.d.).

Es importante señalar que, a partir de **marzo de 2023**, OpenAI anunció la **descontinuación del soporte para la API de Codex**, recomendando a los usuarios la migración hacia modelos más recientes, como **GPT-3.5-Turbo**, que ofrecen una mayor eficiencia y capacidades mejoradas para la asistencia en codificación. A pesar de su discontinuación, Codex representó un avance significativo en el uso de *PLN* para el autocompletado de código y sentó las bases para futuras implementaciones en este campo (*OpenAI Codex / OpenAI*, n.d.).

### 3.2 Amazon Prime Video - X-Ray Recaps

Amazon, fundada en **1994**, ha desarrollado **X-Ray Recaps**, una función innovadora dentro de su plataforma de streaming **Amazon Prime Video**, que emplea inteligencia artificial generativa para **crear resúmenes automáticos** de programas de televisión, facilitando a los espectadores la comprensión y el seguimiento de la trama. Esta herramienta es particularmente útil para los usuarios que retoman una serie después de un tiempo, permitiéndoles recordar los eventos clave sin necesidad de buscar resúmenes externos o volver a ver episodios anteriores (*Prime Video's X-Ray Recaps Feature Keeps You Up to Speed on Your Favorite Shows*, n.d.).

Dirigido a **suscriptores de Amazon Prime Video**, este sistema ofrece resúmenes generados automáticamente para temporadas completas, episodios individuales e incluso momentos específicos dentro de los capítulos. Gracias a esta tecnología, los usuarios pueden acceder a **resúmenes contextuales** que destacan información relevante sobre personajes, ubicaciones y

Asignatura	Datos del alumno	Fecha
Procesamiento del Lenguaje Natural	Apellidos: Pinillos Rubio	07 de febrero de 2025
	Nombre: Jose Manuel	

eventos importantes, mejorando su experiencia de visualización (*Prime Video's X-Ray Recaps Feature Keeps You Up to Speed on Your Favorite Shows*, n.d.).

Desde el punto de vista del **modelo de negocio**, **X-Ray Recaps** se integra como una característica adicional dentro de la suscripción estándar de Amazon Prime Video, sin costo extra para los usuarios. Al añadir valor a su plataforma de streaming, Amazon refuerza su competitividad frente a otros servicios del sector, diferenciándose mediante el uso de **inteligencia artificial aplicada al entretenimiento** (*Prime Video's X-Ray Recaps Feature Keeps You Up to Speed on Your Favorite Shows*, n.d.).

La tecnología que hace posible **X-Ray Recaps** se basa en modelos de **IA generativa proporcionados por Amazon Bedrock**, un servicio gestionado de **AWS** diseñado para crear y escalar aplicaciones de inteligencia artificial. Además, combina el uso de modelos de **Amazon Bedrock con modelos personalizados entrenados en Amazon SageMaker**, lo que le permite analizar segmentos de video en conjunto con subtítulos y diálogos para generar descripciones precisas y relevantes. Para garantizar la calidad del contenido, Amazon ha implementado **medidas de seguridad** que aseguran que los resúmenes sean concisos y no contengan *spoilers*, protegiendo la experiencia del espectador (*Prime Video's X-Ray Recaps Feature Keeps You Up to Speed on Your Favorite Shows*, n.d.).

Esta integración de tecnologías avanzadas de **PLN y aprendizaje automático** permite que **X-Ray Recaps** ofrezca resúmenes altamente precisos y adaptados al contexto de cada programa, optimizando la experiencia de los usuarios de **Amazon Prime Video** y marcando un avance significativo en la aplicación de la inteligencia artificial dentro del entretenimiento digital (*Prime Video's X-Ray Recaps Feature Keeps You Up to Speed on Your Favorite Shows*, n.d.).

Asignatura	Datos del alumno	Fecha
Procesamiento del Lenguaje Natural	Apellidos: Pinillos Rubio	07 de febrero de 2025
	Nombre: Jose Manuel	

## 4. REFLEXIÓN SOBRE LA EVOLUCIÓN TECNOLÓGICA Y SU IMPACTO EN EL MERCADO

La generación automática de resúmenes ha evolucionado significativamente en los últimos años, impulsada por mejoras en las tecnologías y metodologías utilizadas. Los primeros enfoques se basaban en métodos **extractivos**, como *TextRank*, que seleccionaban fragmentos clave del texto original sin modificar su estructura. Sin embargo, la llegada del **aprendizaje profundo** ha permitido la transición hacia modelos **abstractivos**, capaces de reformular la información y generar resúmenes más naturales y coherentes. Este avance ha sido posible gracias al desarrollo de modelos *transformer*, como **BART**, **PEGASUS** y **T5**, que han demostrado una mayor capacidad para sintetizar información de manera precisa. Además, la disponibilidad de **grandes corpus de datos estandarizados**, como *CNN/DailyMail* y *XSum*, ha mejorado la capacidad de generalización de estos modelos, permitiendo evaluaciones más precisas y comparaciones sistemáticas.

