-Editorial-
CILADI
Centro de Investigación Latinoamericano
para el Desarrollo e Innovación

HERRAMIENTAS DE INTELIGENCIA ARTIFICIAL: INVESTIGACIÓN, ARQUITECTURA Y DISEÑO

RICHARD ESTEBAN SARZOSA SOTO

2024

Copyright © 2024

Sarzosa Soto Richard Esteban

Todos los derechos reservados.

ISBN: 978-9942-7292-6-2

La presente obra fue revisada por pares académicos ciegos conforme al proceso editorial de la Editorial CILADI.

Índice de contenidos

HERRAMIENTAS DE INTELIGENCIA ARTIFICIAL: INVESTIGACIÓN, ARQUITECTURA Y DISEÑO 1

1. Introducción 1

2. Inteligencia artificial (IA): Modelados predictivos y generativos 3

2.1. Conceptos básicos para comprender la IA 3

2.1.1. Machine learning (ML) 3

2.1.2. Deep learning (DL) 4

2.1.3. Prompting 6

2.2. Modelos generativos de lenguaje 7

2.2.1. ChatGPT 8

2.2.2. Gemini 12

2.2.3. Perplexity AI 16

2.2.4. Copilot 19

2.3. Lectura de documentos extensos 24

2.3.1. NotebookLM 24

2.4. Revisión y generación de contenido científico 27

2.4.1. Scispace 27

2.4.2. Connected Papers 30

2.4.3. Semantic Scholar 33

2.4.4. ResearchRabbit 37

2.4.5. Consensus 40

2.4.6. Elicit 44

2.4.7. Litmaps 46

2.4.8. Iris 50

2.5. Modelos generativos de imagen 55

2.5.1. Modelos Generales 55

2.5.2. Arquitectura y diseño 69

3. Conclusiones, reflexiones y recomendaciones 77

3.1. Conclusiones 77

3.2. Reflexiones 78

3.3. Recomendaciones 78

Referencias bibliográficas 81

Índice de figuras

Figura 1. Esquema funcional del Machine Learning 4
Figura 2. Esquema funcional del Deep Learning 5
Figura 3. Esquema funcional del prompting 7
Figura 4. Esquema funcional de ChatGPT 9
Figura 5. Esquema funcional de Gemini 15
Figura 6. Esquema funcional de Copilot 22
Figura 7. Esquema funcional de NotebookLM 27
Figura 8. Esquema funcional de Scispace 29
Figura 9. Esquema funcional de Connected Papers 33
Figura 10. Esquema funcional de Semantic Scholar 36
Figura 11. Esquema funcional de ResearchRabbit 40
Figura 12. Esquema funcional de Consensus 43
Figura 13. Esquema funcional de Elicit 46
Figura 14. Esquema funcional de Litmaps 49
Figura 15. Esquema funcional de Iris 52
Figura 16. Esquema funcional de Midjourney 57
Figura 17. Esquema funcional de Krea 61
Figura 18. Esquema funcional de Leonardo 64
Figura 19. Esquema funcional de Stable Diffusion 67
Figura 20. Esquema funcional de D5-Hi 72
Figura 21. Esquema funcional de PromeAI 75
Figura 22. Esquema funcional de Lookx.ai 76

Índice de tablas

Tabla 1. Cuadro comparativo con los principales aspectos de las herramientas: ChatGPT, Gemini, Perplexity AI, y Copilot 23

Tabla 2. Cuadro comparativo con los principales aspectos de las herramientas: Scispace, Connected Papers, Semantic Scholar, ResearchRabbit, Consensus, Elicit, Litmaps, e Iris..... 53

Tabla 3. Cuadro comparativo con los principales aspectos de las herramientas: Midjourney, Krea, Leonardo, Stable Diffusion 68

Listado de Acrónimos

BERT	Bidirectional encoder representations from transformer (Representación de codificador bidireccional de transformadores)
DL	Deep Learning (Aprendizaje profundo)
DOI	Digital Object Identifier (Identificador de Objeto Digital)
GAN	Generative adversarial network (Red generativa antagónica)
GPT	Generative pre-trained transformer (Transformador generativo preentrenado)
IA	Inteligencia artificial
IAG	Inteligencia artificial generativa
MGL	Modelo generativo de lenguaje
ML	Machine learning (Aprendizaje automático)
PLN	Procesamiento de lenguaje natural
RNC	Redes neuronales convolucionales
RNR	Redes neuronales recurrentes

HERRAMIENTAS DE INTELIGENCIA ARTIFICIAL: INVESTIGACIÓN, ARQUITECTURA Y DISEÑO

1. Introducción

En el vertiginoso desarrollo tecnológico del siglo XXI, la inteligencia artificial (IA) ha emergido como una disciplina que ha redefinido múltiples campos del conocimiento humano. Su impacto se ha extendido a áreas como la investigación científica, el diseño, la arquitectura y otros dominios creativos y analíticos, posicionándose como una herramienta indispensable para abordar desafíos complejos. Este libro propone explorar las capacidades, aplicaciones y limitaciones de diversas herramientas de inteligencia artificial, con un enfoque particular en aquellas diseñadas para procesos de investigación, diseño arquitectónico y generación creativa.

El texto se estructura en torno a un análisis riguroso de herramientas avanzadas de IA, clasificadas en dos grandes vertientes: los modelos generativos de lenguaje y los modelos de generación de imágenes. Estas tecnologías, basadas en arquitecturas de aprendizaje profundo, han revolucionado la interacción humano-máquina, permitiendo el desarrollo de contenido multimodal que trasciende las barreras del lenguaje y las representaciones visuales tradicionales. En este sentido, herramientas como *ChatGPT, Gemini* y *Copilot* son estudiadas en detalle, destacando su capacidad para generar texto coherente y contextualizado en aplicaciones tan diversas como la educación, la investigación y la productividad empresarial.

La arquitectura del aprendizaje profundo, representada por transformadores y redes neuronales convolucionales, constituye la base tecnológica de estas herramientas, habilitando modelos capaces de comprender y procesar información textual y visual con una precisión sin precedentes. Este libro no solo describe los principios técnicos que sustentan dichas tecnologías, sino que también analiza su implementación práctica en tareas específicas, como la síntesis de literatura científica donde destacan herramientas como *Scispace* o *Connected Papers*, y la exploración de relaciones conceptuales entre publicaciones académicas, con herramientas como *Semantic Scholar* o *ResearchRabbit*.

En el ámbito del diseño, la inteligencia artificial ha permitido a los profesionales no solo automatizar procesos creativos, sino también explorar nuevos horizontes de innovación. Herramientas como *Midjourney, Stable Diffusion* y *Krea* son claros ejemplos de cómo los modelos generativos de imágenes han abierto oportunidades para la experimentación visual, integrando técnicas tradicionales con enfoques modernos en diseño gráfico, arquitectura y producción multimedia. Este libro examina detalladamente estas tecnologías, identificando sus fortalezas, aplicaciones específicas y desafíos éticos.

El propósito de esta obra no se limita a describir el estado del arte de las tecnologías de IA. También busca fomentar una reflexión crítica sobre las implicaciones éticas y sociales de su implementación. Aspectos como la privacidad, la transparencia en los modelos de entrenamiento y la mitigación de sesgos son discutidos con profundidad, proporcionando a los lectores una comprensión integral de los riesgos y beneficios asociados con el uso de estas herramientas. Finalmente, este libro está diseñado para ser una guía práctica y teórica para investigadores, diseñadores, arquitectos y estudiantes interesados en integrar la inteligencia artificial en sus disciplinas. A través de un enfoque analítico y estructurado, se presenta no solo un panorama de las herramientas disponibles, sino también ejemplos concretos de su uso en contextos reales. Con esto, se espera contribuir al desarrollo de competencias que permitan a los profesionales adaptarse a un entorno en constante cambio, impulsado por la evolución de la inteligencia artificial en la actualidad.

2. Inteligencia artificial (IA): Modelados predictivos y generativos

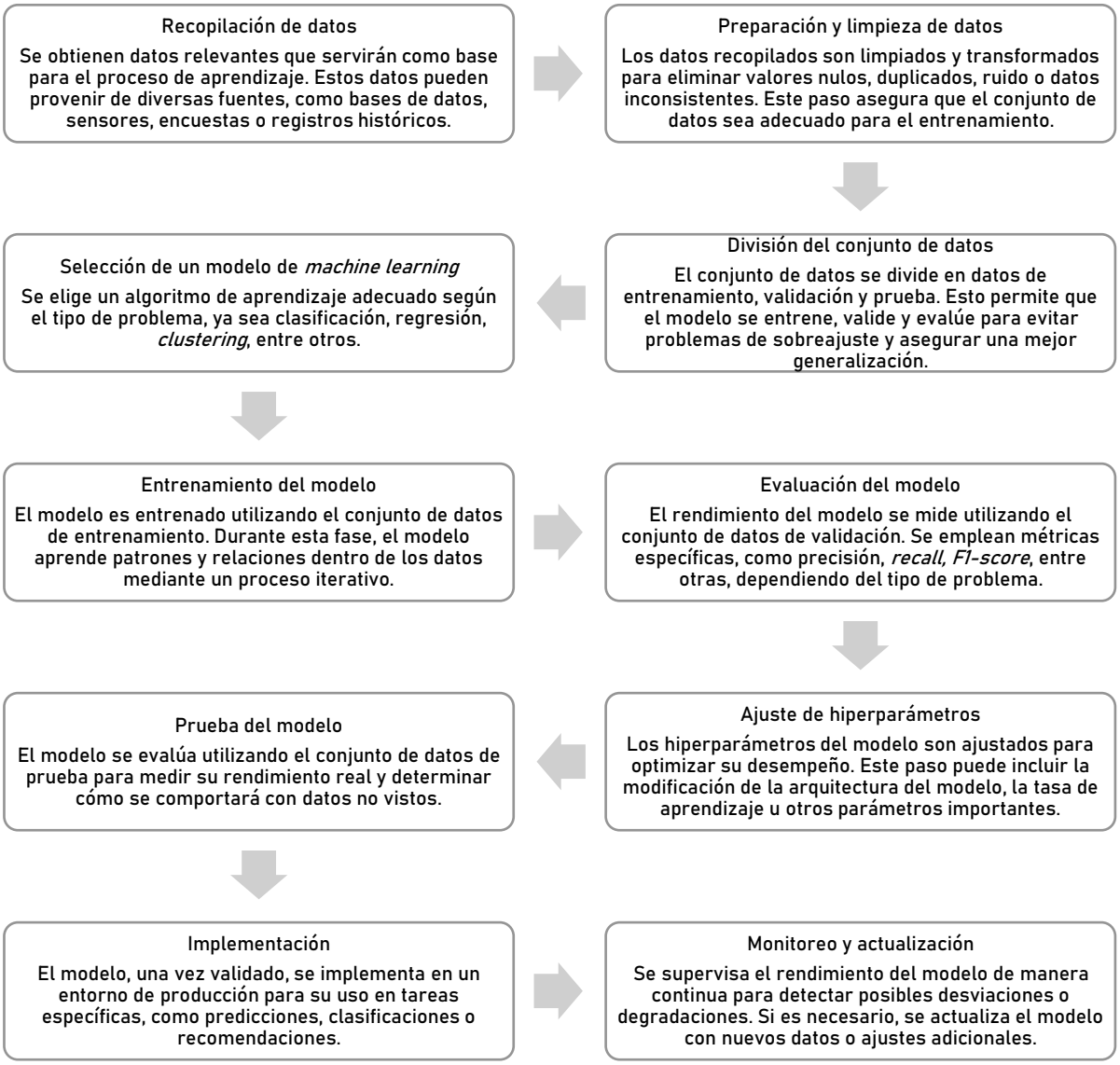
2.1. Conceptos básicos para comprender la IA

2.1.1. Machine learning (ML)

El *Machine Learning* (ML), traducido al español como "aprendizaje automático", es una subdisciplina de la inteligencia artificial (IA) que se enfoca en desarrollar algoritmos capaces de aprender a partir de datos. A diferencia de la programación convencional, en la cual el programador instruye a la máquina cómo actuar ante diferentes situaciones, el ML permite que la máquina descubra patrones y tome decisiones por sí misma sin haber sido explícitamente programada para cada tarea. Según Alpaydin (2020), el *machine learning* se fundamenta en la aplicación de modelos estadísticos para el análisis de datos, con el objetivo de extraer información que permita realizar predicciones o clasificaciones precisas. Esto lo diferencia de enfoques tradicionales de IA, como los sistemas basados en reglas, que dependen de un conocimiento explícito y predefinido. Cabe acotar que el campo de aplicaciones del ML es vasto, desde sistemas de recomendación, como los usados en plataformas de *streaming*, hasta el reconocimiento de voz y visión por ordenadores. Bajo ese contexto, los tipos de ML se dividen en tres categorías principales:

- **Aprendizaje supervisado:** Aquí, el modelo aprende a partir de un conjunto de datos etiquetados, donde se conoce el resultado o etiqueta de cada ejemplo. Este tipo de ML es usado comúnmente en la predicción y clasificación, como en la detección de fraudes bancarios o en el diagnóstico médico automatizado (Goodfellow et al., 2016).
- **Aprendizaje no supervisado:** En este caso, el modelo trabaja con datos no etiquetados y debe descubrir patrones inherentes a los datos. Un uso típico es la agrupación de clientes en estrategias de marketing (Murphy, 2012).
- **Aprendizaje por refuerzo:** Este enfoque implica que el modelo aprende a través de la interacción con su entorno, mejorando sus acciones con base en recompensas o castigos. Se usa en juegos, robótica y vehículos autónomos (Sutton & Barto, 2018).

Figura 1. Esquema funcional del *Machine Learning*



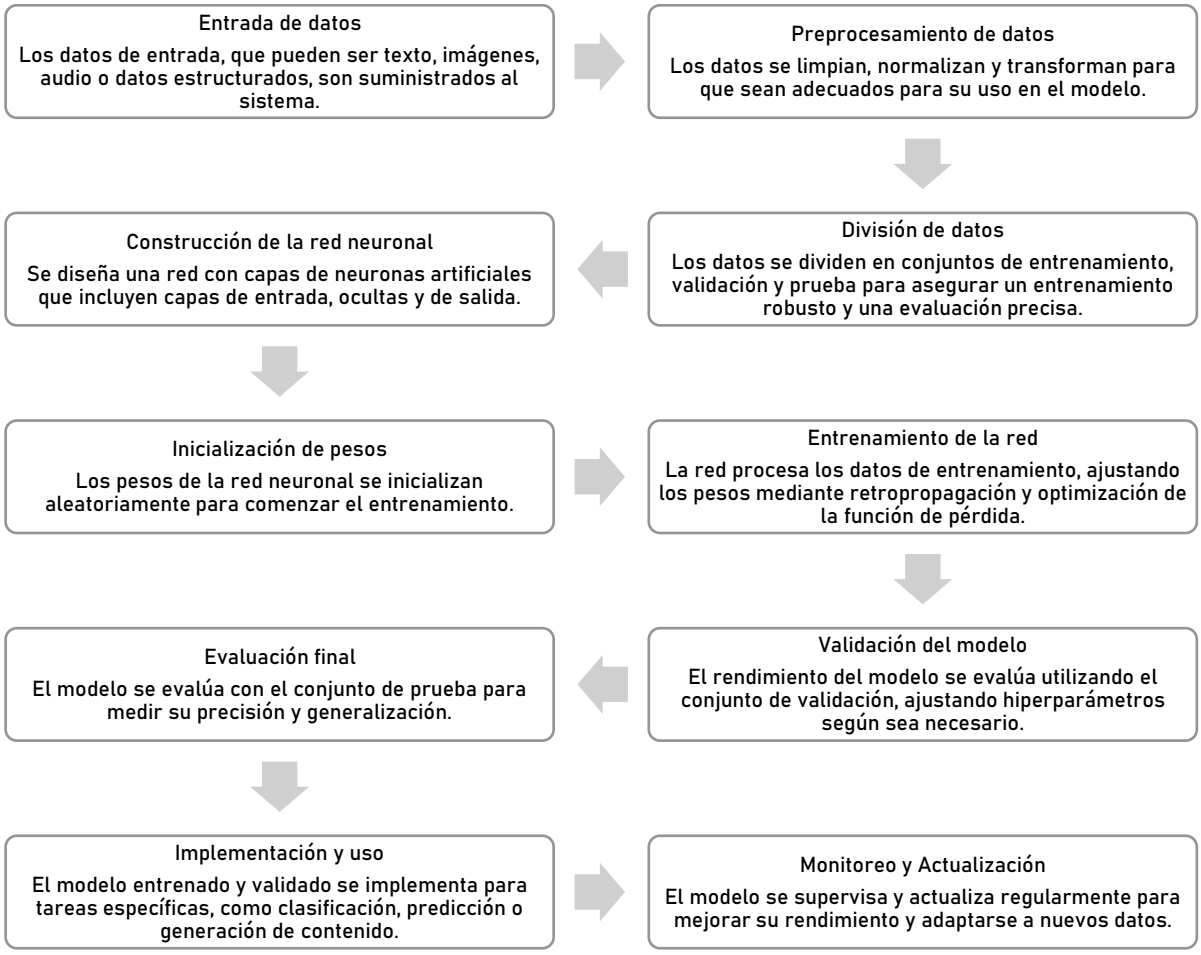
Fuente: Elaborado por el autor.

2.1.2. Deep learning (DL)

El *Deep Learning* (DL), conocido en español como "aprendizaje profundo", es una rama del *Machine Learning* (ML) que se basa en redes neuronales artificiales con múltiples capas, conocidas como redes neuronales profundas. A diferencia del ML tradicional, que utiliza modelos más simples y depende de la extracción manual de características, el DL automatiza la extracción de características relevantes a partir de los datos en bruto. Como indica Schmidhuber (2015), el aprendizaje profundo evidencia una capacidad sobresaliente para abordar problemas de elevada

complejidad, gracias a su estructura jerárquica, la cual le permite aprender representaciones de los datos en múltiples niveles de abstracción.

Figura 2. Esquema funcional del Deep Learning



Fuente: Elaborado por el autor.

Una de las principales diferencias entre ML y DL es el manejo de grandes volúmenes de datos. Mientras que los algoritmos del ML suelen depender de conjuntos de datos más pequeños y necesitan de la intervención humana en la preprocesación de los datos, el DL puede analizar datos masivos y no estructurados, como imágenes y texto, de forma más eficiente gracias a sus múltiples capas de procesamiento no lineal. En ese contexto, las principales arquitecturas de *Deep Learning* son:

- **Redes neuronales convolucionales (RNC):** Se utilizan predominantemente para tareas de procesamiento de imágenes y visión por ordenador. Las RNC están diseñadas para capturar características espaciales mediante capas de convolución y agrupación (LeCun et al., 2015).

- **Redes neuronales recurrentes (RNR):** Ideales para datos secuenciales, como el procesamiento del lenguaje natural o series temporales. Las RNR retienen información de entradas anteriores mediante bucles dentro de su estructura (Mikolov et al., 2013).
- **Transformers:** Introducidos en el contexto del procesamiento del lenguaje natural, los *transformers* se han convertido en la arquitectura dominante en modelos de lenguaje como GPT, debido a su capacidad para procesar datos de manera eficiente y paralela (Vaswani et al., 2017).

2.1.3. Prompting

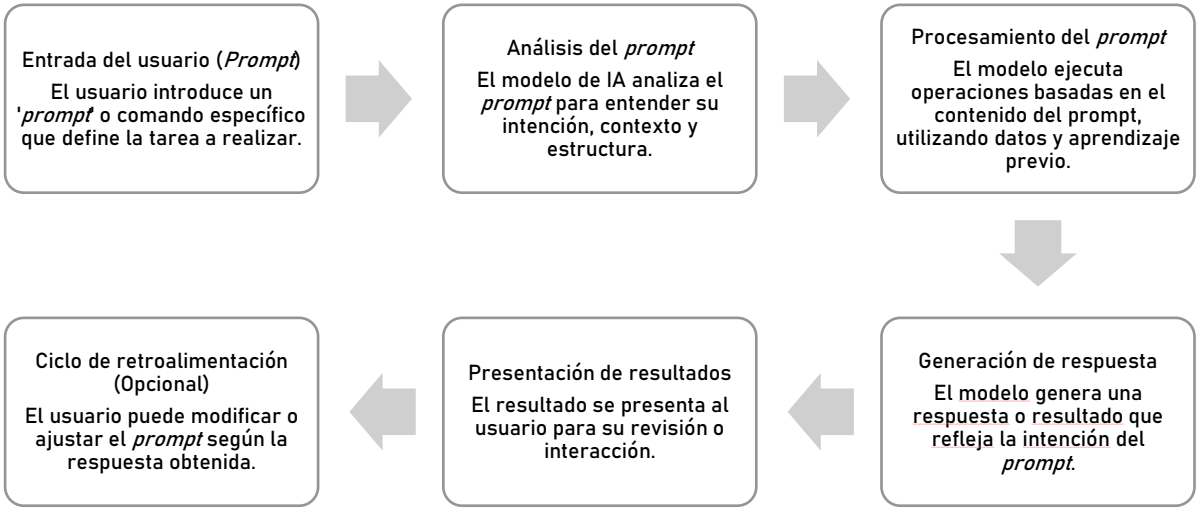
El término "*prompt*" se refiere a las instrucciones o indicaciones que se le proporcionan a un modelo de lenguaje o una IA generativa para que produzca un resultado específico. Para Brown et al. (2020), un *prompt* puede consistir en una frase inicial simple o en un conjunto más detallado de instrucciones que orientan la salida del modelo. La precisión y claridad del aviso influyen de manera directa en la calidad de la respuesta generada. En modelos avanzados como GPT (*Generative Pre-trained Transformer* o traducido al español como: Transformador generativo preentrenado), la estructura del indicador determina qué información se selecciona de la base de datos que el modelo ha procesado.

Cabe señalar que los modelos de lenguaje, como GPT, han experimentado mejoras en su capacidad para interpretar y responder a indicaciones, gracias al avance en la arquitectura *transformer* y al entrenamiento a gran escala. El uso de avisos ha evolucionado desde instrucciones simples a sistemas más complejos de interacción, que permiten generar contenido coherente y ajustado a las necesidades específicas del usuario (Radford et al., 2019). Este desarrollo ha permitido aplicaciones en diversos campos, como la escritura automatizada, la asistencia virtual y la generación creativa de contenido.

A pesar de su eficacia, los *prompt* también presentan limitaciones. La ambigüedad o falta de precisión en la indicación puede generar resultados incorrectos o irrelevantes, lo que requiere de un esfuerzo adicional por parte del usuario para ajustar las instrucciones. Además, la IA puede generar respuestas o señales incorrectas si no se diseñan adecuadamente los *prompt*, lo que puede limitar la

confiabilidad en aplicaciones críticas como el periodismo o la ciencia (Bender et al., 2021).

Figura 3. Esquema funcional del *prompting*



Fuente: Elaborado por el autor.

2.2. Modelos generativos de lenguaje

Los modelos generativos de lenguaje representan un avance crucial en el ámbito de la inteligencia artificial y el procesamiento de lenguaje natural, diseñados para generar texto coherente y contextualizado a partir de entradas proporcionadas por los usuarios. Estos modelos, basados en arquitecturas complejas, han transformado la manera en que interactuamos con sistemas de IA, impulsando su uso en aplicaciones como asistentes virtuales, generación de contenido, educación, entre otros ámbitos. Entre los modelos destacados se encuentran *ChatGPT*, desarrollado por OpenAI, y *Gemini*, un modelo de *Google DeepMind* que extiende sus capacidades a un enfoque multimodal para manejar texto, imágenes y potencialmente audio.

ChatGPT, como parte de la familia de modelos *GPT*, utiliza una arquitectura basada en el procesamiento auto-regresivo para generar texto palabra por palabra, considerando el contexto proporcionado por el usuario. Su desarrollo ha sido impulsado por la creciente demanda de herramientas que permitan interacciones naturales y conversaciones complejas. Este modelo se entrena mediante un proceso que incluye una fase de preentrenamiento con grandes cantidades de datos textuales y un ajuste fino que lo adapta a tareas específicas, como el diálogo conversacional y

la generación de texto. Por otro lado, *Gemini* destaca por su capacidad multimodal, integrando datos de diferentes tipos, como texto e imágenes, lo que lo hace especialmente útil para tareas que requieren una comprensión más rica y compleja del contexto. Esta capacidad de manejar múltiples modalidades de datos simultáneamente permite aplicaciones que van desde la generación de contenido multimedia hasta la descripción de imágenes y su integración en contextos más amplios, como la medicina o el entretenimiento.

Los modelos generativos de lenguaje, sin embargo, enfrentan ciertos desafíos técnicos y éticos. Desde el alto costo computacional y la necesidad de datos extensivos para su entrenamiento hasta los riesgos de reproducir sesgos y generar desinformación, su desarrollo requiere un equilibrio constante entre innovación y responsabilidad ética. Las aplicaciones de estos modelos están destinadas a expandirse en múltiples sectores, pero con un enfoque cada vez mayor en la regulación y el control para evitar el mal uso y asegurar una implementación ética. Bajo el contexto presentado, a continuación se desarrolla un breve repaso de las herramientas: *ChatGPT, Gemini, Perplexity AI,* y *Copilot*.

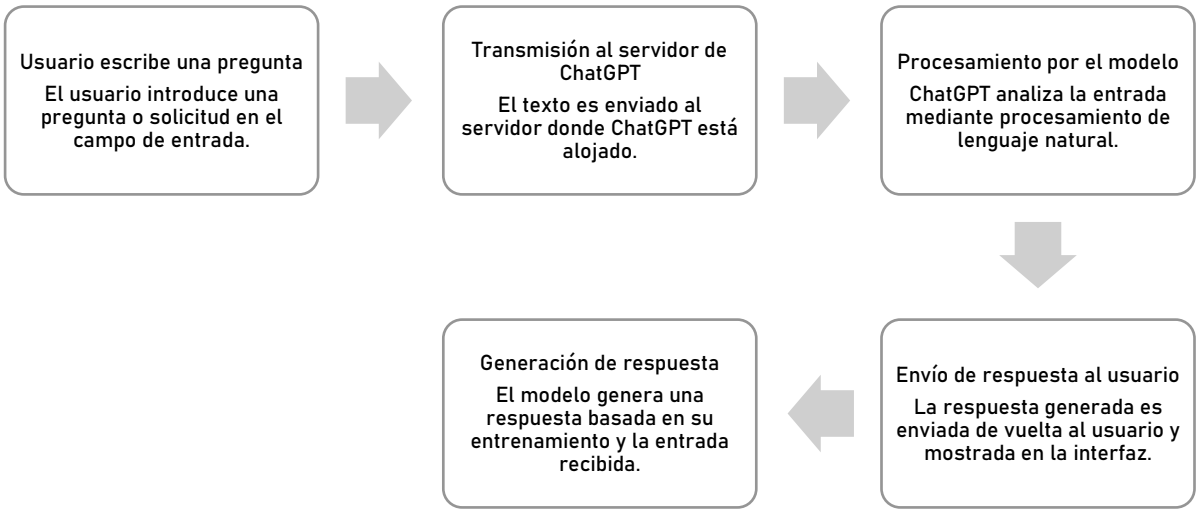
2.2.1. ChatGPT

ChatGPT es un modelo generativo de lenguaje (MGL) basado en la arquitectura *Transformer*, diseñado para generar texto coherente y fluido en conversaciones. Este MGL fue desarrollado por OpenAI y es una implementación de la serie de modelos GPT. Radford et al. (2019) sostiene que el primer modelo GPT fue presentado en 2018, y desde entonces ha evolucionado a través de varias versiones, con mejoras en el tamaño de los parámetros y la capacidad para gestionar conversaciones complejas. Según el autor, la arquitectura de GPT se basa en un proceso de aprendizaje no supervisado, donde el modelo es preentrenado con grandes cantidades de texto y luego ajustado mediante técnicas de aprendizaje supervisado o por refuerzo (Radford et al., 2019). Cabe acotar que el desarrollo de *ChatGPT* fue impulsado por la creciente demanda de herramientas más eficientes para procesar el lenguaje natural, permitiendo una interacción más natural y humana en aplicaciones como asistentes virtuales y *chatbots*.

ChatGPT, como los demás modelos de la familia GPT, se basa en la arquitectura *Transformer* (Vaswani et al., 2017). La principal característica de esta arquitectura es el mecanismo de atención, que permite al modelo centrarse en diferentes partes del texto de entrada a medida que genera una respuesta. La versión de *ChatGPT* que analiza esta obra es un modelo de grandes dimensiones, con miles de millones de parámetros que son ajustados durante el proceso de entrenamiento. Cada "parámetro" es un valor numérico que el modelo ajusta para generar predicciones precisas, basadas en los datos de entrenamiento. Bajo ese contexto, el funcionamiento de *ChatGPT* sigue tres fases fundamentales:

a. **Preentrenamiento:** En esta fase, el modelo se expone a grandes cantidades de texto disponible en la web. Este proceso permite que el modelo aprenda patrones en el lenguaje, lo que le facilita generar texto coherente.
b. **Ajuste fino (*Fine-tuning*):** Aquí, el modelo es entrenado en tareas específicas con ejemplos más detallados, en ocasiones utilizando retroalimentación de evaluadores humanos. Este ajuste es esencial para mejorar su capacidad de seguir instrucciones y producir respuestas más adecuadas en diferentes contextos.
c. **Generación:** Finalmente, cuando el modelo recibe una instrucción, utiliza el contexto proporcionado para generar texto palabra por palabra, aprovechando su conocimiento previo del lenguaje.

Figura 4. Esquema funcional de ChatGPT



Fuente: Elaborado por el autor.

Cabe acotar que *ChatGPT* pertenece a la categoría de modelos de lenguaje generativo, pero tiene características distintivas respecto a otros modelos en este campo. Un punto clave de diferenciación es su capacidad para gestionar interacciones conversacionales prolongadas. Devlin et al. (2019) sostiene que, en comparación con modelos anteriores de lenguaje, como BERT (*Bidirectional Encoder Representations from Transformer* o traducido al español como: Representación de Codificador Bidireccional de Transformadores), el cual está diseñado para tareas de comprensión de texto (no generación); *ChatGPT* está orientado a generar texto, lo que lo hace más adecuado para aplicaciones interactivas.

Además, modelos como BERT emplean un enfoque de codificación bidireccional, lo que significa que consideran el contexto en ambas direcciones (antes y después de una palabra). En cambio, GPT utiliza un enfoque unidireccional, donde predice la próxima palabra obtenida únicamente en el contexto anterior. Esta diferencia en la estructura de atención que permite GPT, y en particular *ChatGPT*, permite que sea más eficiente en tareas de generación de texto coherente. Por otro lado, es necesario señalar que las aplicaciones de *ChatGPT* son diversas y abarcan múltiples facetas, como:

- **Asistentes virtuales:** Uno de los usos más extendidos es en el desarrollo de asistentes virtuales, capaces de mantener conversaciones fluidas con los usuarios, proporcionando respuestas precisas en contextos como el servicio al cliente, soporte técnico o *marketing*.
- **Generación de contenido:** *ChatGPT* se suele utilizar para generar borradores de artículos, publicaciones en redes sociales, descripciones de productos y otros contenidos textuales. Esto optimiza los procesos creativos y productivos en el ámbito del marketing digital y la producción editorial.
- **Educación:** En el campo educativo, *ChatGPT* se suele utilizar para responder preguntas de estudiantes, explicar conceptos y ofrecer tutorías personalizadas.
- **Investigación científica:** También se emplea para resumir artículos científicos o procesar grandes volúmenes de texto, facilitando el análisis y la extracción de información clave (Brown et al., 2020).

A pesar de sus avances, *ChatGPT* enfrenta una serie de limitaciones técnicas y éticas. En cuanto a las limitaciones técnicas, el modelo puede generar contenido que, aunque gramaticalmente correcto, sea semánticamente erróneo o incoherente. También puede reproducir sesgos presentes en los datos de entrenamiento, lo que puede llevar a resultados no deseados o socialmente problemáticos (Bender et al., 2021). En complemento, desde una perspectiva ética, existe preocupación por el uso de *ChatGPT* en la generación de desinformación, *spam* o contenido engañoso. Además, el hecho de que el modelo aprenda de vastos conjuntos de datos no siempre controlados plantea interrogantes sobre la privacidad y la gestión de los datos utilizados durante su entrenamiento (Tokayev, 2023).

En relación a su entrenamiento, *ChatGPT* implica el uso de grandes cantidades de datos textuales recopilados de diversas fuentes, incluidos libros, sitios web y foros. Durante el proceso de preentrenamiento, el modelo aprende a predecir la siguiente palabra en una secuencia de texto, utilizando la técnica de aprendizaje auto-regresivo. Para Brown et al. (2020), este proceso requiere el uso de vastos recursos computacionales y suele realizarse en infraestructuras especializadas, como superordenadores y sistemas de procesamiento en paralelo. El modelo también pasa por una fase de ajuste fino, donde se optimizan sus respuestas a preguntas específicas utilizando conjuntos de datos etiquetados y retroalimentación humana. Este ajuste le permite a *ChatGPT* ser más preciso y adaptarse mejor a los contextos conversacionales.

Cabe acotar que el futuro del desarrollo de modelos generativos de lenguaje como *ChatGPT* se enfrenta a varios desafíos. Uno de los principales radica en mejorar su capacidad para generar respuestas más precisas y menos sesgadas. También es necesario abordar la eficiencia energética y los altos costos computacionales asociados con el entrenamiento de modelos tan grandes (Strubell et al., 2019). Otro desafío importante es el desarrollo de herramientas y sistemas para la detección de contenido generado por IA, lo que podría permitir mitigar el riesgo de desinformación y garantizar que se mantenga la transparencia en su uso. En el futuro, se espera que los modelos de lenguaje generativo se integren con mayor profundidad en áreas como la medicina, el derecho y la educación, donde actualmente se exploran sus

aplicaciones, pero con mayores controles y regulaciones éticas (Floridi & Chiriatti, 2020).

2.2.2. Gemini

Gemini es un modelo generativo de lenguaje desarrollado por *Google DeepMind*, conocido por su capacidad multimodal, lo que significa que puede manejar y generar múltiples tipos de datos, como texto, imágenes, y, en futuras versiones, también audio. Este modelo fue diseñado para superar algunas de las limitaciones de los modelos anteriores, como GPT-3 y GPT-4, ofreciendo una mayor integración entre diferentes modalidades de información. Según sus desarrolladores, el objetivo de *Gemini* radica mejorar la interacción humano-máquina, permitiendo que el modelo comprenda y genere respuestas más contextuales y adaptadas a situaciones más complejas (DeepMind, 2024). El desarrollo de *Gemini* se enmarca en el esfuerzo de *Google DeepMind* por avanzar en una IA que no solo sea capaz de procesar y generar texto, sino también de integrar datos visuales y auditivos para una mayor comprensión de la realidad. Esta característica multimodal lo distingue de otros modelos generativos, posicionándolo como una herramienta clave en tareas que requieren la integración de diferentes tipos de datos (DeepMind, 2024).

La arquitectura de *Gemini* se basa en una adaptación del *Transformer*, la misma tecnología que sustenta otros modelos de lenguaje de gran escala, como GPT. Sin embargo, lo que distingue a *Gemini* es su capacidad multimodal, que le permite procesar múltiples formas de datos simultáneamente. Esta estructura le permite manejar tanto información visual como textual en un único marco de procesamiento, lo cual no era posible con modelos como GPT-3, los cuales se enfocan exclusivamente en el procesamiento de texto. En términos de funcionamiento, *Gemini* emplea un sistema jerárquico de atención, similar al *Transformer* estándar, pero con la capacidad añadida de atención cruzada entre diferentes modalidades de datos. Esto le permite generar respuestas que no solo sean coherentes en términos de lenguaje, sino también visualmente y, eventualmente, auditivamente coherentes. Esta capacidad es esencial para aplicaciones que requieren la integración de información compleja, como la descripción de imágenes o el desarrollo de contenido multimedia (Vaswani et al., 2017; DeepMind, 2024).

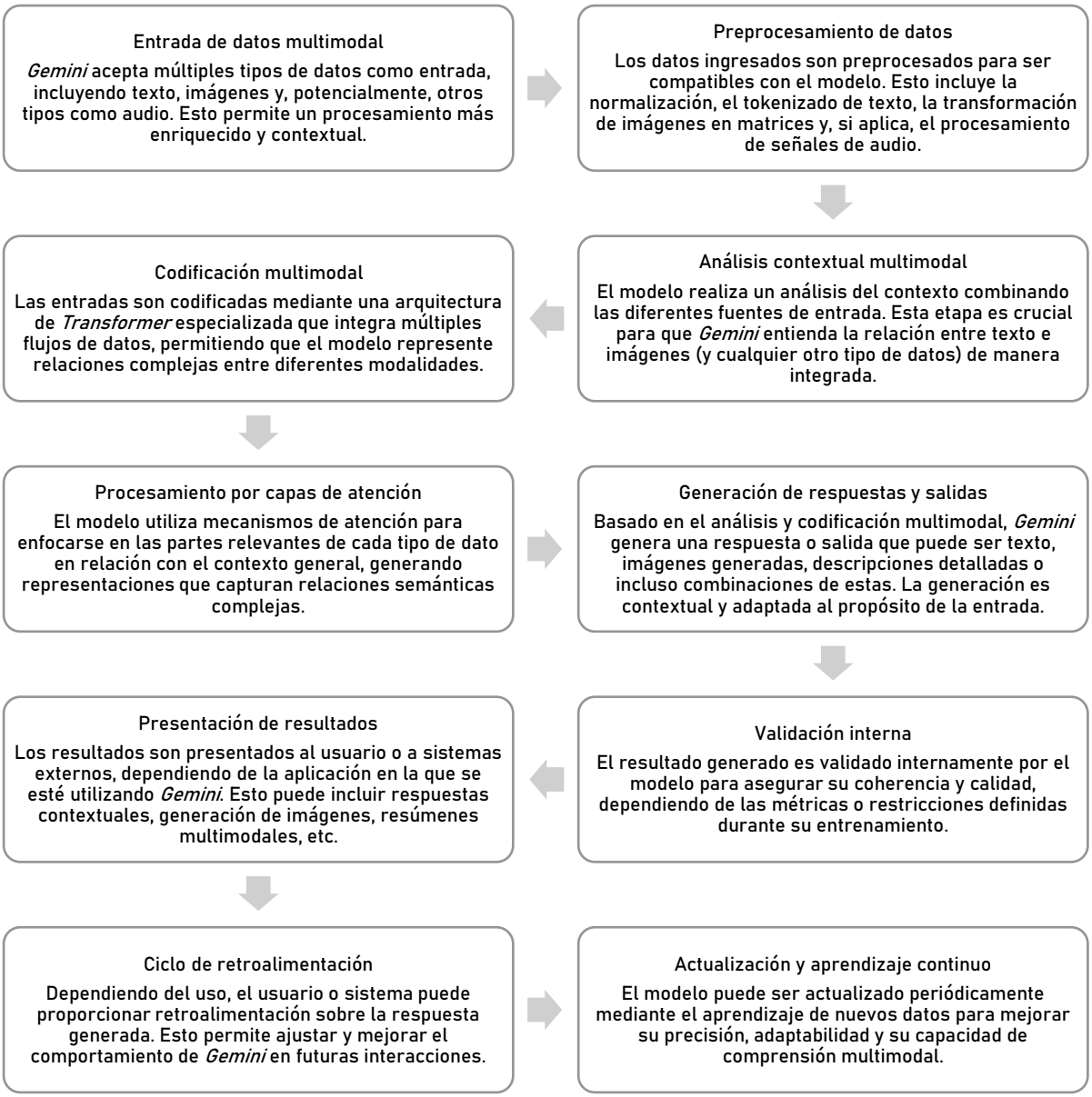
La diferencia clave entre *Gemini* y otros modelos generativos como GPT radica en su enfoque multimodal. Mientras que GPT-3 y GPT-4 están diseñados exclusivamente para la generación de texto, *Gemini* puede manejar múltiples tipos de entrada y generar respuestas que integran texto e imágenes, y potencialmente audio. Esto lo convierte en una herramienta más versátil para tareas que requieren una comprensión más rica del contexto, como la generación automática de subtítulos para videos o el desarrollo de descripciones detalladas de imágenes (Brown et al., 2020; DeepMind, 2024). Además, *Gemini* ha sido diseñado con un enfoque más robusto en la interacción contextual. A diferencia de GPT, que genera texto en función de una secuencia lineal de palabras, *Gemini* puede analizar y generar respuestas basadas en múltiples fuentes de información al mismo tiempo. Esta capacidad lo hace más adecuado para aplicaciones en las que se requiere una combinación de información textual y visual para generar contenido útil (DeepMind, 2024).

En relación a las aplicaciones de *Gemini*, se puede aseverar que son vastas y abarcan varios sectores. En el campo de la salud, *Gemini* puede utilizarse para analizar imágenes médicas y generar descripciones automáticas de hallazgos clínicos. Esta combinación de capacidades visuales y textuales es particularmente útil en la radiología y otras especialidades donde las imágenes médicas desempeñan un papel clave en el diagnóstico. En el marketing digital, *Gemini* se utiliza para generar contenido multimedia que combine imágenes y texto, como el desarrollo automático de anuncios publicitarios o la redacción de descripciones de productos para plataformas de comercio electrónico. Su capacidad para procesar tanto texto como imágenes le permite generar contenido más coherente y contextualizado que los modelos de texto simples (DeepMind, 2024). En el ámbito del arte y el entretenimiento, *Gemini* se utiliza para generar contenido visual y narrativo de manera combinada, lo que es útil en el desarrollo de guiones o en la generación automática de escenas en videojuegos o películas. Esta capacidad de integrar múltiples modalidades le permite generar experiencias inmersivas en las que la narrativa y los elementos visuales trabajan juntos de manera más efectiva (DeepMind, 2024).

Bajo ese contexto, y a pesar de sus avances, *Gemini* enfrenta varias limitaciones técnicas y éticas. Una de las principales limitaciones técnicas es el alto costo computacional asociado con su entrenamiento y uso, dado que procesar múltiples modalidades simultáneamente requiere una cantidad significativa de recursos computacionales. Al igual que otros modelos de lenguaje a gran escala, el uso de *Gemini* puede estar fuera del alcance de organizaciones con recursos limitados, lo que plantea una barrera para su adopción generalizada (Strubell et al., 2019). En términos éticos, existen preocupaciones sobre la generación de contenido engañoso o el uso de *deepfakes*, dada la capacidad de *Gemini* para generar imágenes y texto de manera simultánea. Esto aumenta el potencial de mal uso en la generación de desinformación, lo que podría tener graves consecuencias en áreas como la política o la seguridad pública. Además, al igual que otros modelos de inteligencia artificial, *Gemini* puede reflejar sesgos presentes en los datos con los que fue entrenado, lo que puede resultar en la perpetuación de estereotipos o prejuicios (Bender et al., 2021).

En relación al entrenamiento de *Gemini*, este se desarrolla utilizando grandes conjuntos de datos multimodales que incluyen tanto texto como imágenes y, en versiones futuras, también audio (DeepMind, 2024). Estos conjuntos de datos son extraídos de una variedad de fuentes, como libros, sitios web, bases de datos de imágenes y otros recursos multimedia. Durante el proceso de entrenamiento, *Gemini* aprende a correlacionar información de diferentes modalidades, lo que le permite generar respuestas integradas que combinan múltiples tipos de datos. Este proceso de entrenamiento es altamente intensivo en términos computacionales y generalmente se lleva a cabo en infraestructuras de alto rendimiento, como supercomputadoras. La capacidad del modelo para procesar diferentes tipos de datos simultáneamente lo convierte en una herramienta poderosa, pero también plantea desafíos en términos de escalabilidad y eficiencia energética (Brown et al., 2020; DeepMind, 2024).

Figura 5. Esquema funcional de *Gemini*



Fuente: Elaborado por el autor.

En ese sentido, el futuro de *Gemini* es prometedor, pero enfrenta varios desafíos como mejorar la eficiencia energética y reducir los costos computacionales asociados con su entrenamiento y operación. Así, a medida que los modelos multimodales como *Gemini* se vuelven más complejos, también aumentan los recursos necesarios para mantenerlos, lo que plantea interrogantes sobre la sostenibilidad de su desarrollo (Strubell et al., 2019). Otro desafío clave es garantizar que el contenido generado por *Gemini* sea ético y no perpetúe sesgos o desinformación. Para ello, sería fundamental desarrollar mejores mecanismos de

control y supervisión que aseguren que el modelo se utilice de manera responsable. También se espera que el futuro de *Gemini* incluya mejoras en su capacidad para comprender contextos más complejos y generar contenido de manera más precisa y adaptada a situaciones específicas (DeepMind, 2024).

2.2.3. Perplexity AI

Perplexity AI es un modelo generativo de lenguaje diseñado para responder preguntas utilizando un enfoque conversacional y basado en la búsqueda de información. A diferencia de los modelos de lenguaje convencionales, que suelen basarse exclusivamente en el entrenamiento sobre grandes conjuntos de datos textuales, *Perplexity AI* combina la generación de lenguaje natural con la búsqueda en la web, proporcionando respuestas informadas a partir de información actualizada. El objetivo de *Perplexity AI* es mejorar la precisión y relevancia de las respuestas en tareas de búsqueda conversacional y procesamiento de lenguaje natural (Perplexity AI, 2024). Cabe acotar que el desarrollo de *Perplexity AI* surgió como respuesta a las limitaciones de los modelos puramente generativos, que a menudo producen información desactualizada o incorrecta debido a la naturaleza de su entrenamiento sobre datos estáticos. Al integrar la capacidad de búsqueda, *Perplexity AI* introduce una nueva dimensión a la generación de lenguaje, proporcionando un enfoque híbrido que equilibra la generación de texto con la búsqueda de datos en tiempo real.

El funcionamiento de *Perplexity AI* se basa en una combinación de tecnologías de procesamiento del lenguaje natural y capacidades de búsqueda. A diferencia de modelos como GPT, que generan respuestas únicamente a partir de los patrones aprendidos en los datos de entrenamiento, *Perplexity AI* tiene acceso a una vasta base de datos en línea y utiliza su capacidad de búsqueda para ofrecer respuestas actualizadas y más precisas. Este enfoque mejora la precisión de las respuestas generadas al incluir información externa en tiempo real. En términos de arquitectura, *Perplexity AI* utiliza un modelo basado en *transformers* para la generación de lenguaje, al igual que GPT y otros modelos generativos. Sin embargo, lo que distingue a *Perplexity AI* es su capacidad para interactuar con motores de búsqueda, lo que permite que el modelo responda a preguntas complejas con datos que no están

necesariamente contenidos en su modelo entrenado. Esta capacidad lo convierte en una herramienta más versátil y adaptable para tareas que requieren información actualizada (Perplexity AI, 2024).