Desde el punto de vista de los resultados obtenidos, el uso de métricas como **ROUGE**, **BLEU** y **METEOR** ha permitido evidenciar mejoras significativas en la coherencia y calidad de los resúmenes generados por los modelos más recientes. No obstante, uno de los principales desafíos continúa siendo la **fidelidad al contenido original**, ya que algunos modelos pueden introducir información inexistente en el texto fuente, comprometiendo la precisión y confiabilidad del resumen. A pesar de esto, la evolución tecnológica ha permitido superar muchas de las limitaciones que presentaban los enfoques tradicionales, acercándose cada vez más a la calidad de resúmenes elaborados por humanos.

La conexión entre la investigación en generación automática de resúmenes y su implementación en productos comerciales es evidente. Los avances en **PLN** han permitido que las soluciones comerciales incorporen modelos más avanzados, dejando atrás los métodos basados en reglas y estadísticas para adoptar enfoques basados en **deep learning**. Ejemplo de ello es **X-Ray Recaps de Amazon Prime Video**, que emplea modelos entrenados en **Amazon Bedrock** y **SageMaker** para generar resúmenes de contenido audiovisual. De manera similar, **OpenAI Codex**, aunque centrado en el autocompletado de código, demuestra cómo los modelos basados en *transformers*, como **GPT-3**, pueden aplicarse en la generación contextualizada de texto.

Asignatura	Datos del alumno	Fecha
Procesamiento del Lenguaje Natural	Apellidos: Pinillos Rubio	07 de febrero de 2025
	Nombre: Jose Manuel	

Sin embargo, la implementación comercial de estos modelos aún enfrenta **desafíos en términos de eficiencia computacional y escalabilidad**. Mientras que en la investigación se busca mejorar la precisión y fidelidad de los resúmenes, en el ámbito comercial se prioriza la experiencia del usuario y la integración de los modelos en servicios funcionales y accesibles. Esto puede implicar ciertos ajustes en los modelos utilizados, sacrificando en algunos casos fidelidad a cambio de mayor fluidez o velocidad de procesamiento. A pesar de estas diferencias, la evolución de las soluciones comerciales sigue de cerca las investigaciones en PLN, y es previsible que los futuros avances continúen mejorando la precisión y adaptabilidad de los modelos a diferentes necesidades del mercado.

A pesar de los avances logrados, la generación automática de resúmenes sigue enfrentando **retos pendientes** que dificultan su aplicación óptima en diversos contextos. Además del problema de la **fidelidad al contenido**, otro desafío importante es el **control de sesgos en los modelos de PLN**. Dado que estos sistemas aprenden a partir de grandes volúmenes de datos, la presencia de sesgos implícitos en los textos de entrenamiento puede influir en la manera en que se sintetiza la información. Para mitigar este problema, es crucial mejorar los procesos de curación de datos y desarrollar enfoques que permitan una mayor transparencia en la generación de resúmenes.

Otro aspecto a considerar es la **optimización del rendimiento computacional**. Los modelos más avanzados requieren un uso intensivo de recursos, lo que dificulta su escalabilidad e integración en plataformas de uso masivo. Si bien se han desarrollado técnicas para reducir costos computacionales y mejorar la eficiencia de los modelos, sigue siendo necesario encontrar un equilibrio entre calidad y viabilidad operativa. Asimismo, aunque métricas como **ROUGE, BLEU y METEOR** han sido útiles para evaluar el desempeño de los modelos, estas no siempre reflejan con precisión la percepción humana de la calidad del resumen. Por ello, el desarrollo de **nuevas métricas de evaluación** sigue siendo un desafío clave en la investigación.

Podemos observar, como los avances en la generación automática de resúmenes han permitido el desarrollo de modelos más sofisticados y precisos, acercándose a la calidad de los resúmenes humanos. Sin embargo, aún persisten desafíos que deben ser abordados para lograr resúmenes más fieles al contenido original, menos influenciados por sesgos y más eficientes desde el punto de vista computacional.

Asignatura	Datos del alumno	Fecha
Procesamiento del Lenguaje Natural	Apellidos: Pinillos Rubio	07 de febrero de 2025
	Nombre: Jose Manuel	

## BIBLIOGRAFÍA

- Cajueiro, D. O., Nery, A. G., Tavares, I., De Melo, M. K., Reis, S. A. dos, Weigang, L., & Celestino, V. R. R. (2023). *A comprehensive review of automatic text summarization techniques: method, data, evaluation and coding*. <https://arxiv.org/abs/2301.03403v4>
- Carmona, M. Á. Á., Aranda, R., Diaz-Pacheco, Á., & Ceballos-Mejía, J. de J. (2022). *Generador automático de resúmenes científicos en investigación turística*. <https://doi.org/10.1590/SCIELOPREPRINTS.4194>
- Nikolaevna, Y., René, L., & García Hernández, A. (n.d.). *Generación automática de resúmenes Automatic Generation of Text Summaries Challenges, proposals and experiments*.
- OpenAI Codex | OpenAI. (n.d.). Retrieved February 12, 2025, from <https://openai.com/index/openai-codex/>
- Prime Video's X-Ray Recaps Feature Keeps You Up to Speed on Your Favorite Shows. (n.d.). Retrieved February 12, 2025, from <https://www.aboutamazon.com/news/entertainment/amazon-prime-video-x-ray-recaps>