Bajo ese contexto, *Perplexity AI* tiene aplicaciones amplias en varios sectores, especialmente aquellos que requieren respuestas precisas y actualizadas basadas en búsquedas web. Entre las principales aplicaciones se incluyen:

- **Educación y aprendizaje**: *Perplexity AI* puede ser utilizado como asistente educativo, proporcionando respuestas informadas a preguntas complejas en tiempo real. Esto lo convierte en una herramienta útil para estudiantes y profesionales que buscan información específica y precisa.
- **Periodismo y medios**: Los periodistas pueden utilizar *Perplexity AI* para acceder rápidamente a información actualizada y verificar datos mientras investigan un tema. Su capacidad para generar respuestas basadas en la búsqueda de datos en línea lo hace valioso en situaciones donde la precisión y la actualidad son cruciales.
- **Atención al cliente**: *Perplexity AI* también puede ser implementado en sistemas de atención al cliente, donde se requiere información relevante y específica en tiempo real. Su capacidad para buscar en la web permite resolver consultas que van más allá de las capacidades de los *chatbots* tradicionales basados en respuestas predefinidas.

Una de las diferencias clave entre *Perplexity AI* y otros modelos generativos como GPT y *Gemini* es su capacidad para realizar búsquedas en la web en tiempo real. Mientras que los modelos estándar de GPT y *Gemini* dependen de los datos con los que fueron entrenados y no pueden acceder a nueva información una vez finalizado su entrenamiento, *Perplexity AI* introduce la capacidad de enriquecer sus respuestas mediante la recuperación de datos en tiempo real. Otro punto diferenciador es su enfoque en la búsqueda conversacional. GPT, por ejemplo, está optimizado para generar texto coherente a partir de patrones de lenguaje, pero carece de la capacidad para verificar o actualizar la información que genera. *Gemini*, por otro lado, es un modelo multimodal que integra texto, imágenes y otros tipos de datos, pero no está

diseñado específicamente para tareas de búsqueda. *Perplexity AI*, en contraste, está orientado hacia la generación de respuestas basadas en búsqueda, lo que lo convierte en una herramienta más adecuada para tareas que requieren información actualizada y validada en tiempo real (Perplexity AI, 2024).

A pesar de sus ventajas, *Perplexity AI* enfrenta ciertas limitaciones técnicas y éticas. Una de las principales limitaciones técnicas es la dependencia de la calidad de los datos en línea. Dado que *Perplexity AI* realiza búsquedas en la web para generar respuestas, la precisión de la información depende de las fuentes disponibles, que pueden no ser siempre fiables o estar sesgadas. Esta característica puede introducir un riesgo de desinformación, similar a los desafíos que enfrentan otros sistemas de búsqueda. En cuanto a las consideraciones éticas, una preocupación relevante es la privacidad de los datos. Dado que *Perplexity AI* realiza búsquedas en la web y accede a múltiples fuentes de información, existe la posibilidad de que recupere datos privados o no autorizados. Además, al igual que otros modelos generativos, *Perplexity AI* puede perpetuar sesgos presentes en los datos recuperados, lo que plantea desafíos éticos en su implementación en aplicaciones sensibles como la justicia o la atención médica (Perplexity AI, 2024).

Cabe acotar que el entrenamiento de *Perplexity AI* sigue un enfoque similar al de otros modelos basados en *transformers*, donde grandes cantidades de datos textuales se utilizan para enseñar al modelo los patrones del lenguaje. Sin embargo, lo que distingue a *Perplexity AI* es su integración con motores de búsqueda. El modelo es optimizado continuamente para mejorar la recuperación de información precisa y relevante a partir de las búsquedas en línea. El proceso de optimización de *Perplexity AI* incluye no solo el ajuste fino de sus capacidades de generación de texto, sino también el mejoramiento de los algoritmos de búsqueda, asegurando que las fuentes recuperadas sean las más adecuadas para la tarea. Además, *Perplexity AI* emplea sistemas de retroalimentación que permiten evaluar la relevancia y la precisión de las respuestas generadas, lo que ayuda a reducir la generación de contenido incorrecto o irrelevante (Perplexity AI, 2024).

Vale recalcar que *Perplexity AI* representa un avance importante en la evolución de los modelos generativos de lenguaje, especialmente por su capacidad para integrar

búsqueda y generación de texto en tiempo real. Su enfoque híbrido, que combina la generación de lenguaje con la búsqueda en la web, podría marcar una tendencia hacia la generación de modelos de lenguaje más adaptables, capaces de acceder a nueva información de manera continua. En el futuro, podría ser probable identificar una mayor adopción de modelos híbridos como *Perplexity AI*, donde las capacidades generativas no estén limitadas por los datos estáticos utilizados en el entrenamiento. Este tipo de modelos tiene el potencial de transformar la manera de interacción con la inteligencia artificial en áreas como la investigación académica, la atención médica, el marketing y la educación, proporcionando información con mayor precisión y actualización (Perplexity AI, 2023).

2.2.4. Copilot

Copilot es un sistema de inteligencia artificial generativa (IAG) desarrollado por la empresa *Microsoft*, diseñado para integrarse en aplicaciones de productividad como *Microsoft 365*, incluyendo *Word, Excel, PowerPoint, Outlook* y *Teams*. Su propósito es asistir a los usuarios en el desarrollo de contenido, análisis de datos, gestión de tareas y comunicación, aprovechando las capacidades avanzadas de modelos de lenguaje como GPT-4 de *OpenAI* (Microsoft, 2024). El origen de *Copilot* se basa en la necesidad de automatizar tareas repetitivas y mejorar la eficiencia, permitiendo a los usuarios concentrarse en tareas de mayor valor estratégico. El lanzamiento de *Copilot* se vincula al avance en modelos generativos y su capacidad para interactuar con aplicaciones empresariales. Al integrarse directamente con herramientas que millones de usuarios ya utilizan, el objetivo de *Copilot* radica en revolucionar la forma en que las personas manejan información y colaboran en sus lugares de trabajo (Microsoft, 2024).

Copilot funciona al integrarse con el ecosistema de aplicaciones de *Microsoft 365*, donde utiliza capacidades de procesamiento de lenguaje natural (PLN) para generar texto, realizar análisis de datos y automatizar flujos de trabajo. Se apoya en grandes modelos de lenguaje como GPT-4, los cuales fueron entrenados en vastos conjuntos de datos textuales y son capaces de generar texto coherente a partir de instrucciones en lenguaje natural. A diferencia de otros modelos generativos, como GPT-4 de *OpenAI* en su forma original, *Copilot* está específicamente optimizado para

su uso en tareas empresariales, proporcionando respuestas contextuales basadas en datos de la empresa o el usuario, como documentos, correos electrónicos o *chats* en *Microsoft Teams*. Este enfoque le permite ofrecer soluciones personalizadas que no están disponibles en modelos más generales que no están integrados en sistemas empresariales concretos (Microsoft, 2024).

Cabe acotar que las aplicaciones de *Copilot* son diversas y están diseñadas para mejorar la productividad en una amplia gama de tareas empresariales. Algunas de las aplicaciones clave incluyen:

- **Automatización de tareas de oficina**: En *Word*, *Copilot* puede generar borradores de documentos, corregir textos y sugerir mejoras estilísticas, mientras que en *Excel* puede analizar datos complejos, desarrollar gráficos y realizar cálculos avanzados a partir de instrucciones simples en lenguaje natural.
- **Análisis de datos**: En *Excel*, *Copilot* facilita el desarrollo de análisis detallados a partir de grandes conjuntos de datos, automatizando el desarrollo de fórmulas y recomendaciones. Este uso es especialmente útil para analistas financieros y de datos que buscan automatizar procesos repetitivos (Microsoft, 2024).
- **Gestión de proyectos y colaboración**: En aplicaciones como *Microsoft Teams* y *Outlook*, *Copilot* asiste en la gestión de correos electrónicos, desarrollo de agendas de reuniones y priorización de tareas, permitiendo a los usuarios concentrarse en actividades más estratégicas.

Al automatizar tareas rutinarias y generar informes, presentaciones y otros documentos a partir de entradas mínimas, *Copilot* permite a los empleados ahorrar tiempo y reducir el esfuerzo necesario para completar trabajos tediosos. Esto no solo busca mejorar la productividad, sino que también está direccionado a permitir a los trabajadores concentrarse en tareas que requieren más creatividad y pensamiento estratégico (Khan, 2024). Además, *Copilot* optimiza el trabajo colaborativo en equipos al generar resúmenes de conversaciones, distribuir tareas automáticamente, y desarrollar agendas basadas en correos electrónicos y

reuniones previas. Esto resulta en una mejor coordinación entre equipos y una reducción en los errores de comunicación, permitiendo que las organizaciones funcionen de manera más eficiente (Microsoft, 2024).

El impacto de *Microsoft Copilot* en entornos empresariales ha sido significativamente satisfactorio. En ese sentido, al automatizar tareas rutinarias y generar informes, presentaciones y otros documentos a partir de entradas mínimas, *Copilot* permite a los empleados ahorrar tiempo y reducir el esfuerzo necesario para completar trabajos tediosos. Esto no solo se refleja como un indicador para mejorar la productividad, sino que también permite a los trabajadores concentrarse en tareas que requieren de mayor creatividad y pensamiento estratégico (Wang et al., 2022). Además, *Copilot* optimiza el trabajo colaborativo en equipos al generar resúmenes de conversaciones, distribuir tareas automáticamente, y desarrollar agendas basadas en correos electrónicos y reuniones previas. Esto resulta en una mejor coordinación entre equipos y una reducción en los errores de comunicación, permitiendo que las organizaciones funcionen con mayor eficiencia.

A pesar de sus muchas ventajas, *Microsoft Copilot* enfrenta varias limitaciones técnicas y éticas. Desde un punto de vista técnico, una de las principales limitaciones es la dependencia del contexto local de los usuarios para generar resultados relevantes. *Copilot* debe estar bien integrado con los datos y documentos del usuario para ofrecer recomendaciones personalizadas. Si esta integración no es adecuada, las sugerencias podrían ser irrelevantes o inexactas. En cuanto a las consideraciones éticas, una preocupación central es la privacidad de los datos. Dado que *Copilot* accede a información empresarial sensible para generar respuestas contextuales, existe el riesgo de que datos confidenciales sean manejados inadecuadamente. *Microsoft* ha implementado estrictas políticas de privacidad y seguridad para mitigar estos riesgos, pero siempre existe la posibilidad de vulneraciones (Microsoft, 2024). Además, la dependencia de *Copilot* en modelos entrenados en grandes conjuntos de datos públicos ha planteado preguntas sobre la equidad y el sesgo en las respuestas generadas, un desafío que comparte con otros sistemas de IA generativa.

Cabe acotar que el entrenamiento de *Microsoft Copilot* se basa en el modelo GPT-4 de *OpenAI*, que fue entrenado en una vasta cantidad de datos textuales provenientes de internet, libros, artículos y otras fuentes públicas. Una vez que el modelo base ha sido entrenado, se adapta específicamente para su integración en aplicaciones de *Microsoft 365*, optimizando su capacidad para manejar tareas empresariales y flujos de trabajo típicos de estos entornos (Microsoft, 2024). Además, *Microsoft Copilot* se personaliza para cada usuario en función de los datos empresariales y personales a los que tiene acceso, como documentos, correos electrónicos y *chats* de *Teams*. Esta capacidad de personalización permite que *Copilot* ofrezca sugerencias más precisas y alineadas con las necesidades específicas de cada usuario (Microsoft, 2024).

Figura 6. Esquema funcional de *Copilot*

Fuente: Elaborado por el autor.

Según Microsoft (2024), *Copilot* marca un hito en el uso de la inteligencia artificial generativa para la productividad y la automatización en entornos empresariales. Su

capacidad para integrarse profundamente en herramientas de uso diario y mejorar la eficiencia de los empleados lo posiciona como un modelo a seguir para el desarrollo futuro de tecnologías IA asistenciales. En ese sentido, a medida que más empresas adopten *Copilot* y otros sistemas similares, podría ser probable que la automatización de tareas complejas y repetitivas se convierta en una norma en todos los sectores (Microsoft, 2024). Según sus desarrolladores, el impacto a largo plazo de *Microsoft Copilot* también podría incluir un cambio en la forma en que los empleados interactúan con la tecnología, permitiéndoles delegar tareas técnicas y centrarse en actividades más estratégicas y creativas. Esto plantea un nuevo paradigma de trabajo en el que la inteligencia artificial no solo apoya, sino que colabora activamente con los seres humanos en la generación de contenido de valor.

A modo de resumen, a continuación se establece un cuadro comparativo con los principales aspectos de las herramientas: *ChatGPT, Gemini, Perplexity AI,* y *Copilot*.

Tabla 1. Cuadro comparativo con los principales aspectos de las herramientas: ChatGPT, Gemini, Perplexity AI, y Copilot

Aspecto	*ChatGPT*	*Gemini*	*Perplexity AI*	*Copilot*
Descripción general	Modelo generativo de lenguaje basado en el procesamiento de lenguaje natural para generación de texto y diálogo.	Modelo multimodal de *Google DeepMind* diseñado para manejar texto, imágenes y audio.	Modelo diseñado para proporcionar respuestas a preguntas y realizar búsquedas contextuales.	Asistente de programación diseñado para facilitar la escritura de código mediante recomendaciones inteligentes.
Desarrollador	*OpenAI*	*Google DeepMind*	*Perplexity*	*GitHub, OpenAI, Microsoft*
Funcionalidad principal	Generación de texto y respuestas conversacionales a partir de entradas textuales.	Generación de contenido multimodal y comprensión avanzada de contextos.	Búsqueda de información y generación de respuestas basadas en datos disponibles.	Sugerencia y autocompletado de código en tiempo real.
Arquitectura subyacente	Arquitectura *Transformer* basada en el aprendizaje auto-regresivo.	Arquitectura avanzada de *Transformer* con capacidades multimodales.	Basado en tecnología de lenguaje natural y algoritmos de búsqueda contextual.	Basado en modelos de lenguaje de *OpenAI* adaptados a programación.

Aspecto	*ChatGPT*	*Gemini*	*Perplexity AI*	*Copilot*
Capacidades	Capacidad de mantener conversaciones, generar texto coherente y responder preguntas en lenguaje natural.	Procesamiento y generación de texto, imágenes y potencialmente otros tipos de datos.	Generación de respuestas precisas a preguntas basadas en contextos predefinidos.	Generación y sugerencia de código en múltiples lenguajes de programación.
Aplicaciones comunes	Asistentes virtuales, chatbots, generación de contenido textual.	Integración en aplicaciones multimedia, descripción de imágenes, generación de contenido gráfico.	Búsquedas informativas, respuestas basadas en contexto.	Asistencia en el desarrollo de software, autocompletado de código.
Ventajas	Interacciones naturales y consistencia en la generación de texto.	Capacidad de comprender y procesar múltiples tipos de datos simultáneamente.	Capacidad para proporcionar respuestas específicas y contextualizadas.	Facilita la programación, reduce tiempo de desarrollo y mejora la productividad.
Desafíos y limitaciones	Requiere grandes volúmenes de datos de entrenamiento. Puede presentar sesgos inherentes y riesgos de desinformación.	Complejidad técnica elevada y dependencia de datos multimodales de alta calidad.	Limitado a información disponible, puede carecer de profundidad en respuestas complejas.	Puede generar código incorrecto o inseguro, limitado por la calidad de los ejemplos de código en su entrenamiento.
Tipo de entrada soportada	Texto	Texto, imágenes y audio.	Texto	Texto
Uso multimodal	No soporta uso multimodal nativamente.	Sí, soporta uso multimodal.	No soporta uso multimodal nativamente.	No soporta uso multimodal nativamente.
Casos de uso destacados	Chatbots, asistentes conversacionales.	Aplicaciones multimedia, desarrollo de contenido gráfico.	Herramientas de búsqueda y consultas informativas.	Asistentes de desarrollo de software y entornos de codificación.

Fuente: Elaborado por el autor.

2.3. Lectura de documentos extensos

2.3.1. NotebookLM

NotebookLM es un modelo de lenguaje avanzado desarrollado con el objetivo de mejorar la forma en que los usuarios interactúan y organizan información escrita. Se enmarca dentro de los modelos de lenguaje generativo que emplean redes neuronales profundas, siendo una evolución directa de sistemas como GPT (Vaswani

et al., 2017). La principal innovación de *NotebookLM* radica en su capacidad para manejar información estructurada y no estructurada en entornos colaborativos, permitiendo a los usuarios procesar y generar texto de manera contextualizada y organizada (Google, 2024). La evolución de estos modelos ha estado fuertemente influenciada por el desarrollo de arquitecturas *transformer*, que han demostrado una capacidad superior para capturar dependencias a largo plazo en el lenguaje natural, en comparación con los enfoques anteriores basados en redes neuronales recurrentes (RNN) o de memoria a largo plazo (LSTM). Este progreso ha sido facilitado por la disponibilidad de grandes cantidades de datos y el incremento en la capacidad de procesamiento computacional, lo cual ha permitido la aparición de modelos con cientos de miles de millones de parámetros, capaces de producir texto coherente y relevante a partir de una variedad de *prompts*.

Cabe acotar que el modelo *NotebookLM* está diseñado específicamente para facilitar la organización de conocimiento y colaborar en entornos de investigación y trabajo académico. A diferencia de otros modelos de lenguaje generativo, como GPT-4 o BERT, que se centran en la generación de texto fluido a partir de datos de entrenamiento generales, *NotebookLM* sobresale en su capacidad para gestionar y procesar grandes volúmenes de información estructurada (Google, 2024). Esto lo hace particularmente útil para tareas académicas y de investigación, donde la síntesis, resumen y análisis de grandes cantidades de datos son esenciales. Entre las aplicaciones principales de *NotebookLM* se incluyen la redacción de artículos científicos, la síntesis de información en documentos largos, y la generación de informes que requieran consistencia temática y conceptual. Además, su integración con sistemas de gestión de conocimientos le permite extraer, analizar y organizar datos con mayor precisión que otros modelos de propósito general (Google, 2024).

En relación a su arquitectura, *NotebookLM* se basa en un *transformer*, una red neuronal profunda que emplea mecanismos de atención auto-regresiva. Esta arquitectura permite que el modelo pueda "prestar atención" a diferentes partes de la secuencia de entrada para generar texto que mantenga coherencia y relevancia (Vaswani et al., 2017). Una de las principales innovaciones de *NotebookLM* es su capacidad para realizar una mejor integración de fuentes heterogéneas de

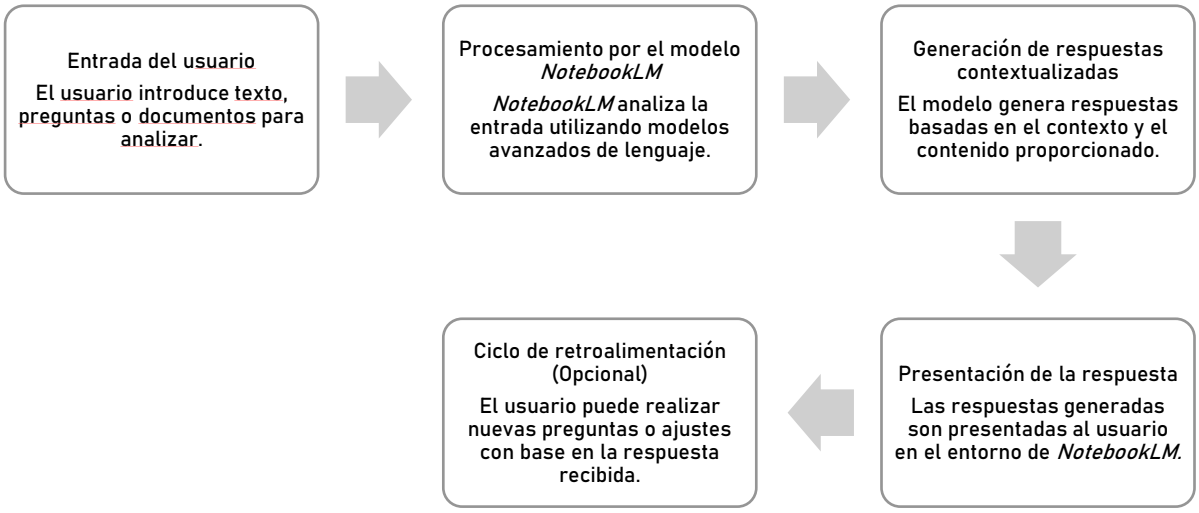
información, combinando datos estructurados (como bases de datos) con texto no estructurado (como artículos o notas). A su vez, el uso de "*multi-head attention*" traducido al español como atención multicabezal, permite al modelo capturar diferentes contextos simultáneamente, lo que es crucial para tareas donde se debe mantener consistencia temática a lo largo de un documento extenso. Además, *NotebookLM* ha sido entrenado con un enfoque particular en la "memoria" de largo plazo, lo que le permite gestionar tareas que requieren recordar y hacer referencia a información mencionada en secciones previas de un documento (Google, 2024).

Otro punto a destacar radica en que, en el ámbito académico, el uso de *NotebookLM* plantea varios desafíos éticos y sociales. Uno de los problemas principales es la cuestión de la autoría y la originalidad. Dado que este tipo de modelos pueden generar texto basado en grandes volúmenes de datos preexistentes, surge la preocupación sobre la propiedad intelectual y los derechos de autor. Si bien *NotebookLM* puede ayudar en la generación de informes o artículos, es crucial que los investigadores supervisen el contenido generado para evitar plagio involuntario. Otro desafío importante es el sesgo inherente a los datos utilizados para entrenar estos modelos. Si bien *NotebookLM* es capaz de generar texto coherente y relevante, su precisión y neutralidad dependen en gran medida de los datos que ha procesado. Si los datos de entrenamiento contienen sesgos, el modelo reproducirá estos sesgos en sus respuestas, lo que puede llevar a resultados problemáticos en contextos académicos sensibles o cuando se trata de temas polémicos (Floridi & Chiriatti, 2020).

En relación a la integración a los procesos académicos, *NotebookLM* puede integrarse mediante la asistencia en la generación de borradores de artículos, la organización de notas de investigación y la síntesis de información compleja. Al actuar como un asistente digital, permite a los académicos ahorrar tiempo en tareas de escritura, mientras se concentran en el análisis crítico y en el desarrollo teórico (Román Acosta, 2023). Es particularmente útil en la redacción de literatura relacionada, donde es necesario sintetizar información de múltiples fuentes. Finalmente, cabe acotar que la integración de *NotebookLM* en el flujo de trabajo académico también puede mejorar la colaboración entre investigadores, ya que

permite compartir documentos generados por IA que pueden ser editados y refinados en conjunto, promoviendo un entorno de trabajo más eficiente. Sin embargo, su uso debe estar siempre acompañado por una revisión exhaustiva de parte del investigador, garantizando la precisión y la integridad del contenido producido.

Figura 7. Esquema funcional de NotebookLM



Fuente: Elaborado por el autor.

2.4. Revisión y generación de contenido científico

2.4.1. Scispace

Scispace es una plataforma impulsada por inteligencia artificial que tiene como objetivo simplificar la comprensión y acceso a la investigación científica. Se especializa en procesar grandes volúmenes de publicaciones científicas y generar resúmenes, análisis de contexto, y facilitar la búsqueda de referencias relevantes (Scispace, 2024). Su algoritmo está basado en procesamiento de lenguaje natural (PLN) que permiten extraer información relevante de textos complejos, facilitando su interpretación. Este tipo de plataformas aplica tecnologías de procesamiento de lenguaje natural y aprendizaje profundo para analizar literatura científica de manera eficiente (Otter et al., 2020). Cabe acotar que este tipo de enfoque dentro de las aplicaciones de inteligencia artificial representa un avance importante para la investigación, ya que reduce el tiempo necesario para revisar grandes cantidades de información, mejorando el acceso a conocimientos clave y acelerando el progreso científico (Dong el at., 2021).

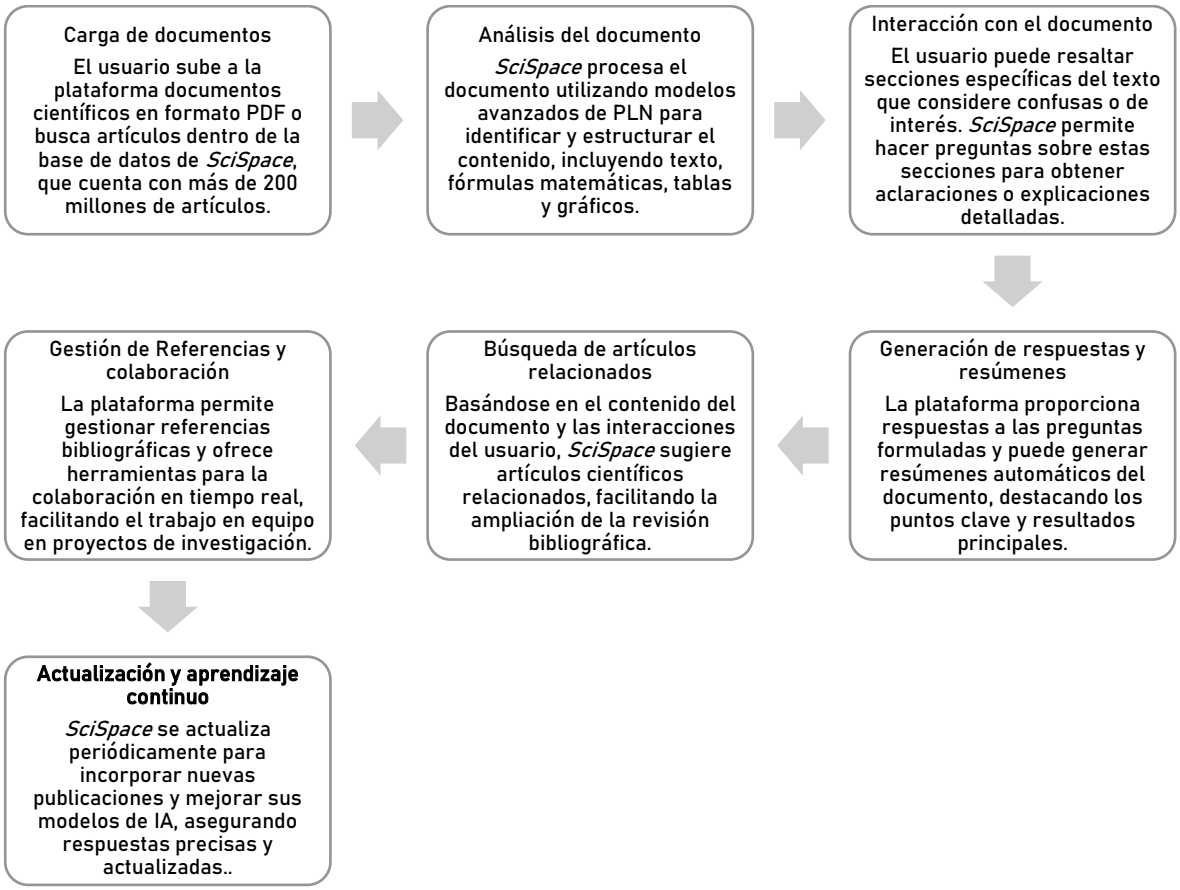
Scispace ofrece varias funcionalidades diseñadas para apoyar a investigadores en diversas áreas del conocimiento. Entre sus características más destacadas se encuentran la capacidad de generar resúmenes automáticos de artículos, la comparación de múltiples estudios de manera estructurada, y la identificación de citas y referencias relevantes dentro de una base de datos masiva de publicaciones científicas (Scispace, 2024). Además de su capacidad para extraer y resumir información científica, *Scispace* facilita la colaboración académica al permitir a los usuarios compartir análisis de textos y resultados en tiempo real. Esta herramienta es aplicable a una amplia gama de áreas científicas, incluyendo ciencias de la computación, medicina, biología, y ciencias sociales, donde el acceso rápido a estudios previos y referencias es crucial para el desarrollo de nuevas investigaciones (Scispace, 2024).

Por su parte, en relación a la mejora del proceso de investigación científica, *Scispace* tiene la capacidad de automatizar tareas repetitivas y técnicas que usualmente consumen tiempo, como la búsqueda y la comparación de literatura. Al utilizar técnicas avanzadas de aprendizaje automático, *Scispace* permite que los investigadores accedan rápidamente a datos relevantes y sintetizados, lo que facilita la toma de decisiones informadas en las primeras etapas de un proyecto de investigación (Scispace, 2024). La plataforma también proporciona herramientas de visualización de datos y relaciones entre estudios, lo que permite a los investigadores identificar patrones, conexiones y lagunas en la literatura existente. Esto no solo mejora la eficiencia en el proceso de revisión de la literatura, sino que también promueve una mayor colaboración entre investigadores de diferentes disciplinas, ya que las referencias y los estudios relevantes son fácilmente accesibles y comparables (Abdul et al., 2021).

El uso de herramientas como *Scispace* plantea varios desafíos éticos en el ámbito académico. Uno de los principales problemas es la cuestión de la autoría y la propiedad intelectual. Al automatizar parte del proceso de análisis y síntesis de información, los investigadores pueden estar menos conscientes de las fuentes originales, lo que puede dar lugar a un uso inadecuado de las citas y referencias (Bender et al., 2021). Además, la confianza excesiva en resúmenes generados por

inteligencia artificial podría conducir a la omisión involuntaria de detalles críticos en los estudios revisados. Otro desafío ético es el posible sesgo en los algoritmos de inteligencia artificial utilizados por plataformas como *Scispace*. Estos sistemas están entrenados con datos previamente disponibles, lo que puede perpetuar sesgos existentes en la literatura científica, especialmente si el conjunto de datos no es lo suficientemente diverso o está incompleto. Por lo tanto, es esencial que los investigadores continúen realizando una revisión crítica y manual del contenido generado por la IA, garantizando la integridad académica de sus trabajos (Floridi & Cowls, 2019).

Figura 8. Esquema funcional de Scispace



Fuente: Elaborado por el autor.

Cabe acotar que *Scispace* puede integrarse en los flujos de trabajo de los investigadores mediante la optimización de la búsqueda y revisión de literatura, la organización de referencias y la generación de resúmenes automáticos (Scispace, 2024). Estas funciones permiten que los investigadores ahorren tiempo en tareas administrativas y enfoquen su energía en el análisis y desarrollo de nuevas ideas. La

plataforma es particularmente útil en la fase inicial de investigación, donde es necesario revisar un amplio espectro de publicaciones en poco tiempo (Scispace, 2024). Además, la integración de *Scispace* con otras plataformas de gestión de conocimiento y *software* de investigación puede mejorar el proceso de colaboración. Bajo ese contexto, los equipos de investigación pueden compartir resúmenes generados automáticamente, ajustar búsquedas de manera colectiva, y mantener una base de datos centralizada de la literatura más relevante para un proyecto específico (Soliman et al., 2023).

2.4.2. Connected Papers

Connected Papers es una herramienta basada en inteligencia artificial que facilita la exploración y visualización de literatura científica (Connected Papers, 2024). Su objetivo principal radica en ayudar a los investigadores a descubrir artículos académicos relevantes relacionados con un tema o un documento específico, generando un gráfico que muestra cómo se conectan los trabajos de investigación en función de sus similitudes (Connected Papers, 2024). Esta plataforma no clasifica los artículos basados en citas directas, sino en un análisis del contenido textual mediante técnicas de procesamiento de lenguaje natural (PLN). Al proporcionar una vista gráfica de los artículos relacionados, *Connected Papers* permite a los investigadores navegar de manera eficiente por la literatura académica, identificando trabajos clave que de otro modo podrían pasar desapercibidos. Herramientas como esta plantean un valor agregado en el ámbito académico, ya que reducen la sobrecarga cognitiva asociada a la revisión exhaustiva de literatura, un aspecto que es cada vez más desafiante debido al creciente volumen de publicaciones científicas (Segura-Bedmar, 2021).

El funcionamiento de *Connected Papers* se basa en un modelo de aprendizaje automático que utiliza técnicas avanzadas de procesamiento de lenguaje natural (PLN). En lugar de centrarse en las citas tradicionales para conectar artículos académicos, la plataforma realiza un análisis semántico del contenido de los documentos. Mediante el uso de redes neuronales, el algoritmo analiza las similitudes entre los trabajos en función de palabras clave, temas tratados y el enfoque metodológico. Esto genera un mapa de "artículos científicos conectados",

donde los artículos más similares se sitúan más cerca en el gráfico, permitiendo a los investigadores identificar rápidamente las publicaciones más relevantes para su estudio (Connected Papers, 2024). El uso de redes neuronales y modelos preentrenados como BERT, son clave para este análisis de similitud semántica. Estos modelos permiten al sistema entender el contexto de las palabras dentro de los textos académicos, capturando la relación entre diferentes artículos científicos de manera más efectiva que los métodos tradicionales basados únicamente en citas (Devlin et al., 2019).

Cabe acotar que las principales aplicaciones de *Connected Papers* se centran en la exploración y organización eficiente de la literatura científica (Connected Papers, 2024). Al generar gráficos visuales que conectan artículos de investigación, la plataforma facilita la identificación de estudios clave y ayuda a los investigadores a comprender rápidamente cómo se relacionan entre sí diferentes publicaciones dentro de un campo específico. Esto es particularmente útil para aquellos que inician una revisión de literatura, ya que les permite construir una base sólida sobre el estado del arte en un área determinada sin depender exclusivamente de las citas directas o recomendaciones automatizadas (Connected Papers, 2024). Además, *Connected Papers* ofrece una herramienta poderosa para identificar vacíos en la literatura. Al visualizar cómo se conectan los artículos académicos, los investigadores pueden detectar áreas que no han sido suficientemente exploradas, lo que les permite enfocar sus esfuerzos en la generación de conocimiento nuevo y relevante (Otter et al., 2020). Asimismo, la plataforma es utilizada en etapas avanzadas del ciclo de investigación para verificar que se han considerado todos los enfoques metodológicos y teóricos pertinentes, promoviendo la rigurosidad en el trabajo académico (Connected Papers, 2024).

En relación a los principales beneficios de *Connected Papers* en comparación con otras herramientas de búsqueda académica, destaca su enfoque en la visualización de la relación semántica entre documentos. Mientras que bases de datos académicos como *Google Scholar* o *Web of Science* se centran en el número de citas y referencias, *Connected Papers* utiliza análisis semántico para ofrecer una comprensión más profunda de cómo se conectan los trabajos por su contenido

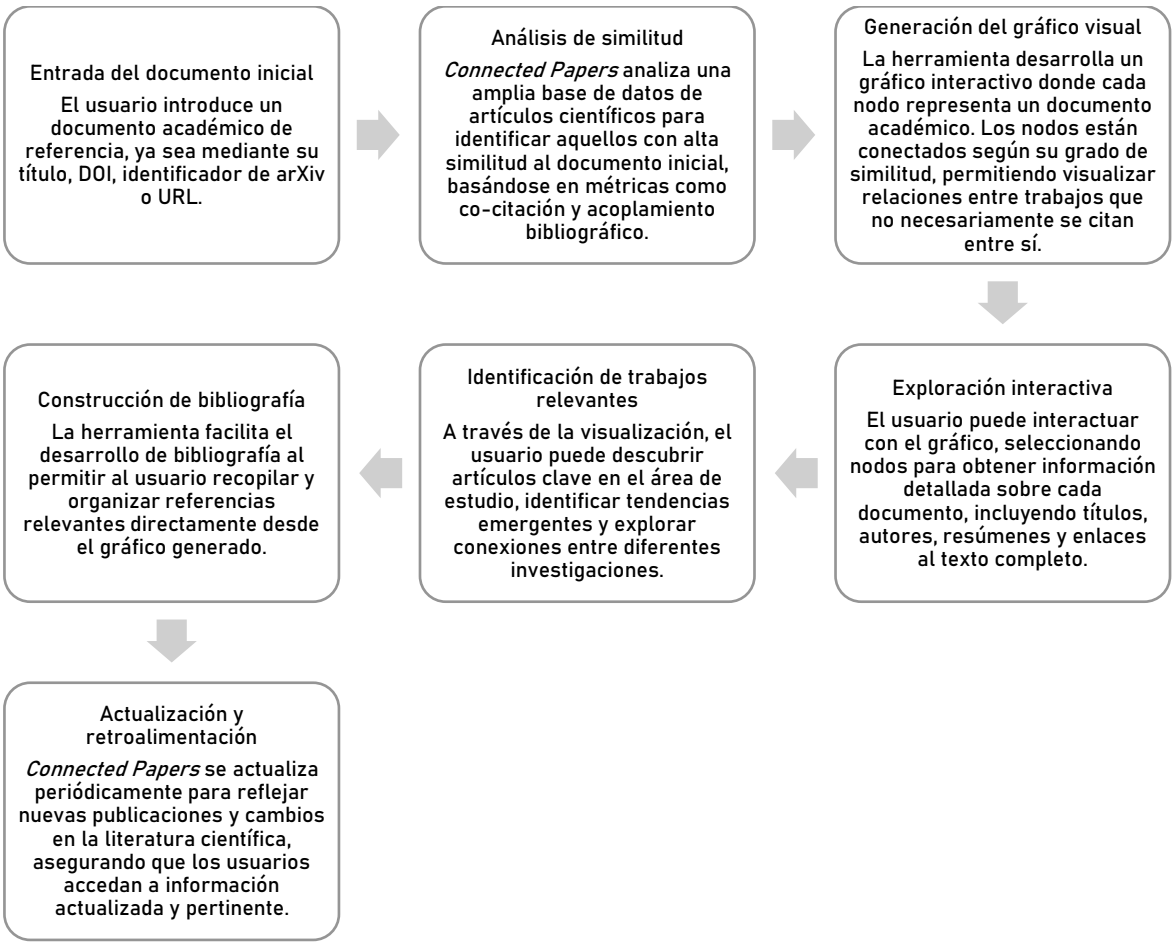
(Connected Papers, 2024). Esta diferencia metodológica es crucial, ya que muchos artículos de investigación que no están fuertemente citados pueden ser, sin embargo, conceptualmente relevantes para un área de estudio. Además, la interfaz visual que proporciona *Connected Papers* simplifica el proceso de descubrir literatura relacionada, permitiendo a los usuarios navegar a través de los resultados de forma intuitiva y encontrar trabajos complementarios con facilidad. Esto es especialmente útil para investigadores multidisciplinarios que necesitan integrar conocimientos de diferentes campos.

Al igual que otras herramientas impulsadas por inteligencia artificial, *Connected Papers* enfrenta desafíos éticos relacionados con el sesgo de los algoritmos y la dependencia de las fuentes de datos disponibles. Aunque la plataforma proporciona un análisis basado en similitudes semánticas, su precisión depende en gran medida de la calidad y diversidad de los datos con los que ha sido entrenado el modelo (Connected Papers, 2024). Si los conjuntos de datos no son lo suficientemente inclusivos o si existen sesgos en la literatura utilizada para entrenar el modelo, las recomendaciones y conexiones pueden reflejar esos sesgos. Además, existe el riesgo de que los investigadores dependan excesivamente de estas herramientas automatizadas, lo que podría llevar a la omisión de trabajos relevantes que no sean identificados por el sistema. Por lo tanto, es fundamental que los usuarios combinen el uso de herramientas como *Connected Papers* con un análisis crítico y manual de la literatura para garantizar la exhaustividad y la neutralidad en su revisión (Bender et al., 2021).

Finalmente, en relación al flujo de trabajo de la investigación académica, *Connected Papers* se integra de manera efectiva en el flujo de trabajo de los investigadores académicos en diversas fases del proceso de investigación. Durante la fase de revisión de literatura, los usuarios pueden emplear la plataforma para descubrir artículos que no hubieran sido identificados a través de búsquedas convencionales, lo que les permite obtener una comprensión más completa sobre un tema específico. La visualización de conexiones entre artículos también es útil para organizar las fuentes en torno a temas específicos, facilitando la escritura de artículos de revisión o el desarrollo de hipótesis. En las etapas posteriores, la herramienta puede ser

utilizada para actualizar la revisión de la literatura existente, garantizando que se consideren los últimos avances en un campo de estudio. Asimismo, al ser una herramienta visual y accesible, facilita la colaboración entre equipos de investigación, quienes pueden compartir gráficos de artículos académicos conectados para discutir nuevas ideas o enfoques metodológicos.

Figura 9. Esquema funcional de *Connected Papers*



Fuente: Elaborado por el autor.

2.4.3. Semantic Scholar

Semantic Scholar es un motor de búsqueda académica impulsado por inteligencia artificial diseñado para mejorar la forma en que los investigadores acceden, interpretan y utilizan la literatura científica (Semantic Scholar, 2024). Desarrollado por el *Allen Institute for AI*, su propósito principal es ayudar a los investigadores a navegar por grandes volúmenes de publicaciones científicas, proporcionando resultados más relevantes y enfocados en aspectos clave como las citas influyentes, los resúmenes automatizados y las recomendaciones de trabajos relacionados

(Semantic Scholar, 2024). Una de las características más distintivas de *Semantic Scholar* es su capacidad para identificar automáticamente citas clave dentro de un artículo académico, lo que facilita el acceso a la información más relevante dentro de un campo de estudio. Al aplicar técnicas avanzadas de procesamiento de lenguaje natural (PLN) y aprendizaje profundo, la plataforma puede filtrar millones de documentos, ayudando a los usuarios a enfocarse en los estudios más influyentes y actuales en su área.

El motor de *Semantic Scholar* utiliza un enfoque avanzado de inteligencia artificial que combina técnicas de procesamiento de lenguaje natural (PLN), análisis de citas y aprendizaje automático. La plataforma emplea modelos de redes neuronales profundas, que permiten analizar tanto el contenido de los artículos como las citas entre ellos, para generar una comprensión más contextual de los temas investigados (Semantic Scholar, 2024). Uno de los avances más importantes en *Semantic Scholar* es su capacidad para generar resúmenes automáticos de los artículos, destacando las contribuciones clave de cada documento. Esto se logra mediante la aplicación de algoritmos de comprensión de texto como BERT, que permiten a la plataforma identificar las ideas centrales de un texto académico y proporcionar una visión resumida que facilita la toma de decisiones en la búsqueda de literatura (Devlin et al., 2019). Además, *Semantic Scholar* utiliza un sistema de análisis de citas influyentes, donde destaca aquellas referencias dentro de un artículo que han tenido un impacto significativo en el campo. Este tipo de procesos permite a los investigadores enfocarse en los trabajos más relevantes.

Entre las principales funcionalidades de *Semantic Scholar* se encuentran la búsqueda avanzada de artículos científicos, la identificación de citas influyentes y el uso de resúmenes automatizados (Semantic Scholar, 2024). Estas características ayudan a los investigadores a localizar rápidamente la literatura más relevante y comprender el contexto de un artículo académico sin necesidad de leer todo el documento. La plataforma también ofrece un sistema de recomendaciones basado en las búsquedas y los intereses de los usuarios, que sugiere artículos académicos relacionados de manera precisa. El análisis de citas influyentes es una de las herramientas más valiosas de *Semantic Scholar* , ya que permite a los investigadores

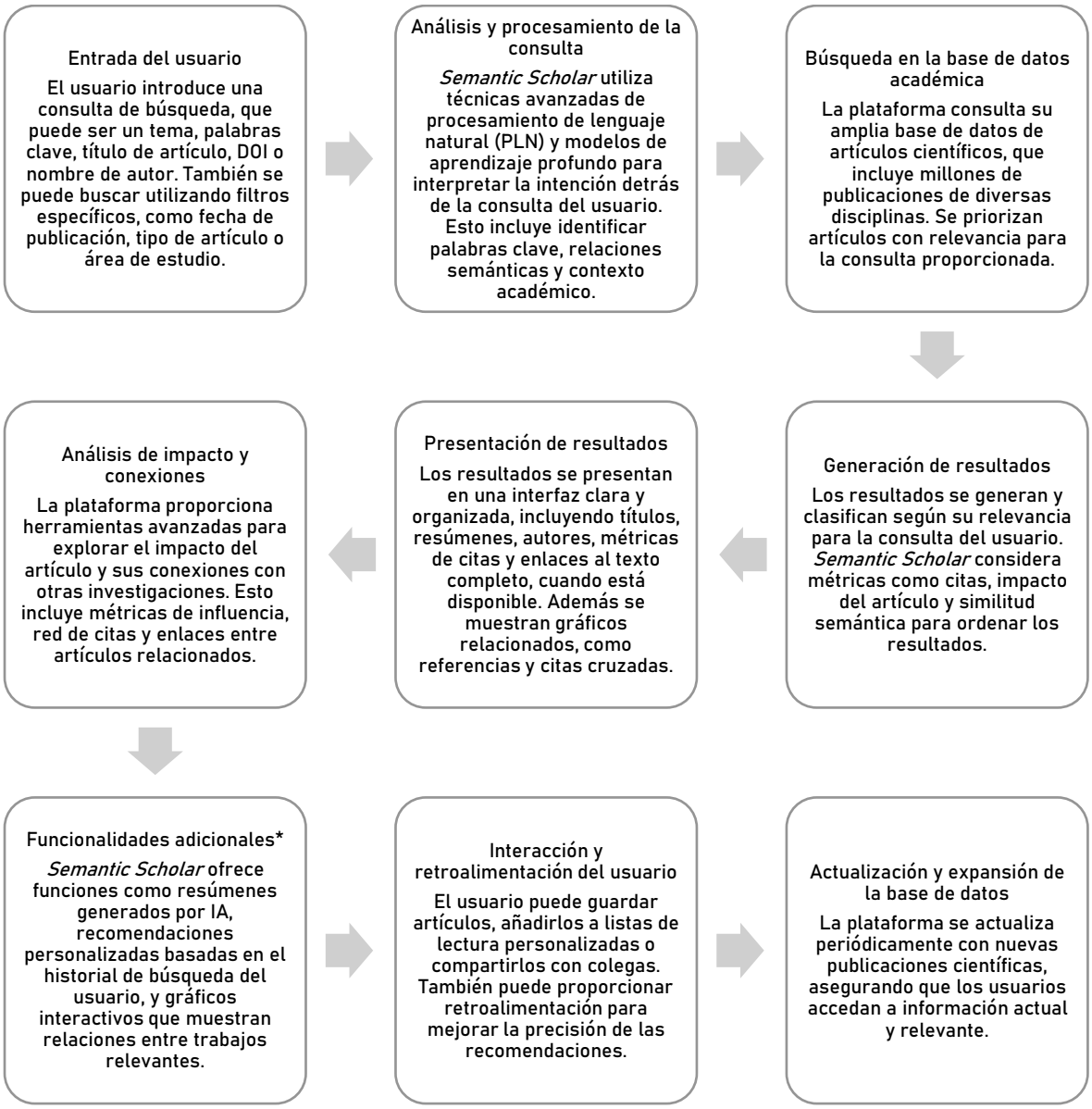
identificar no solo qué artículos académicos son más citados, sino cuáles han tenido un impacto significativo en su campo. Esto es particularmente útil en áreas en rápida evolución, donde mantenerse al día con la literatura puede ser un desafío (Wang et al., 2020). Además, la plataforma permite a los usuarios acceder a datos bibliométricos, como el índice h, lo que proporciona una visión más completa del impacto de un artículo o autor en el campo académico.

Cabe acotar que *Semantic Scholar* ha demostrado ser una herramienta esencial para mejorar el proceso de revisión de literatura en la investigación académica. A través de su capacidad para filtrar la información más relevante, ofrece a los investigadores un enfoque más eficiente y enfocado para identificar los estudios más influyentes y pertinentes (Semantic Scholar, 2024). El uso de resúmenes automáticos permite a los investigadores obtener una comprensión rápida del contenido de los artículos académicos, lo que resulta en una mayor productividad al revisar grandes volúmenes de literatura. Asimismo, la plataforma permite a los usuarios realizar un seguimiento de los avances en campos específicos a través de la funcionalidad de alertas personalizadas. Estas alertas notifican al usuario cuando se publican artículos académicos nuevos o cuando un trabajo relevante es citado en la literatura (Semantic Scholar, 2024). Esto asegura que los investigadores se mantengan actualizados en sus áreas de estudio sin la necesidad de realizar búsquedas manuales constantemente.

Uno de los principales desafíos éticos asociados con *Semantic Scholar* es la cuestión del sesgo en los algoritmos de inteligencia artificial. A pesar de que la plataforma busca proporcionar resultados más precisos y relevantes, los algoritmos que utilizan para analizar y recomendar artículos académicos pueden estar influenciados por los datos de entrenamiento, lo que podría llevar a la sobre-representación de ciertos autores o enfoques metodológicos en detrimento de otros. Este sesgo algorítmico es una preocupación creciente en el campo de la inteligencia artificial aplicada a la investigación (Bender et al., 2021). Otro desafío es el posible impacto en la diversidad de la investigación. Al centrarse en las citas influyentes y los artículos más populares, *Semantic Scholar* podría desincentivar la exploración de trabajos menos citados pero con mayor innovación, lo que podría limitar la diversidad de enfoques

dentro de un campo. Esto resalta la necesidad de un uso equilibrado y crítico de la plataforma por parte de los investigadores, quienes deben complementarla con otros métodos de búsqueda y revisión de literatura.

Figura 10. Esquema funcional de *Semantic Scholar*



Fuente: Elaborado por el autor.

Finalmente, es necesario recalcar que *Semantic Scholar* se integra eficazmente en el flujo de trabajo de los investigadores mediante la provisión de herramientas que facilitan el acceso rápido a la literatura relevante y la organización de referencias clave. Durante las primeras etapas de la investigación, la plataforma permite a los investigadores explorar el estado del arte en su campo y descubrir artículos

académicos relacionados a través de su sistema de recomendaciones inteligentes. En etapas posteriores, los resúmenes generados por la plataforma permiten a los investigadores centrarse en los aspectos más relevantes de los artículos sin tener que revisarlos por completo (Semantic Scholar, 2024). Además, *Semantic Scholar* facilita la colaboración entre investigadores al permitir el acceso a un vasto corpus de publicaciones científicas. A través de la función de alertas y la capacidad de seguir el trabajo de autores específicos, los equipos de investigación pueden mantenerse actualizados y compartir artículos relevantes con mayor facilidad, mejorando así la calidad del trabajo colaborativo.

2.4.4. ResearchRabbit

ResearchRabbit es una plataforma de búsqueda y organización de información científica que utiliza inteligencia artificial para ayudar a los investigadores a descubrir literatura relevante y seguir creaciones en tiempo real dentro de áreas específicas de investigación (ResearchRabbit, 2024). Su objetivo es optimizar la forma en que los investigadores encuentran, rastrean y organizan publicaciones académicas y referencias bibliográficas. A diferencia de los motores de búsqueda tradicionales, *ResearchRabbit* permite a los usuarios construir "mapas de investigación" interactivos, donde las relaciones entre artículos académicos, autores y temas pueden visualizarse fácilmente (ResearchRabbit, 2024). La plataforma es especialmente útil para facilitar el descubrimiento de nueva literatura y seguir las tendencias emergentes dentro de variedad de campos de estudio. A su vez, herramientas como *ResearchRabbit* permiten a los usuarios no solo encontrar artículos académicos relacionados, sino también establecer conexiones entre investigaciones que podrían no estar directamente vinculadas por citas, pero que comparten temas o enfoques metodológicos similares (Wang et al., 2020). Esto ofrece una perspectiva más amplia de un campo académico y ayuda a los investigadores a mantenerse al día en un entorno donde la cantidad de publicaciones aumenta exponencialmente.

La inteligencia artificial en *ResearchRabbit* se basa principalmente en técnicas de procesamiento de lenguaje natural (PLN) y aprendizaje profundo. Utiliza algoritmos de aprendizaje automático para analizar el contenido de los artículos científicos,

generar recomendaciones basadas en patrones textuales y conectar investigaciones que comparten enfoques o áreas temáticas (ResearchRabbit, 2024). La plataforma también emplea métodos de análisis semántico para identificar relaciones entre trabajos que no están conectados explícitamente a través de citas. Para los investigadores este proceso permite descubrir literatura relevante que de otra manera podría ser pasada por alto (Otter et al., 2020). Cabe acotar que una de las tecnologías más destacadas utilizadas en *ResearchRabbit* es el desarrollo de redes de investigación visuales que muestran cómo los artículos académicos, autores y temas están interconectados. Estas redes son dinámicas y permiten a los usuarios explorar un campo de investigación desde diferentes perspectivas, identificando relaciones que no son evidentes en las bases de datos de búsqueda tradicionales (Devlin et al., 2019).

Las principales funcionalidades de *ResearchRabbit* incluyen el desarrollo de bibliotecas personalizadas, la visualización de redes de relaciones entre artículos académicos y autores, y la capacidad de rastrear la evolución de un campo de investigación en tiempo real. Los investigadores pueden seguir temas, autores o artículos específicos, y la plataforma les notifica cuando se publican nuevos trabajos relevantes o cuando se citan estudios en los que están interesados (ResearchRabbit, 2024). Esta capacidad de mantenerse actualizada en tiempo real es una ventaja significativa en campos de investigación de rápido crecimiento. Otra funcionalidad clave es la capacidad de descubrir artículos académicos y autores que pueden no estar directamente relacionados por citas (ResearchRabbit, 2024). A su vez, *ResearchRabbit* permite a los usuarios identificar conexiones temáticas y metodológicas entre investigaciones que podrían no ser obvias en una revisión tradicional de la literatura. Esto amplía el alcance del análisis bibliográfico y permite a los investigadores explorar áreas emergentes o enfoques novedosos en su campo (Wang et al., 2020).

En relación a la mejora del proceso de revisión de literatura, *ResearchRabbit* lo desarrolla a partir de la automatización y visualización de relaciones entre trabajos científicos. La plataforma no solo facilita la búsqueda de artículos académicos relevantes, sino que también ayuda a los investigadores a organizar y estructurar la

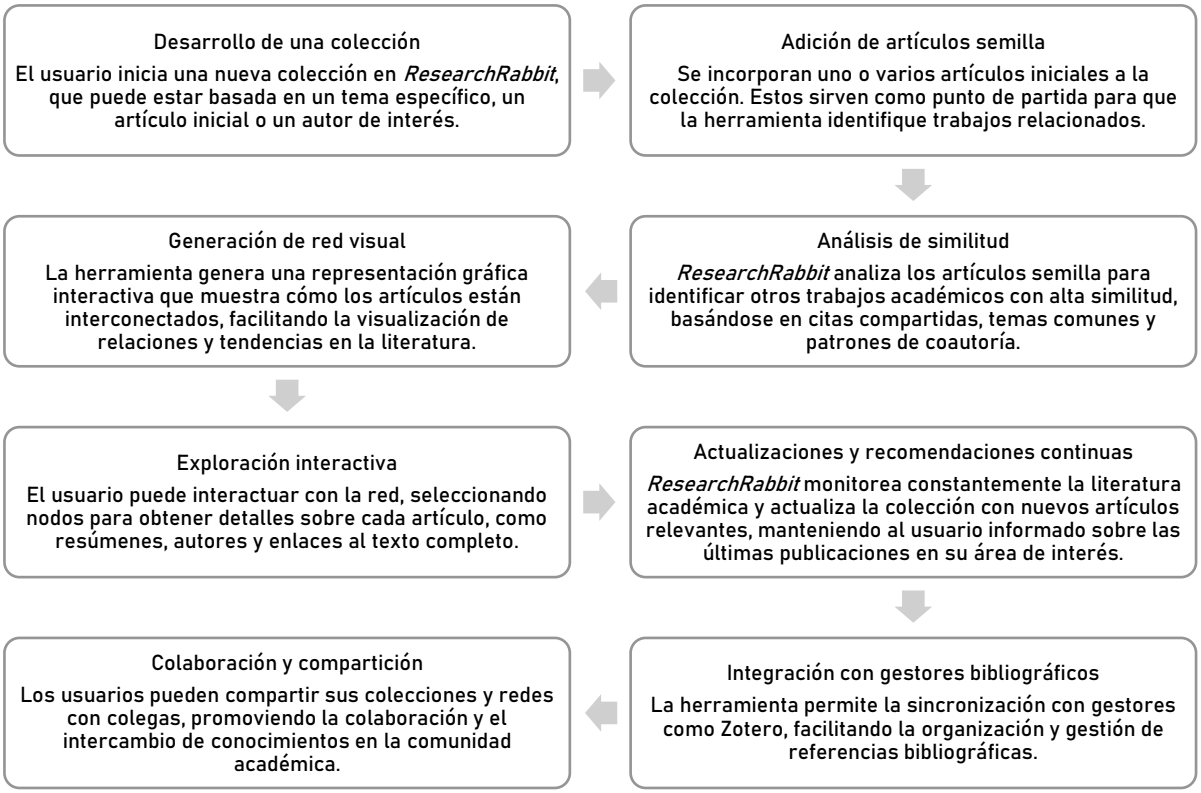
información que encuentran. Al permitir el desarrollo de "mapas de investigación" interactivos, los usuarios pueden visualizar cómo los artículos académicos se relacionan entre sí, lo que les permite identificar rápidamente los estudios clave dentro de un campo determinado. Este enfoque visual y estructurado reduce el tiempo que los investigadores necesitan para revisar la literatura, mejorando la eficiencia en las primeras etapas de un proyecto de investigación. Además, la capacidad de recibir notificaciones sobre nuevas publicaciones o citas permite a los investigadores mantenerse al día con los avances más recientes en sus áreas de estudio, asegurando que sus revisiones sean más completas y actualizadas.

Cabe acotar que el uso de *ResearchRabbit* plantea desafíos éticos similares a las de otras plataformas impulsadas por inteligencia artificial en la investigación científica. Uno de los principales desafíos es el sistema incorporado en los algoritmos de búsqueda y recomendación. Si bien la plataforma facilita la identificación de artículos académicos relacionados, la precisión y la relevancia de sus recomendaciones dependen en gran medida de los datos con los que ha sido entrenado el modelo. Esto puede llevar a la sobre-representación de ciertos enfoques o autores, excluyendo investigaciones menos citadas pero igualmente relevantes (Bender et al., 2021). Otro desafío ético es la posible dependencia excesiva de los investigadores en estas herramientas automatizadas. Si bien *ResearchRabbit* puede agilizar el proceso de búsqueda de literatura, es crucial que los usuarios mantengan un enfoque crítico y complementen los resultados proporcionados por la plataforma con búsquedas manuales. Esto es necesario para garantizar la integridad académica y evitar la omisión de trabajos relevantes que podrían no haber sido identificados por el algoritmo.

Finalmente, es necesario resaltar que *ResearchRabbit* se integra eficazmente en el flujo de trabajo de los investigadores al proporcionar una plataforma que facilita la búsqueda, organización y seguimiento de la literatura académica. Durante las primeras etapas de la investigación, los usuarios pueden utilizar *ResearchRabbit* para descubrir artículos clave y explorar cómo estos están interrelacionados. Las redes de investigación visuales proporcionan una estructurada de las conexiones entre trabajos académicos, lo que permite a los investigadores identificar tendencias

emergentes o vacíos en la literatura. En las etapas posteriores, *ResearchRabbit* ayuda a los investigadores a gestionar su bibliografía y mantenerse al día con las publicaciones más recientes en sus áreas de interés, lo que garantiza que sus proyectos se mantengan actualizados y relevantes. Además, la capacidad de colaborar en tiempo real con otros usuarios mejora la calidad de la investigación colaborativa, ya que los equipos pueden compartir hallazgos y referencias con mayor eficiencia.

Figura 11. Esquema funcional de *ResearchRabbit*



Fuente: Elaborado por el autor.

2.4.5. Consensus

Consensus es una plataforma impulsada por inteligencia artificial diseñada para ayudar a los investigadores a extraer información relevante de grandes volúmenes de literatura científica (Consensus, 2024). Su objetivo principal es facilitar el acceso a conclusiones basadas en evidencia, permitiendo a los usuarios identificar rápidamente el consenso o disenso en torno a un tema específico dentro de la comunidad científica. A través de un sistema de procesamiento de lenguaje natural (PLN), *Consensus* puede analizar artículos científicos y extraer conclusiones claves

de manera más eficiente que los métodos tradicionales de búsqueda (Consensus, 2024). El propósito de *Consensus* en la investigación científica radica en simplificar el proceso de revisión de literatura, y, al mismo tiempo, mejorar la toma de decisiones basada en evidencia. Al proporcionar resúmenes automatizados de las conclusiones más relevantes en torno a un tema, este tipo de plataformas permiten a los investigadores concentrarse en interpretar los resultados en lugar de invertir tiempo en la búsqueda y análisis manual de información (Otter et al., 2020).

La inteligencia artificial en *Consensus* utiliza técnicas avanzadas de procesamiento de lenguaje natural (PLN) y aprendizaje automático. El sistema analiza textos científicos para identificar patrones, resúmenes y conclusiones clave, lo que permite a los investigadores obtener rápidamente una visión general de lo que ha sido identificado en un campo específico (Consensus, 2024). Para lograr esto, la plataforma se basa en modelos preentrenados, como BERT, que permiten al sistema comprender el contexto de los textos académicos y extraer información relevante de manera precisa (Devlin et al., 2019). Otra clave de tecnología utilizada por *Consensus* es el análisis semántico, que permite al sistema identificar relaciones conceptuales entre diferentes artículos académicos (Consensus, 2024). Esto permite a los usuarios no solo identificar el contenido en un tema, sino también explorar áreas de divergencia o controversia dentro de la comunidad científica.

Por otro lado, las principales aplicaciones de *Consensus* se centran en la optimización del proceso de búsqueda de información y la extracción automática de conclusiones relevantes (Consensus, 2024). Esto beneficia a los investigadores en diversas etapas del ciclo de investigación. Durante la fase de revisión de la literatura, *Consensus* permite a los usuarios obtener rápidamente un resumen de las principales conclusiones en torno a un tema, reduciendo así el tiempo necesario para realizar una revisión exhaustiva de la literatura (Consensus, 2024). Esta funcionalidad es especialmente útil en áreas científicas que generan grandes volúmenes de publicaciones en poco tiempo. Además, *Consensus* es utilizado para identificar vacíos en la investigación, lo que permite a los investigadores centrarse en áreas donde el contenido científico aún no ha sido alcanzado. Esta capacidad de destacar áreas de incertidumbre o controversia es clave para orientar

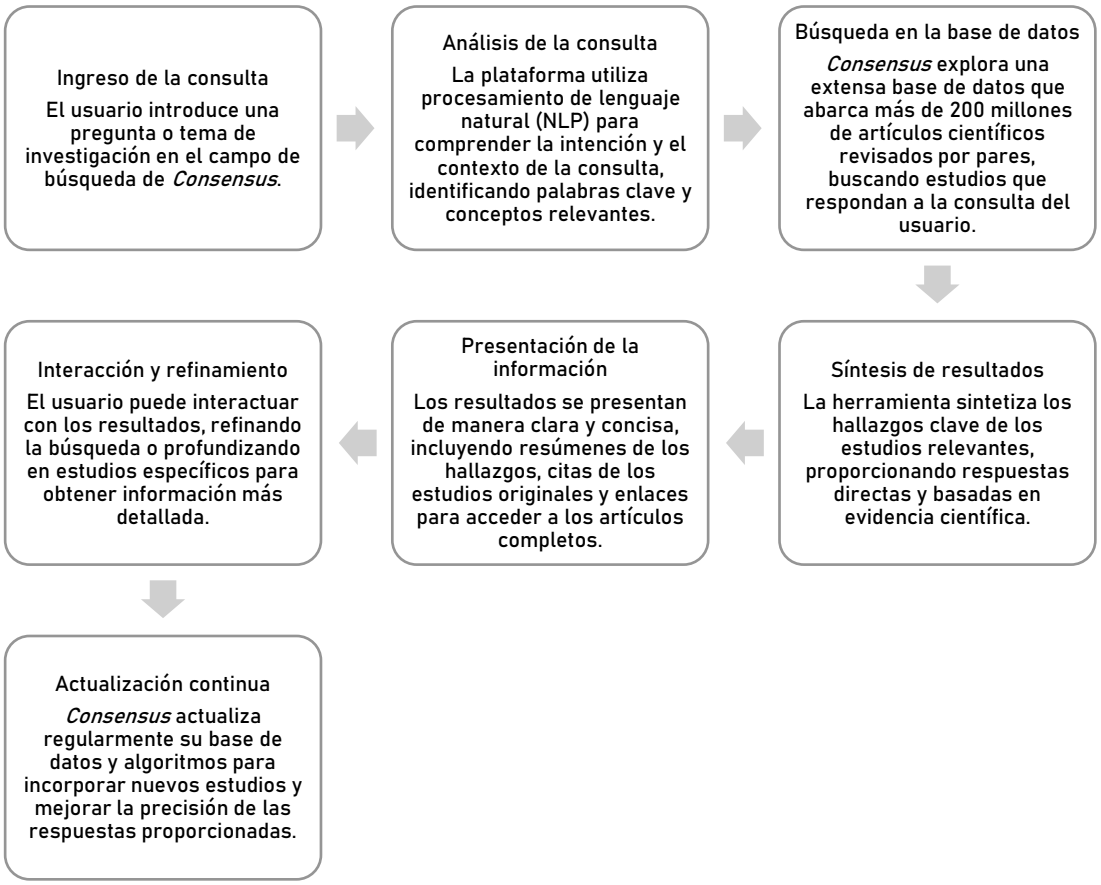
investigaciones futuras y diseñar estudios que contribuyan al avance del conocimiento en campos específicos.

Cabe acotar que *Consensus* facilita el proceso de toma de decisiones basado en evidencia al proporcionar resúmenes automatizados de los puntos claves encontrados en la literatura científica (Consensus, 2024). En lugar de que los investigadores deban leer una gran cantidad de artículos académicos para formarse una idea general de lo que ha sido explorado en un campo específico, la plataforma extrae las conclusiones más relevantes, lo que permite a los investigadores evaluar rápidamente la dirección general de las investigaciones previas (Consensus, 2024). Al ofrecer una visión clara y objetiva del contenido científico, la plataforma ayuda a los investigadores a fundamentar sus decisiones en datos empíricos y resultados contrastados. Asimismo, herramientas como *Consensus* promueven una mayor transparencia en la investigación, ya que permite visualizar de manera clara tanto las áreas de acuerdo como las de controversia dentro de un campo (Wang et al., 2020).

En relación al uso de *Consensus* en la investigación científica, se identifican desafíos éticos relacionados con el sesgo en los algoritmos y la interpretación automática de resultados científicos. Aunque la plataforma puede procesar grandes volúmenes de literatura científica de manera eficiente, los algoritmos que utilizan pueden estar influenciados por los datos de entrenamiento, lo que podría sesgar los resultados hacia ciertos enfoques o autores. Este sesgo algorítmico es una preocupación constante en las plataformas impulsadas por inteligencia artificial, ya que puede afectar la neutralidad de las conclusiones presentadas (Bender et al., 2021). Otro desafío ético es la posible reducción de la revisión crítica de la literatura por parte de los investigadores. Dado que *Consensus* automatiza la extracción de conclusiones, existe el riesgo de que los usuarios dependan en exceso de los resúmenes proporcionados por la plataforma, sin realizar una lectura detallada de los estudios originales. Esto podría llevar a la omisión de matices importantes en la interpretación de los resultados científicos, lo que subraya la necesidad de un uso complementario de la plataforma junto con métodos tradicionales de revisión académica.

Finalmente, cabe acotar que *Consensus* se integra eficazmente en el flujo de trabajo de los investigadores durante las etapas iniciales de la revisión de literatura y el análisis de datos. Al proporcionar un acceso rápido a los tópicos relevantes en la literatura, los investigadores pueden estructurar sus revisiones de manera más eficiente y centrada en áreas donde existen vacíos de conocimiento. Además, la plataforma facilita el descubrimiento de nuevas publicaciones relevantes a través de su sistema de recomendaciones basado en el análisis de la literatura previa. En comparación con otras herramientas, *Consensus* ofrece una ventaja significativa en la visualización y extracción de conclusiones a nivel macro, permitiendo a los usuarios identificar rápidamente tendencias generales en un campo específico (Consensus, 2024). Mientras que otras plataformas se enfocan principalmente en la búsqueda y citación, *Consensus* destaca por su enfoque en la interpretación de resultados, lo que es crucial para la toma de decisiones fundamentadas en la investigación científica (Consensus, 2024).

Figura 12. Esquema funcional de *Consensus*



Fuente: Elaborado por el autor.

2.4.6. Elicit

Elicit es una herramienta de inteligencia artificial diseñada para apoyar a los investigadores en la revisión de literatura y el análisis de datos científicos (Elicit, 2024). Su objetivo principal radica en mejorar la toma de decisiones basada en evidencia, al automatizar la búsqueda y la organización de información relevante con mayor eficiencia. Cabe acotar que *Elicit* emplea técnicas avanzadas de procesamiento de lenguaje natural (PLN) para interpretar y sintetizar grandes volúmenes de información académica, permitiendo a los investigadores acceder rápidamente a las respuestas más relevantes dentro de varios campos de estudio (Elicit, 2024). El propósito de *Elicit* en la investigación académica radica en reducir el tiempo necesario para realizar una revisión exhaustiva de la literatura, ayudando a los investigadores a encontrar estudios relevantes, analizar sus métodos y resultados, y generar preguntas de investigación bien fundamentadas (Elicit, 2024).

La inteligencia artificial en *Elicit* se basa en modelos de procesamiento de lenguaje natural (PLN) y aprendizaje automático. La plataforma utiliza algoritmos para procesar y extraer información clave de textos académicos, sintetizando el contenido en un formato que permite a los usuarios obtener respuestas claras y rápidas a las preguntas de investigación (Elicit, 2024). Al igual que otros sistemas de PLN avanzados, *Elicit* emplea redes neuronales profundas, como BERT, que le permiten analizar el contenido de los artículos y generar resúmenes de los puntos más relevantes (Devlin et al., 2019). Una característica clave de *Elicit* es su capacidad para generar respuestas basadas en evidencia. Al analizar múltiples fuentes de datos y estudios, la plataforma ofrece una visión clara de las conclusiones científicas en torno a un tema, lo que ayuda a los investigadores a tomar decisiones informadas sobre el diseño de sus propios estudios o la interpretación de datos existentes.

En relación a las aplicaciones en la investigación científica, *Elicit* destaca principalmente en la revisión de literatura y en el análisis crítico de estudios previos. Su capacidad para sintetizar grandes volúmenes de información permite a los investigadores evaluar rápidamente el estado del arte en un campo específico y generar preguntas de investigación fundamentadas. Además, herramientas como *Elicit* son útiles para identificar patrones y tendencias en la investigación existente,

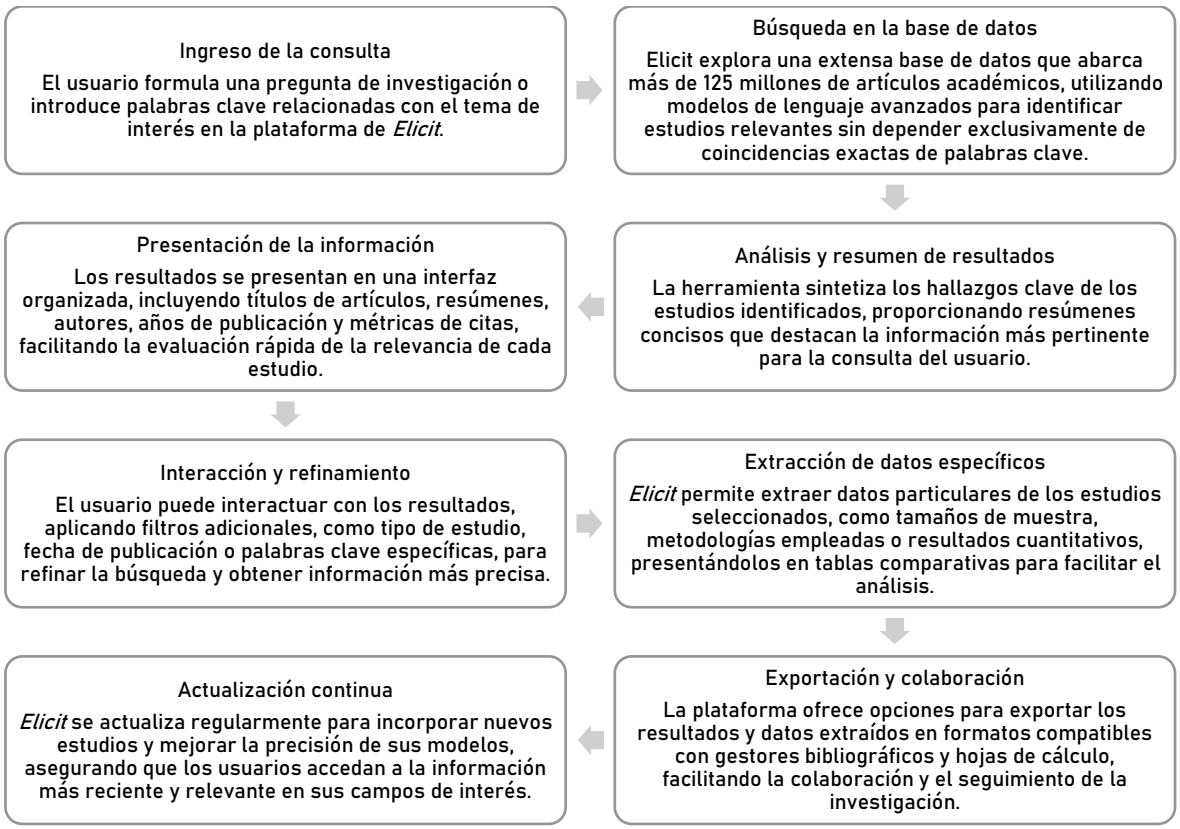
lo que facilita la detección de vacíos de conocimiento que pueden ser abordados en estudios futuros. A su vez, una de las contribuciones más significativas de *Elicit* es su capacidad para mejorar la toma de decisiones basada en evidencia al permitir que los investigadores accedan rápidamente a conclusiones relevantes y bien fundamentadas (Elicit, 2024). Al sintetizar información de múltiples estudios, la plataforma facilita la identificación de patrones y tendencias que respaldan decisiones críticas en el proceso de investigación. Esta capacidad es particularmente valiosa en campos donde se requiere una gran cantidad de datos para fundamentar las decisiones, como la medicina y las ciencias sociales (Wang et al., 2020). *Elicit* también promueve una mayor transparencia en el proceso de toma de decisiones, ya que proporciona resúmenes claros de las fuentes y estudios que se utilizan para generar conclusiones. Esto permite a los investigadores realizar un seguimiento de las decisiones basado en evidencia y justificar sus elecciones metodológicas de manera más clara.

Cabe acotar que uno de los principales desafíos éticos asociados con el uso de *Elicit* es el sesgo algorítmico. Al igual que muchas herramientas impulsadas por inteligencia artificial, *Elicit* depende de los datos de entrenamiento para generar respuestas y recomendaciones. Si los datos utilizados están sesgados o incompletos, el sistema podría generar conclusiones que no reflejan de manera precisa la realidad del campo de estudio. Esto es especialmente preocupante en áreas donde los estudios disponibles no representan de manera equitativa a todas las poblaciones o enfoques metodológicos (Bender et al., 2021). Otro desafío ético es la dependencia excesiva de los investigadores en las recomendaciones automáticas de *Elicit*. Aunque la plataforma puede proporcionar resúmenes útiles de la literatura, es importante que los investigadores mantengan un enfoque crítico y complementen los resultados automatizados con una revisión exhaustiva y manual de los estudios originales. Esto asegura que las decisiones de investigación no estén basadas únicamente en las conclusiones generadas por la IA.

Finalmente, en relación al flujo de trabajo de los investigadores, *Elicit* se integra en desde las primeras etapas de la revisión de la literatura hasta la toma de decisiones basada en datos. Durante la fase inicial de investigación, la plataforma permite a los

usuarios realizar búsquedas rápidas de estudios relevantes y sintetizar información clave para generar preguntas de investigación bien fundamentadas. Esto ahorra tiempo y mejora la eficiencia del proceso de revisión de literatura, ya que permite a los investigadores centrarse en los estudios más relevantes para su campo de estudio. En las etapas posteriores, *Elicit IA* ayuda a los investigadores a evaluar la calidad de los estudios previos, ofreciendo información detallada sobre sus métodos, resultados y conclusiones. Esto facilita la toma de decisiones informadas sobre el diseño de nuevos estudios y la interpretación de los resultados obtenidos.

Figura 13. Esquema funcional de *Elicit*



Fuente: Elaborado por el autor.

2.4.7. Litmaps

Litmaps es una herramienta de inteligencia artificial diseñada para ayudar a los investigadores a rastrear la evolución de la literatura científica a través de la visualización interactiva de citas y publicaciones (Litmaps, 2024). Su principal objetivo radica en permitir que los usuarios sigan el desarrollo de un campo de investigación, identificando cómo los trabajos académicos están conectados entre sí en términos de citas y referencias (Litmaps, 2024). Esto facilita la comprensión de la

estructura y el avance de un área de estudio específica, ayudando a los investigadores a descubrir nuevas publicaciones y detectar tendencias emergentes. El valor principal de *Litmaps IA* radica en su capacidad para proporcionar mapas dinámicos de la literatura científica. A diferencia de las bases de datos tradicionales, esta herramienta ofrece una visión gráfica de las conexiones entre los estudios, lo que permite a los investigadores identificar rápidamente los artículos académicos más influyentes y las áreas que están generando más atención en la comunidad científica.

La inteligencia artificial en *Litmaps* se basa en técnicas avanzadas de procesamiento de lenguaje natural (PLN) y aprendizaje automático. Estas tecnologías permiten analizar grandes volúmenes de publicaciones científicas y mapear cómo se citan entre sí, generando una representación visual de las conexiones entre trabajos académicos (Litmaps, 2024). Uno de los componentes clave de la plataforma es su capacidad para actualizar dinámicamente los mapas a medida que se publican nuevos artículos, lo que garantiza que los investigadores siempre tendrán acceso a la información más reciente (Litmaps, 2024). El uso de redes neuronales, como BERT, permite a *Litmaps* entender el contenido semántico de los artículos, mejorando la precisión en la identificación de conexiones relevantes entre trabajos. Además, este tipo de plataformas emplean algoritmos de análisis de citas para destacar aquellos estudios que han tenido un impacto significativo en el campo, ayudando a los investigadores a centrarse en las publicaciones más influyentes (Devlin et al., 2019).

Cabe acotar que las aplicaciones principales de *Litmaps* en la investigación científica incluyen la identificación de artículos clave, la visualización de la evolución de un campo de estudio y la detección de tendencias emergentes (Litmaps, 2024). La capacidad de la plataforma para generar mapas interactivos de citas permite a los investigadores ver de forma clara cómo los estudios están relacionados, lo que facilita la identificación de trabajos relevantes que podrían no haber sido detectados mediante búsquedas tradicionales en bases de datos (Litmaps, 2024). Otra aplicación importante es la posibilidad de seguir el progreso de una investigación a lo largo del tiempo. *Litmaps* notifica a los usuarios cuando se publican nuevos estudios que están conectados con los artículos académicos que ya han sido mapeados, lo que garantiza

que los investigadores estén siempre al día con los últimos desarrollos en su área de estudio.

En relación al proceso de revisión de literatura, *Litmaps* proporciona una visualización interactiva de cómo los estudios están conectados a través de citas (Litmaps, 2024). Esta capacidad de visualización facilita a los investigadores identificar rápidamente los trabajos más influyentes y las áreas donde se están generando nuevos conocimientos. A diferencia de los métodos tradicionales de búsqueda bibliográfica, que requieren que los investigadores revisen manualmente grandes volúmenes de texto, *Litmaps* organiza automáticamente la literatura relevante en mapas visuales, lo que reduce el tiempo necesario para realizar una revisión exhaustiva de la literatura (Litmaps, 2024). Además, la capacidad de la plataforma para actualizar los mapas en tiempo real permite que los investigadores se mantengan al día con las publicaciones más recientes y ajusten sus revisiones en función de los nuevos estudios que surjan. Esto es particularmente valioso en campos en rápida evolución, donde la literatura científica crece a un ritmo acelerado (Wang et al., 2020).

Cabe acotar que uno de los principales desafíos éticos asociados con el uso de *Litmaps* es el sesgo interno en los algoritmos utilizados para mapear y destacar las publicaciones más influyentes. Al igual que otras herramientas basadas en inteligencia artificial, los resultados producidos por *Litmaps* están condicionados por los datos de entrenamiento y los patrones de citación dentro de la literatura científica. Esto podría llevar a una sobre-representación de ciertos autores o enfoques metodológicos, dejando de lado investigaciones valiosas pero menos citadas (Bender et al., 2021). Otro desafío es la dependencia excesiva de los investigadores en las visualizaciones generadas por la plataforma. Si bien *Litmaps* facilita la identificación de conexiones entre estudios, es crucial que los investigadores complementen esta herramienta con un análisis crítico y manual de la literatura para garantizar que las conclusiones estén bien fundamentadas y que no se omitan trabajos relevantes.

Figura 14. Esquema funcional de *Litmaps*

Fuente: Elaborado por el autor.

Finalmente, en relación a la integración en el flujo de trabajo de los investigadores, *Litmaps* proporciona una plataforma visual que ayuda a organizar y seguir la literatura relevante a lo largo de un proyecto de investigación. Durante las primeras etapas de revisión de literatura, los usuarios pueden utilizar la plataforma para generar mapas que representen las conexiones entre estudios y autores clave en su campo de estudio. A medida que el proyecto avanza, la capacidad de *Litmaps* para actualizar dinámicamente los mapas permite a los investigadores mantenerse al día con los estudios emergentes y ajustar su revisión de manera acorde. En las etapas posteriores, la herramienta ayuda a los investigadores a organizar y presentar sus hallazgos de manera más clara, destacando las conexiones entre los estudios que

han influido en su trabajo. Además, la capacidad de compartir los mapas con otros investigadores facilita la colaboración, mejorando la calidad de la investigación colectiva.

2.4.8. Iris

Iris es una plataforma de inteligencia artificial diseñada para mejorar la manera en que los investigadores descubren, organizan y analizan la información científica (Iris, 2024). Su objetivo principal radica en proporcionar a los usuarios una herramienta que optimiza el acceso a la literatura relevante mediante técnicas avanzadas de procesamiento de lenguaje natural (PLN) y algoritmos de aprendizaje automático (Iris, 2024). A diferencia de los motores de búsqueda académica tradicionales, *Iris* ofrece una experiencia más personalizada al publicar publicaciones científicas basadas en los intereses y necesidades específicas de los investigadores, ayudando así a reducir el tiempo dedicado a la revisión de literatura y encontrar estudios críticos en etapas tempranas de la investigación.

La inteligencia artificial en *Iris* se basa en el uso de modelos de procesamiento de lenguaje natural (PLN) que permiten analizar grandes volúmenes de textos científicos, identificar temas relevantes, y extraer patrones y relaciones entre artículos. La plataforma emplea redes neuronales profundas, como BERT, que son fundamentales para la comprensión semántica de los artículos y para generar recomendaciones de publicaciones basadas en el análisis contextual de los estudios previos (Devlin et al., 2019). El uso de algoritmos de aprendizaje automático recomendados permite a *Iris* mejorar constantemente sus capacidades a medida que más usuarios interactúan con la plataforma (Iris, 2024). Esto asegura que las sugerencias de artículos académicos y referencias sean cada vez más precisas y personalizadas, lo que facilita una búsqueda eficiente y una organización más intuitiva de la literatura científica.

Cabe acotar que las aplicaciones más destacadas de *Iris* incluyen la automatización de la revisión de literatura, la identificación de tendencias emergentes en un campo de estudio y la recomendación de artículos científicos clave (Iris, 2024). La plataforma es utilizada por investigadores que buscan obtener una visión rápida y detallada de las publicaciones relevantes para su trabajo, lo que les permite

enfocarse en los estudios más influyentes y actuales sin necesidad de realizar búsquedas manuales exhaustivas. Además, *Iris* es útil para identificar vacíos en la literatura, lo que permite a los investigadores dirigir sus esfuerzos hacia áreas donde existe menor conocimiento. Esta capacidad es particularmente valiosa en campos en rápida evolución, donde la cantidad de información publicada crece constantemente y puede resultar difícil de manejar.

En relación a la mejora del proceso de revisión de literatura, *Iris* automatiza la búsqueda de artículos científicos relevantes y ofrece recomendaciones basadas en las preferencias y necesidades del investigador (Iris, 2024). En lugar de realizar búsquedas en manuales extensos, los usuarios pueden utilizar la plataforma para recibir sugerencias personalizadas de publicaciones que se alinean con su área de estudio. Este tipo de función no solo ahorra tiempo, sino que también permite una revisión más profunda y completa de la literatura al destacar estudios que podrían haber pasado desapercibidos en búsquedas tradicionales (Wang et al., 2020). Además, *Iris* actualiza automáticamente sus recomendaciones a medida que se publican nuevos estudios, lo que permite a los investigadores mantenerse al día con los avances más recientes en sus campos de interés. Esto es crucial para asegurar que las revisiones de literatura sean completas y reflejen el estado actual del conocimiento científico.

Cabe acotar que uno de los desafíos éticos más importantes asociados con *Iris* es el sesgo en los algoritmos de recomendación. Dado que la plataforma se basa en el análisis de patrones de citas y en el comportamiento de los usuarios, los algoritmos pueden favorecer ciertos enfoques o autores que ya tienen una alta visibilidad, lo que podría excluir información menos conocida pero igualmente valiosa. Este tipo de sesgo algorítmico puede afectar la diversidad en la revisión de la literatura y limitar el acceso a perspectivas alternativas en la investigación científica (Bender et al., 2021). A su vez, existe el riesgo de que los investigadores dependan excesivamente de las recomendaciones automatizadas, lo que podría llevar a una revisión superficial de la literatura. Pese a que *Iris* facilita el acceso a información relevante, es fundamental que los investigadores complementen su uso con una revisión crítica y manual de los estudios para asegurar la calidad y precisión en sus conclusiones.

Figura 15. Esquema funcional de *Iris*

Fuente: Elaborado por el autor.

Finalmente, cabe acotar que *Iris* se integra fácilmente en el flujo de trabajo de los investigadores al proporcionar una plataforma eficiente para la organización y descubrimiento de literatura relevante. Durante las primeras fases de investigación, los usuarios pueden utilizar la plataforma para recibir recomendaciones automáticas de artículos que se alineen con sus intereses y campos de estudio. A medida que el proyecto avanza, *Iris* actualiza sus sugerencias en función de nuevos estudios, lo que permite a los investigadores mantener su revisión de literatura actualizada sin necesidad de realizar búsquedas continuas. La plataforma también facilita la colaboración entre investigadores, ya que permite compartir listas de referencias y

recomendaciones personalizadas, lo que mejora la coordinación y calidad de los proyectos colaborativos. Además, la interfaz intuitiva de *Iris* permite una gestión más efectiva de la literatura, ayudando a los investigadores a organizar su bibliografía y seguir los estudios más relevantes de forma más estructurada.

A modo de resumen, a continuación se establece un cuadro comparativo con los principales aspectos de las herramientas: *Scispace, Connected Papers, Semantic Scholar, ResearchRabbit, Consensus, Elicit, Litmaps, e Iris.*

Tabla 2. Cuadro comparativo con los principales aspectos de las herramientas: *Scispace, Connected Papers, Semantic Scholar, ResearchRabbit, Consensus, Elicit, Litmaps, e Iris.*

Aspecto	*Scispace*	*Connected Papers*	*Semantic Scholar*	*ResearchRabbit*	*Consensus*	*Elicit*	*Litmaps*	*Iris*
Descripción general	Herramienta para analizar y comprender documentos científicos.	Plataforma para explorar redes de citas académicas mediante gráficos visuales.	Plataforma de búsqueda académica impulsada por IA para análisis de impacto.	Herramienta para explorar redes de artículos con recomendaciones interactivas.	Motor de búsqueda científica que proporciona respuestas basadas en evidencia a preguntas específicas.	Herramienta de IA para búsqueda y análisis de literatura científica con enfoque en datos clave.	Herramienta para la visualización de redes de citas y seguimiento de publicaciones.	Herramienta que utiliza IA para crear mapas conceptuales y sintetizar literatura científica.
Funcionalidad principal	Generación de explicaciones y resúmenes simplificados de documentos.	Descubrimiento de artículos relacionados a través de redes de citas.	Búsqueda de literatura científica con análisis de impacto y citas.	Descubrimiento iterativo de artículos mediante mapas visuales.	Búsqueda de literatura con síntesis de respuestas basadas en evidencia.	Búsqueda avanzada, extracción de datos clave y creación de tablas comparativas.	Mapeo interactivo de redes de citas académicas.	Creación de mapas conceptuales y descubrimiento de literatura relevante.
Base de datos utilizada	Bases de datos revisadas por pares.	Redes de citas y referencias bibliográficas abiertas.	Amplias bases académicas de revistas científicas de alto impacto.	Redes de coautoría y citas académicas.	Bases de datos revisadas por pares (más de 200 millones de artículos).	Bases de datos académicas extensas, como Semantic Scholar.	Redes de citas de publicaciones científicas.	Bases de datos científicas abiertas y revisadas por pares.
Capacidades de búsqueda	Búsqueda avanzada basada en texto técnico y términos clave.	Basada en similitud bibliográfica y co-citaciones.	Filtros avanzados por autor, fecha, métricas, y área de estudio.	Conexión temática entre artículos, autores y redes académicas.	Búsqueda contextual y basada en preguntas específicas.	Búsqueda avanzada con extracción específica de datos clave.	Búsqueda basada en relaciones de citas y referencias.	Búsqueda avanzada con enfoque semántico.

Aspecto	*Scispace*	*Connected Papers*	*Semantic Scholar*	*ResearchRabbit*	*Consensus*	*Elicit*	*Litmaps*	*Iris*
Visualización de resultados	Secciones clave con explicaciones detalladas.	Mapas interactivos con nodos y relaciones entre documentos.	Listas estructuradas con métricas de impacto y gráficos de citas.	Gráficos interactivos con nodos que conectan temas y publicaciones.	Resultados claros con resúmenes y enlaces a artículos.	Presenta tablas comparativas y resúmenes claros.	Mapas visuales interactivos y dinámicos.	Mapas conceptuales interactivos y estructurados.
Análisis y resumen de Artículos	Incluye resúmenes detallados y explicaciones simplificadas.	No incluye resúmenes automáticos.	Genera resúmenes automáticos de los artículos.	No ofrece resúmenes, pero muestra conexiones detalladas.	Sí, ofrece resúmenes de hallazgos clave.	Sí, incluye resúmenes y extracción de datos específicos.	No incluye resúmenes automáticos.	Sí, genera resúmenes de artículos y conceptos clave.
Interactividad con el usuario	Permite destacar y analizar secciones de documentos específicos	Explora mapas visuales interactivos.	Ofrece filtros avanzados para personalizar búsquedas.	Exploración iterativa y personalización de resultados.	Refinamiento de preguntas y resultados interactivos.	Personalización y refinamiento iterativo de resultados.	Interacción con nodos y mapas visuales.	Ajuste de mapas y consultas adicionales.
Colaboración y Compartición	No incluye colaboración directa.	Permite compartir mapas y proyectos con otros usuarios.	Los resultados pueden compartirse, pero sin colaboración directa integrada.	Facilita la compartición de colecciones y redes de artículos.	No incluye funciones de colaboración directa.	Exportación y compatibilidad con gestores bibliográficos.	Proyectos y mapas pueden compartirse fácilmente.	Permite compartir mapas conceptuales con colaboradores.
Exportación de Datos	Permite exportar documentos y notas.	Exportación de mapas visuales y referencias.	Compatible con gestores bibliográficos como Zotero o Mendeley.	Exportación de redes y colecciones en formatos estándar.	Exportación limitada a resultados específicos.	Exportación de tablas comparativas y datos clave.	Exportación de mapas y bibliografías.	Exportación de mapas y listas de referencias.
Actualización Continua	Integra nuevas publicaciones científicas regularmente.	Actualiza redes con publicaciones recientes relacionadas.	Se actualiza constantemente con nuevas publicaciones.	Monitorea y actualiza proyectos con nuevas publicaciones.	Integra nuevos estudios en su base de datos regularmente.	Sí, integra datos continuamente.	Rastrea automáticamente publicaciones relacionadas.	Incluye actualizaciones periódicas con nuevas publicaciones.
Uso Multimodal	No admite formatos multimodales.	No incluye capacidades multimodales.	Se enfoca en datos textuales y métricas académicas.	Funciona solo con texto y datos de citas.	No, se enfoca en datos textuales.	No, se centra en texto y datos tabulados.	No, se centra en redes de citas y referencias textuales.	No, se centra en texto y conceptos académicos.

Fuente: Elaborado por el autor.

2.5. Modelos generativos de imagen

2.5.1. Modelos Generales

2.5.1.1. Midjourney

Midjourney es una plataforma de inteligencia artificial generativa que permite a los usuarios desarrollar imágenes a partir de descripciones textuales, utilizando modelos de aprendizaje profundo (Midjourney, 2024). Su objetivo principal radica en ofrecer una herramienta accesible para el desarrollo de imágenes a partir de comandos escritos, lo que abre nuevas posibilidades para diseñadores, artistas y profesionales creativos que buscan generar contenido visual de manera automatizada (Midjourney, 2024). La tecnología detrás de *Midjourney* emplea redes neuronales que interpretan el lenguaje natural y lo transforman en representaciones gráficas, lo que facilita el desarrollo de imágenes personalizadas sin la necesidad de habilidades técnicas avanzadas en diseño gráfico.

Midjourney se basa en modelos generativos de redes neuronales profundas, específicamente en variantes de las *generative adversarial networks* (GANs) (traducido al español como Redes generativas adversativas), y modelos de transformación. Estos algoritmos permiten a la IA generar imágenes a partir de descripciones textuales mediante el procesamiento y la conversión de datos lingüísticos en representaciones visuales. La arquitectura funciona de manera similar a otros modelos de IA generativa, donde un modelo "generador" desarrolla imágenes y otro modelo "discriminador" evalúa la calidad de las mismas. Este proceso iterativo mejora la calidad de las imágenes generadas con el tiempo, ajustando la interpretación de las descripciones textuales para producir resultados visuales más precisos y coherentes (Piñón, 2024). Además, *Midjourney* emplea técnicas de procesamiento de lenguaje natural (PLN) para interpretar los comandos textuales. Esto implica que la IA necesita comprender tanto el contenido semántico como el contexto de las palabras proporcionadas por el usuario, a fin de generar imágenes que coincidan con las expectativas descriptivas (Midjourney, 2024).

Las aplicaciones de *Midjourney* abarcan una amplia gama de campos, desde el arte digital hasta el diseño gráfico, la publicidad y el desarrollo de prototipos visuales. En el ámbito del arte digital, *Midjourney* permite a los artistas explorar nuevas formas

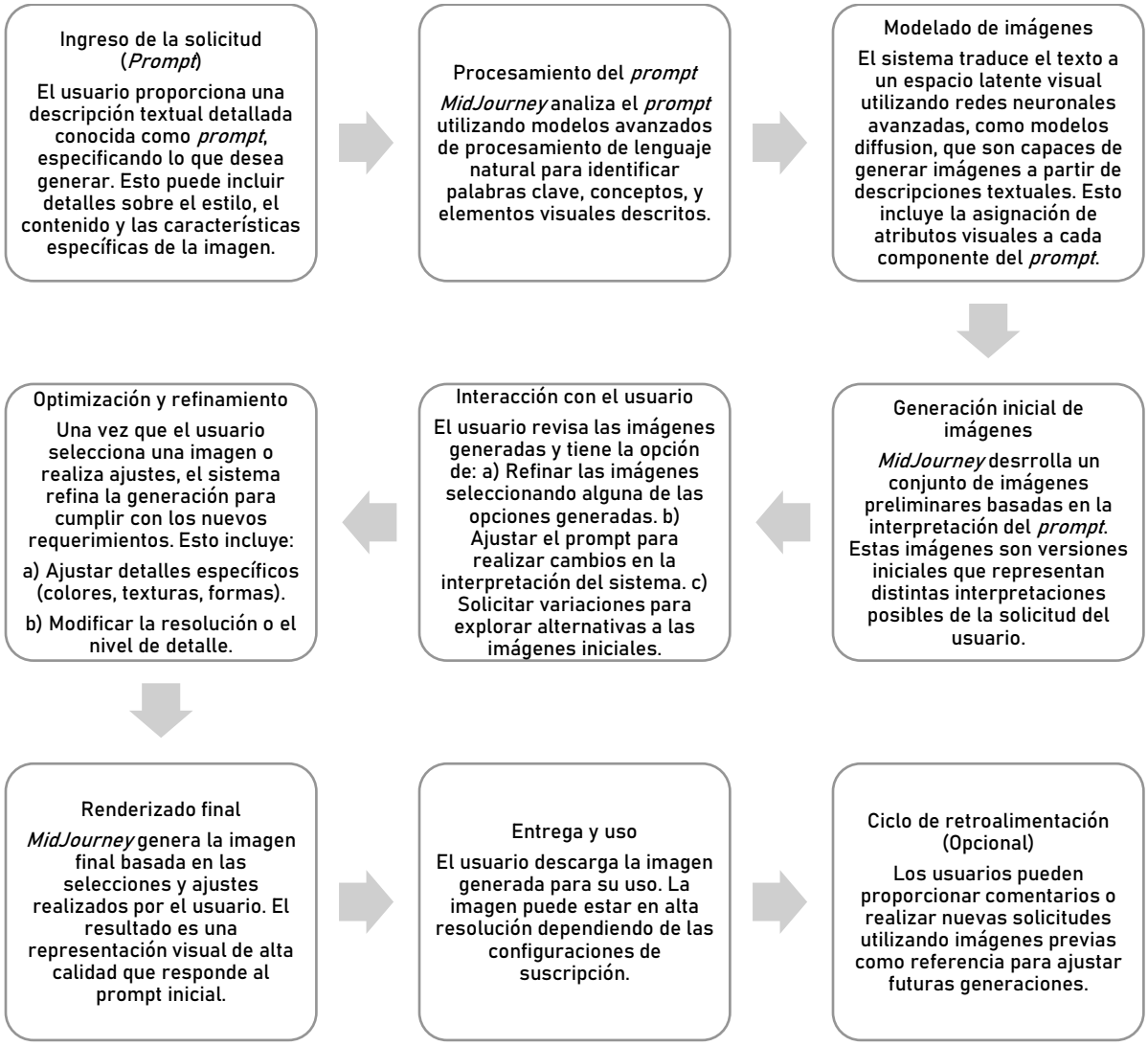
de expresión creativa, generando imágenes que combinan estilos visuales de maneras novedosas, sin las limitaciones tradicionales del dibujo o el modelado (Midjourney, 2024). Esta capacidad facilita la experimentación visual en etapas tempranas de un proyecto artístico o de diseño, permitiendo iteraciones rápidas sobre ideas conceptuales (García, 2022). En industrias como la del diseño gráfico, *Midjourney* es útil para generar representaciones visuales a partir de ideas abstractas, lo que puede ser particularmente valioso en el desarrollo de material publicitario, campañas visuales y productos gráficos. Asimismo, el uso de la IA para generar imágenes puede acelerar el proceso de desarrollo de contenido visual en áreas como el marketing digital y la producción de medios, donde la rapidez y la originalidad son esenciales.

Comparado con otras herramientas de generación de imágenes como *DALL-E* o *Artbreeder*, *Midjourney* se destaca por su enfoque en la accesibilidad y la facilidad de uso para un público general, sin requerir un conocimiento técnico avanzado. Mientras que plataformas como *DALL-E* están diseñadas con arquitecturas más complejas para generar imágenes de alta calidad a partir de entradas textuales detalladas, *Midjourney* equilibra la calidad visual con una interfaz amigable que permite a los usuarios experimentar con descripciones más amplias o abstractas (Midjourney, 2024). Además, *Midjourney* tiende a generar imágenes con un estilo visual más artístico o estilizado, en comparación con la aproximación más realista de *DALL-E*. Otra diferencia clave radica en la capacidad de personalización y control que ofrecen las plataformas. Mientras que *DALL-E* permite un control preciso sobre ciertos aspectos de la imagen generada, como la posición de los objetos o el estilo de iluminación, *Midjourney* proporciona un enfoque más intuitivo y espontáneo, enfocado en el rápido desarrollo y la iteración creativa.

Cabe acotar que el uso de *Midjourney* y otras plataformas de inteligencia artificial generativa plantea varios desafíos éticos y sociales. Uno de los principales problemas es la cuestión de los derechos de autor y la propiedad intelectual. Dado que las imágenes generadas por *Midjourney* se desarrollan a partir de descripciones textuales y están basadas en datos de entrenamiento derivados de imágenes existentes, es posible que algunas creaciones visuales infrinjan derechos de autor o

reproduzcan estilos visuales sin el debido reconocimiento de los autores originales (Ordelín, 2023). Otro desafío importante es el riesgo de sesgo algorítmico en la generación de imágenes. Como los modelos de inteligencia artificial se entrenan en grandes conjuntos de datos de imágenes preexistentes, existe el riesgo de que los sesgos presentes en estos datos se reproduzcan en las imágenes generadas. Esto puede llevar a la representación desproporcionada de ciertos estereotipos o a la falta de diversidad en las imágenes desarrolladas por la IA (Castillo & Beltrán, 2024).

Figura 16. Esquema funcional de *Midjourney*



Fuente: Elaborado por el autor.

Finalmente, cabe acotar que *Midjourney* se integra de manera eficiente en el flujo de trabajo de diseñadores, desarrolladores de contenido y artistas al proporcionar una herramienta rápida y accesible para la generación de imágenes personalizadas

(Midjourney, 2024). En las primeras etapas de un proyecto creativo, los diseñadores pueden usar *Midjourney* para generar ideas visuales preliminares o desarrollar prototipos de conceptos antes de pasar a etapas de refinamiento más detalladas. Esto es especialmente útil en procesos iterativos, donde el rápido desarrollo de imágenes permite explorar múltiples direcciones visuales en poco tiempo. Además, *Midjourney* permite a los usuarios ajustar los resultados generados, lo que facilita la colaboración entre diseñadores y clientes. Los diseñadores pueden compartir iteraciones visuales con sus clientes y recibir retroalimentación rápida, lo que acelera el ciclo de desarrollo y asegura que las expectativas del cliente se cumplan de manera eficiente.

2.5.1.2. Krea

Krea es una plataforma de inteligencia artificial generativa enfocada en el desarrollo de imágenes a partir de descripciones textuales, utilizando modelos avanzados de aprendizaje profundo (Krea, 2024). Su propósito principal radica en facilitar el desarrollo de contenido visual automatizado, permitiendo a los usuarios generar imágenes originales a partir de comandos escritos (Krea, 2024). Al emplear modelos de redes neuronales, *Krea* ofrece una herramienta versátil que es útil para diseñadores, creativos y profesionales en diversos campos que requieren prototipos visuales o conceptos gráficos rápidamente. Su tecnología está diseñada para interpretar descripciones textuales y convertirlas en representaciones gráficas precisas. Cabe acotar que *Krea* emplea una arquitectura de redes neuronales profundas, particularmente modelos generativos como los *Generative adversarial networks* (GANs). El funcionamiento de estos modelos implica la interacción de dos redes neuronales: una red generadora que desarrolla imágenes a partir de los comandos textuales y una red discriminadora que evalúa la calidad de las imágenes generadas. Este proceso se refina de manera iterativa, permitiendo que el modelo mejore continuamente en la generación de imágenes a medida que aprende a interpretar mejor las descripciones textuales proporcionadas. Además de las GANs, *Krea* también utiliza técnicas de procesamiento de lenguaje natural (PLN) para comprender el significado de las entradas textuales. El modelo traduce las descripciones en componentes visuales y estilos, generando imágenes coherentes con el contexto y la intención del usuario (Krea, 2024). Este enfoque híbrido de PLN

y redes neuronales profundas le permite a la IA generar contenido visual adaptado a diversas necesidades creativas.

Cabe acotar Las aplicaciones de *Krea* son diversas y abarcan varios campos creativos y profesionales. En el diseño gráfico y la publicidad, *Krea* permite generar conceptos visuales de manera rápida y económica, lo que es particularmente útil para el desarrollo de campañas publicitarias y la visualización de ideas en las etapas tempranas de un proyecto. En el ámbito del diseño de productos, la plataforma ayuda a los diseñadores a visualizar prototipos o conceptos preliminares sin necesidad de herramientas tradicionales de modelado. En la actualidad, *Krea* también se utiliza en la industria del entretenimiento, específicamente en el desarrollo de arte conceptual para películas, videojuegos y medios digitales. La capacidad de generar imágenes a partir de descripciones detalladas permite a los desarrolladores explorar múltiples estilos visuales y experimentar con enfoques artísticos sin las limitaciones del diseño manual. Esto no solo acelera el proceso de desarrollo, sino que también abre nuevas posibilidades creativas en términos de innovación visual.

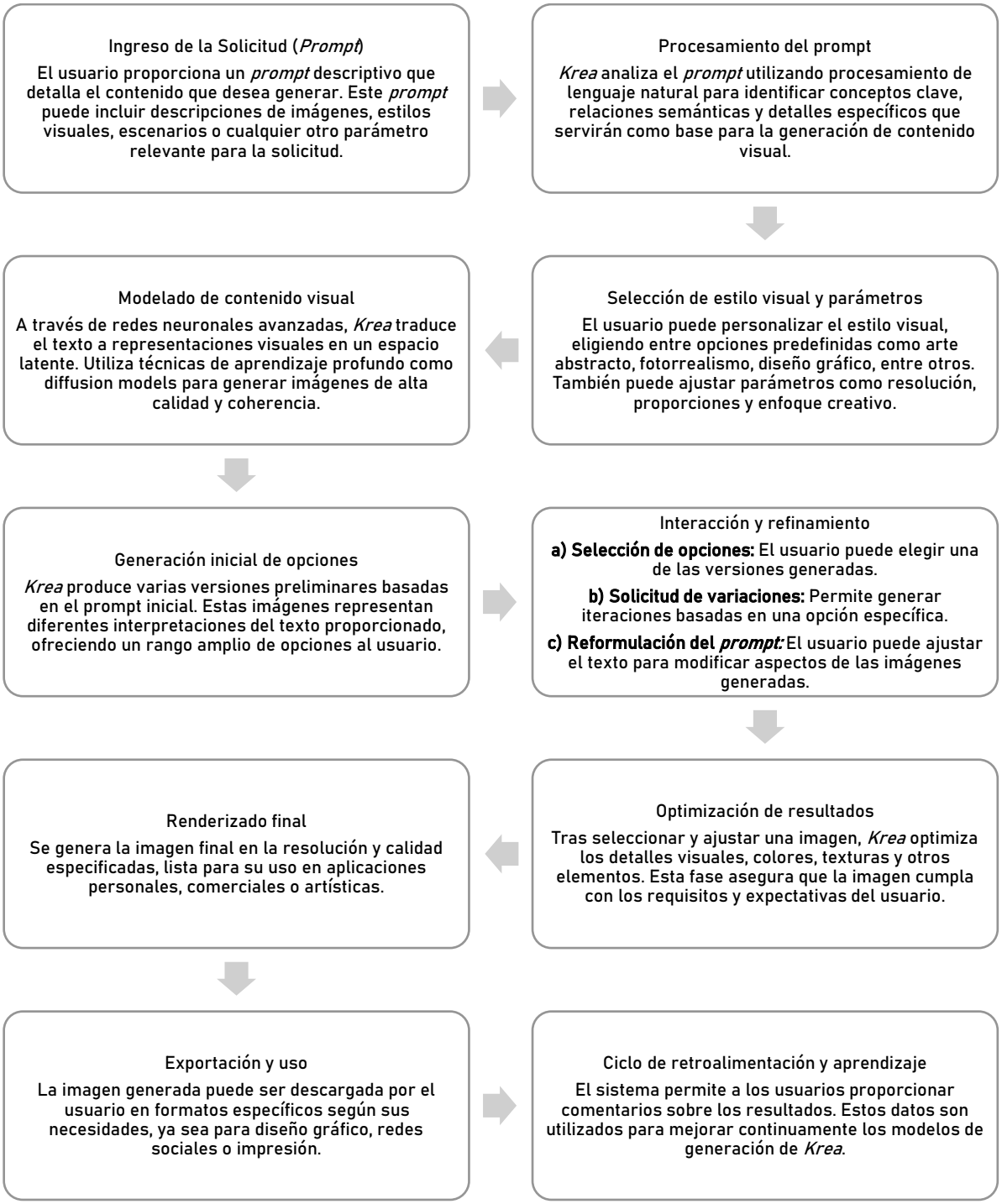
En relación a la similitud con plataformas como *Midjourney* y *DALL-E, Krea* tiene la capacidad de generar imágenes a partir de descripciones textuales. Sin embargo, una diferencia clave radica en su enfoque en la personalización y la accesibilidad para una variedad más amplia de usuarios. Mientras que plataformas como *DALL-E* se centran en la precisión técnica y el desarrollo de imágenes altamente detalladas, *Krea* ofrece una interfaz más flexible que permite a los usuarios experimentar con ideas visuales rápidamente sin requerir comandos textuales altamente específicos. Además, *Krea IA* está diseñada para ser utilizada por profesionales creativos en una variedad de industrias, ofreciendo características que facilitan la integración del contenido visual generado en proyectos de diseño gráfico, marketing o entretenimiento. En comparación con *Midjourney* , que tiene un enfoque más artístico y estilizado, *Krea* busca equilibrar la calidad visual con la rapidez en la generación de imágenes, lo que es útil en escenarios donde el tiempo es un factor clave (Krea, 2024).

Al igual que otras plataformas de inteligencia artificial generativa, *Krea* plantea desafíos éticos relevantes. Uno de los principales problemas es la cuestión de los

derechos de autor y la propiedad intelectual. Dado que las imágenes generadas por *Krea* se basan en un modelo entrenado con grandes volúmenes de datos visuales, existe la posibilidad de que las imágenes generadas sean similares a obras preexistentes, lo que podría llevar a disputas sobre la propiedad de los contenidos desarrollados (Ordelín, 2023). Otro desafío es el riesgo de sesgo algorítmico en las imágenes generadas. Los modelos de inteligencia artificial se entrenan en conjuntos de datos que pueden contener sesgos, lo que puede llevar a que las imágenes generadas reproduzcan estereotipos o limitaciones culturales. Es fundamental que las plataformas como *Krea IA* aborden estos problemas mediante la revisión y el ajuste de los datos de entrenamiento para garantizar que las imágenes generadas representen una diversidad adecuada y no perpetúen prejuicios existentes (Bender et al., 2021).

Finalmente, en relación a la integración al flujo de trabajo de los profesionales creativos, *Krea* lo hace de una manera sencilla, al proporcionar una plataforma accesible para la generación rápida de prototipos visuales y contenido gráfico. En las primeras etapas de un proyecto, los diseñadores pueden utilizar *Krea* para generar imágenes conceptuales a partir de descripciones iniciales, lo que facilita la exploración de múltiples enfoques visuales sin la necesidad de herramientas de diseño más complejas. Esto permite una mayor flexibilidad en la toma de decisiones durante la fase creativa y acelerar el ciclo de desarrollo de productos o campañas visuales. Además, *Krea* permite a los usuarios realizar ajustes en las imágenes generadas, lo que facilita la colaboración entre diseñadores, clientes y otros actores en el proceso creativo. La capacidad de iterar rápidamente sobre ideas visuales ayuda a mejorar la comunicación y garantiza que los proyectos se mantengan alineados con las expectativas del cliente o del equipo creativo.

Figura 17. Esquema funcional de *Krea*



Fuente: Elaborado por el autor.

2.5.1.3. Leonardo

Leonardo es una plataforma de inteligencia artificial enfocada en la generación de contenido visual, diseñada para permitir a los usuarios desarrollar imágenes y gráficos a partir de descripciones textuales (Leonardo, 2024). La finalidad de *Leonardo* radica en simplificar el proceso de diseño y desarrollo visual,

proporcionando a diseñadores, artistas y creativos una herramienta de apoyo que automatiza la generación de imágenes (Leonardo, 2024). Al basarse en técnicas avanzadas de aprendizaje profundo, la plataforma permite el desarrollo de contenido original y personalizado, facilitando un enfoque práctico para la experimentación visual y el desarrollo de conceptos gráficos sin la necesidad de habilidades técnicas avanzadas en diseño.

Cabe acotar que la tecnología detrás de *Leonardo* se basa en modelos de redes neuronales profundas y algoritmos generativos, específicamente las *Generative adversarial networks* (GANs). Estos modelos operan a través de una red generadora que desarrolla imágenes y una red discriminadora que evalúa la calidad de las imágenes generadas, refinando continuamente el proceso mediante un ciclo de retroalimentación. Esta estructura permite que el modelo mejore progresivamente su capacidad para generar imágenes que se ajustan a las descripciones textuales proporcionadas por los usuarios. Además, *Leonardo* emplea procesamiento de lenguaje natural (PLN) para interpretar los comandos textuales. El modelo analiza el contenido semántico y el contexto de las entradas textuales para generar imágenes que reflejen con precisión las expectativas del usuario. Este enfoque híbrido, que combina redes generativas y PLN, es fundamental para la generación precisa de contenido visual en la plataforma (Leonardo, 2024).

En relación a sus aplicaciones prácticas, *Leonardo* se ajusta a una variedad de campos creativos, como el diseño gráfico, la publicidad, el desarrollo de productos y el arte digital (Leonardo, 2024). En el diseño gráfico, facilita el desarrollo de conceptos visuales, permitiendo a los diseñadores generar imágenes que pueden utilizarse como prototipos iniciales o como inspiración en campañas visuales y *branding*. En publicidad, *Leonardo* permite la generación rápida de contenido visual, reduciendo el tiempo y los recursos necesarios para generar representaciones visuales de ideas abstractas o temáticas complejas. Otra aplicación relevante es en la industria del entretenimiento, donde *Leonardo* puede ser utilizado para desarrollar arte conceptual para películas, videojuegos y medios digitales. Esta capacidad de generar imágenes detalladas y estilizadas a partir de descripciones textuales permite a los desarrolladores experimentar con diversos estilos y enfoques visuales,

proporcionando una herramienta adicional para la exploración estética y conceptual en proyectos creativos.

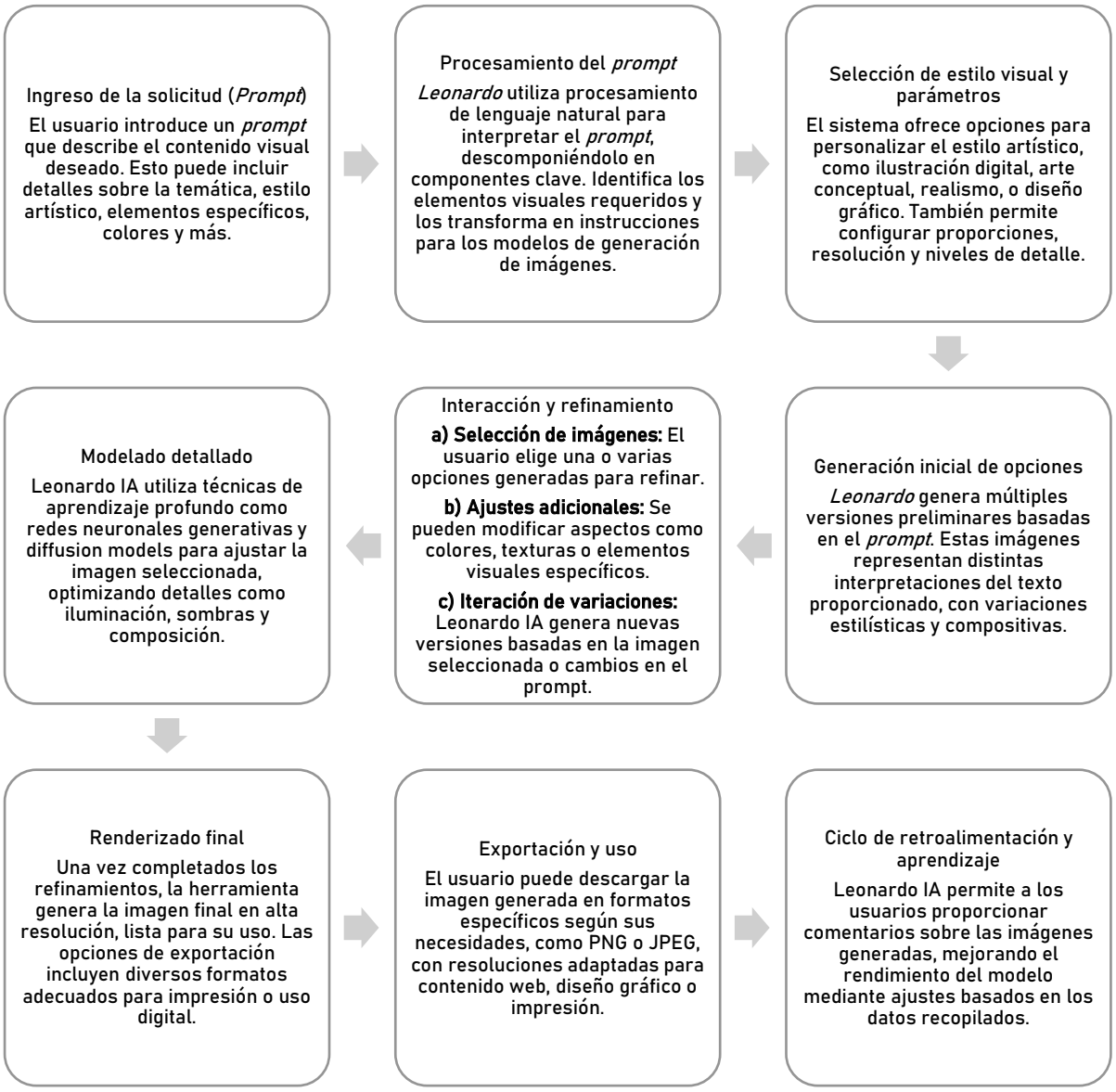
Otro tema a destacar radica en que *Leonardo* se diferencia de otras plataformas como *DALL-E* o *Midjourney* en varios aspectos, principalmente en su enfoque y accesibilidad. Mientras que *DALL-E* emplea una arquitectura compleja diseñada para generar imágenes de alta precisión y detalles específicos, *Leonardo* busca equilibrar la facilidad de uso y la personalización de las imágenes. Esto lo hace adecuado para usuarios que buscan desarrollar contenido visual rápidamente sin requerir descripciones extremadamente detalladas. A diferencia de *Midjourney* , que tiene un enfoque más artístico y estilizado, *Leonardo* está diseñado para ser utilizado en aplicaciones prácticas y comerciales, ofreciendo una gama de opciones que permiten ajustar la apariencia de las imágenes generadas en función de las necesidades del proyecto. Este enfoque práctico convierte a *Leonardo* en una herramienta versátil para diseñadores que requieren una integración rápida y fluida en sus flujos de trabajo.

Al igual que otras plataformas de inteligencia artificial generativa, *Leonardo* plantea desafíos éticos, especialmente en relación con los derechos de autor y el sesgo algorítmico. Un gran problema radica en el potencial de infracción de derechos de autor, ya que las imágenes generadas pueden basarse en datos de entrenamiento que contienen elementos visuales preexistentes. Esto genera la posibilidad de que las imágenes generadas por *Leonardo* reflejen o repliquen obras con propiedad intelectual de terceros, lo que podría llevar a disputas legales. El sesgo algorítmico es otro desafío importante, ya que los modelos de inteligencia artificial como los que utilizan *Leonardo* se entrenan en grandes conjuntos de datos que pueden contener sesgos inherentes. Esto puede llevar a que las imágenes generadas reflejen patrones o estereotipos culturales preexistentes. Para mitigar estos riesgos, es fundamental que los desarrolladores revisen y ajusten continuamente los datos de entrenamiento para garantizar cierta objetividad en los resultados generados.

Finalmente, cabe acotar que *Leonardo* se integra de manera eficaz en el flujo de trabajo de profesionales del diseño, facilitando la generación rápida de contenido visual durante las primeras etapas de desarrollo de un proyecto. En la fase de

conceptualización, los diseñadores pueden utilizar *Leonardo* para explorar diferentes opciones visuales basadas en descripciones textuales, lo que agiliza el proceso de ideación y permite experimentar con múltiples estilos y enfoques. La plataforma también permite la colaboración en tiempo real, lo que facilita la comunicación entre diseñadores, clientes y otros colaboradores en el proyecto. Los usuarios pueden compartir rápidamente los resultados generados, recibir retroalimentación y realizar ajustes en función de las necesidades del cliente o del equipo, lo que ayuda a mejorar la calidad y efectividad de los proyectos creativos.

Figura 18. Esquema funcional de *Leonardo*



Fuente: Elaborado por el autor.

2.5.1.4. Stable Diffusion

Stable Diffusion es un modelo de inteligencia artificial basado en técnicas de difusión profunda diseñado para generar imágenes a partir de descripciones textuales (Stable Diffusion, 2024). Su objetivo principal radica en proporcionar una herramienta accesible para el desarrollo automatizada de contenido visual a partir de indicaciones textuales, permitiendo a artistas, diseñadores y desarrolladores producir imágenes de alta calidad sin necesidad de intervención manual en el proceso creativo (Stable Diffusion, 2024). A diferencia de otros modelos generativos, *Stable Diffusion* emplea un enfoque de proceso de difusión que permite generar imágenes complejas, utilizando el aprendizaje profundo y redes neuronales avanzadas. El modelo de difusión de *Stable Diffusion* se basa en una serie de transformaciones sucesivas aplicadas a un conjunto inicial de ruido aleatorio para aproximar una imagen en etapas (Stable Diffusion, 2024). El proceso de generación se desarrolla en una secuencia de pasos que simulan una "difusión inversa", lo que significa que el modelo aprende a revertir el ruido hacia una representación visual coherente (Stable Diffusion, 2024). Este tipo de método permite al modelo generar imágenes de manera gradual, aplicando transformaciones cada vez más precisas hasta que se alcanza una imagen final. Cabe acotar que este modelo de inteligencia artificial se diferencia de los *Generative Adversarial Networks* (GANs), debido a su capacidad para controlar el nivel de detalle y la coherencia en los resultados sin requerir una red discriminadora. Además, a diferencia de los GANs, el modelo de difusión no depende de una arquitectura competitiva, sino de un proceso iterativo de refinamiento.

En relación a las aplicaciones de *Stable Diffusion*, estas abarcan desde el desarrollo de arte digital hasta el diseño gráfico, la visualización arquitectónica y la generación de prototipos en diseño de productos (Stable Diffusion, 2024). En el ámbito del arte digital, *Stable Diffusion* permite a los artistas experimentar con estilos visuales diversos, generando imágenes que reflejan temas o conceptos abstractos con precisión (Stable Diffusion, 2024). Esta capacidad es particularmente útil en la ilustración conceptual y el desarrollo de paisajes visuales complejos. En diseño gráfico, *Stable Diffusion* se utiliza para desarrollar material visual que puede ser incorporado en campañas publicitarias o utilizado para prototipos. Además, su

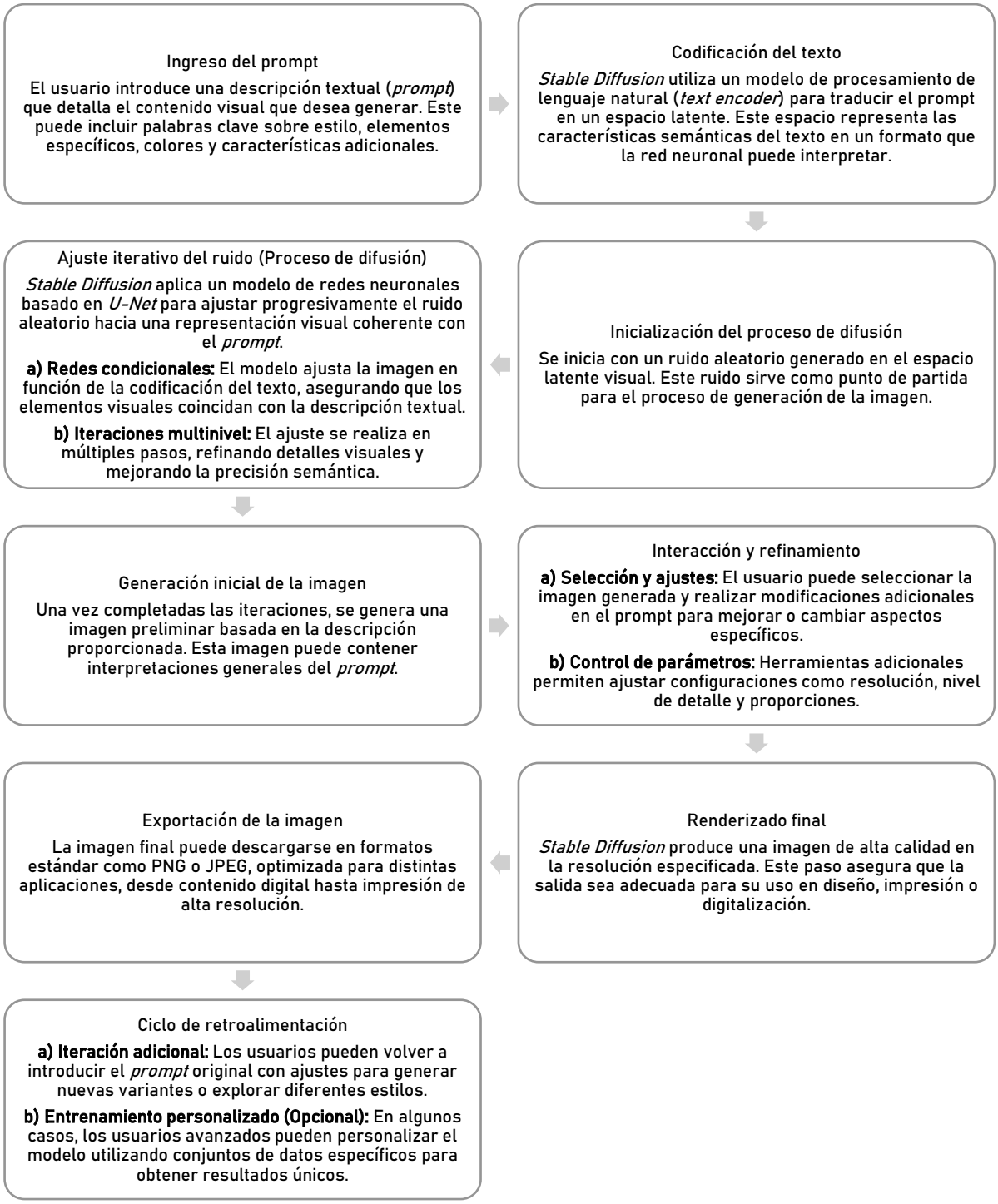
capacidad para generar imágenes detalladas y estilizadas lo hace una herramienta valiosa en la previsualización de escenas arquitectónicas o productos en fase de diseño, ya que permite a los diseñadores probar conceptos visuales antes de invertir en el modelado físico o en etapas avanzadas de producción.

Uno de los desafíos éticos más relevantes en el uso de *Stable Diffusion* es la cuestión de los derechos de autor y la propiedad intelectual. Las imágenes generadas por modelos de inteligencia artificial pueden basarse en datos de entrenamiento que contienen imágenes protegidas por derechos de autor, lo que plantea el riesgo de replicar o emular obras preexistentes sin el permiso del desarrollador original. Este desafío ético requiere atención cuidadosa en la selección de los datos de entrenamiento y en la implementación de regulaciones que protejan la propiedad intelectual. A nivel técnico, el modelo de difusión utilizado por *Stable Diffusion* presenta desafíos en términos de los recursos computacionales necesarios para entrenar y ejecutar el modelo. La generación de imágenes detalladas y coherentes requiere una gran cantidad de procesamiento, lo que puede limitar el acceso a esta tecnología a usuarios con recursos computacionales avanzados. Este aspecto subraya la importancia de optimizar el modelo para mejorar su accesibilidad sin comprometer la calidad de las imágenes generadas.

Cabe acotar que *Stable Diffusion* se integra de manera efectiva en el flujo de trabajo de profesionales en diseño y producción de contenido visual al facilitar el desarrollo de prototipos visuales y exploraciones conceptuales. Durante las primeras etapas de un proyecto, los diseñadores pueden utilizar *Stable Diffusion* para generar imágenes conceptuales a partir de descripciones, lo que les permite visualizar diferentes enfoques visuales antes de avanzar hacia el desarrollo detallado (Stable Diffusion, 2024). Esta funcionalidad es particularmente útil en contextos de diseño iterativo, donde los creativos necesitan experimentar con múltiples estilos y temas antes de seleccionar una dirección final. La plataforma también facilita la colaboración entre diseñadores, clientes y otros miembros del equipo creativo, ya que los resultados generados pueden compartirse y discutirse rápidamente. Esto permite una comunicación fluida y asegura que el contenido visual final esté alineado con las

expectativas del cliente y del equipo, mejorando la eficiencia en la toma de decisiones visuales en proyectos creativos complejos.

Figura 19. Esquema funcional de *Stable Diffusion*



Fuente: Elaborado por el autor.

Es relevante mencionar que *Stable Diffusion* presenta varias ventajas en comparación con otros modelos de inteligencia artificial generativa, como los GANs

y los modelos basados en transformadores. Una de las principales ventajas es su capacidad para controlar el nivel de detalle en las imágenes generadas sin la necesidad de una red discriminadora, lo que permite un enfoque de generación más estable y predecible. Además, el enfoque de difusión utilizado en *Stable Diffusion* permite generar imágenes de alta calidad sin sufrir los problemas de inestabilidad que suelen afectar a los GANs. Sin embargo, *Stable Diffusion* también tiene limitaciones, especialmente en términos de los requisitos de recursos computacionales. La naturaleza iterativa del modelo de difusión implica que la generación de imágenes requiere múltiples etapas de procesamiento, lo que puede resultar en tiempos de ejecución prolongados en comparación con otros modelos generativos. Este aspecto limita su accesibilidad y uso en dispositivos con menos capacidad de procesamiento, destacando la necesidad de una optimización continua del modelo para mejorar su eficiencia y ampliar su aplicación.

A modo de resumen, a continuación se establece un cuadro comparativo con los principales aspectos de las herramientas: *Midjourney, Krea, Leonardo y Stable Diffusion*.

Tabla 3. Cuadro comparativo con los principales aspectos de las herramientas: *Midjourney, Krea, Leonardo, Stable Diffusion*

Aspecto	*Midjourney*	*Krea IA*	*Leonardo IA*	*Stable Diffusion*
Descripción general	Plataforma para la generación de imágenes a partir de descripciones textuales con estilos artísticos personalizados.	Herramienta de generación creativa enfocada en estilos personalizados y variaciones.	Plataforma de diseño visual para generar imágenes de alta calidad y estilos detallados.	Modelo de código abierto para la generación de imágenes basado en el enfoque de difusión.
Funcionalidad principal	Generación de imágenes a partir de texto con iteraciones interactivas.	Desarrollo de contenido visual con soporte para múltiples estilos y ajustes personalizados.	Generación de contenido visual con refinamiento iterativo y alta resolución.	Generación de imágenes a partir de descripciones textuales utilizando modelos de difusión.
Base de datos utilizada	Datos entrenados con enfoques de aprendizaje profundo y redes neuronales generativas.	Modelos de IA entrenados con bases de datos gráficas y artísticas amplias.	Amplias bases de datos visuales y gráficos, con personalización basada en el estilo elegido.	Datos de entrenamiento basados en imágenes públicas y accesibles.

Aspecto	*Midjourney*	*Krea IA*	*Leonardo IA*	*Stable Diffusion*
Capacidades de búsqueda	Basada en texto con palabras clave y descripciones detalladas.	Búsqueda avanzada por texto y configuración de estilos visuales.	Personalización basada en el texto y en parámetros como iluminación, color y textura.	No incluye búsqueda, funciona únicamente mediante *prompt* textual.
Visualización de resultados	Genera imágenes preliminares con opciones iterativas para refinamiento.	Múltiples versiones iniciales para elegir y refinar.	Opciones visuales iterativas con ajustes basados en las selecciones del usuario.	Generación directa de imágenes sin visualización intermedia.
Análisis y resumen de artículos	No aplica, se centra en la generación de imágenes.	No aplica, se centra en la generación de imágenes.	No aplica, se centra en la generación de imágenes.	No aplica, se centra en la generación de imágenes.
Interactividad con el usuario	Permite iteraciones para modificar estilos, detalles y composiciones de las imágenes.	Ofrece refinamientos iterativos y variaciones estilísticas.	Permite ajustes iterativos de estilos, colores y detalles visuales.	Interacción limitada a ajustes mediante reformulación del *prompt*.
Colaboración y compartición	Soporte mediante comunidades activas, como en *Discord*, para compartir resultados y colaboraciones.	No incluye funcionalidades específicas para colaboración.	Facilita el intercambio de imágenes y proyectos con colaboradores.	Al ser de código abierto, permite proyectos colaborativos de desarrollo técnico.
Exportación de datos	Imágenes en alta resolución listas para impresión o contenido digital.	Compatible con formatos estándar como PNG y JPEG.	Exportación en formatos de alta resolución para impresión o uso digital.	Exportación en diversos formatos estándar (PNG, JPEG) según configuración del usuario.
Actualización continua	Actualizaciones regulares para incluir nuevos estilos y capacidades.	Actualizaciones basadas en estilos y mejoras de modelos.	Incluye nuevas funciones y ajustes según las tendencias creativas.	Actualizaciones disponibles mediante la comunidad y desarrolladores externos.
Uso multimodal	Compatible con texto y ajustes visuales mediante descripciones iterativas.	Centrado en texto para la generación visual personalizada.	Permite ajuste de parámetros textuales y visuales en combinación.	Basado exclusivamente en *prompt* textual.

Fuente: Elaborado por el autor.

2.5.2. Arquitectura y diseño

2.5.2.1. D5-Hi

D5-Hi es una herramienta avanzada de visualización que utiliza inteligencia artificial para desarrollar y renderizar imágenes en tiempo real, orientada principalmente a profesionales del diseño, la arquitectura y la producción de contenido visual (D5-Hi, 2024). Su objetivo principal radica en mejorar la precisión y la eficiencia en el desarrollo de representaciones visuales realistas, facilitando el proceso de diseño y

permitiendo a los usuarios experimentar en tiempo real con distintos parámetros, como iluminación, materiales y texturas (D5-Hi, 2024). *D5-Hi* combina técnicas avanzadas de renderizado con el uso de redes neuronales para optimizar el proceso y ofrecer resultados de alta calidad sin la necesidad de recursos computacionales excesivos.

El motor de renderizado de *D5-Hi* utiliza técnicas de aprendizaje profundo y procesamiento de gráficos para generar imágenes realistas de manera eficiente. Su arquitectura combina el trazado de rayos (*ray tracing*) en tiempo real con modelos de redes neuronales convolucionales y algoritmos de inteligencia artificial, lo cual permite a la plataforma generar imágenes de alta calidad sin comprometer el rendimiento. El trazado de rayos simula cómo la luz interactúa con diferentes superficies, estableciendo efectos de iluminación y sombras detalladas. Las redes neuronales, por su parte, optimizan la velocidad de este proceso al predecir y corregir ciertas características visuales en tiempo real, lo que reduce la necesidad de realizar cálculos exhaustivos en cada cuadro (D5-Hi, 2024). La tecnología de *D5-Hi* también incluye técnicas de renderizado basado en física, que permiten simular con precisión las propiedades físicas de materiales como metal, vidrio y madera. Esto contribuye al desarrollo de visualizaciones que no solo son estéticamente detalladas, sino también esencialmente precisas, lo cual es fundamental en campos como la arquitectura y el diseño industrial.

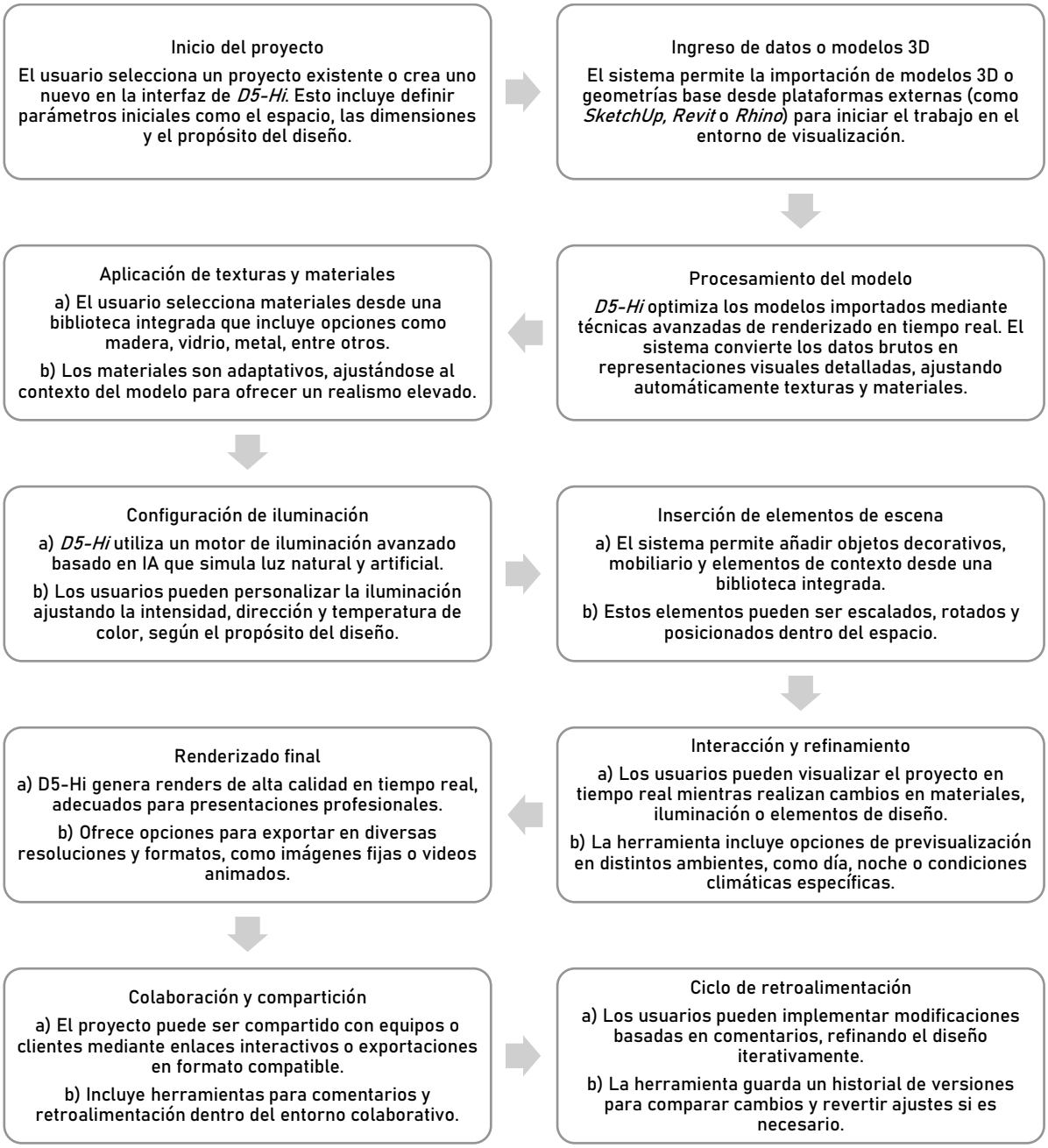
Las aplicaciones de *D5-Hi* en diseño y arquitectura abarcan el desarrollo de renders de alta calidad para proyectos arquitectónicos, diseño de interiores, visualización de productos y desarrollo de prototipos. En arquitectura, *D5-Hi* permite a los diseñadores y arquitectos generar visualizaciones detalladas de edificios y espacios, facilitando el análisis y la comunicación de ideas conceptuales. La capacidad de ajustar parámetros en tiempo real permite que los profesionales experimenten con diferentes configuraciones de iluminación y materiales, lo cual es esencial para evaluar cómo un diseño podría adaptarse a distintas condiciones. En el diseño de productos, *D5-Hi* se utiliza para generar representaciones visuales necesarias de prototipos, permitiendo a los diseñadores ver y modificar aspectos de sus generaciones antes de llevarlas a producción. Esta capacidad de previsualización en

tiempo real ayuda a reducir costos y tiempos de desarrollo, ya que minimiza la necesidad de desarrollar modelos físicos preliminares.

Cabe acotar que el uso de *D5-Hi* presenta varios desafíos éticos y técnicos. A nivel ético, la capacidad para generar visualizaciones realistas plantea el riesgo de representar imágenes que pueden ser interpretadas como reales, lo que podría ser utilizado para manipular la percepción de un proyecto o producto de manera intencional. En contextos comerciales o de marketing, existe la posibilidad de que las representaciones generadas con *D5-Hi* sean percibidas como fieles a la realidad, cuando en realidad son modelos visuales ideales o poco representativos de la versión final del diseño. En términos técnicos, *D5-Hi* enfrenta el desafío de mantener un equilibrio entre la calidad visual y la eficiencia en el procesamiento. A pesar de los avances en técnicas de IA para el renderizado, el procesamiento en tiempo real de gráficos complejos sigue siendo una tarea intensiva en recursos. Esto limita su aplicación en dispositivos con baja capacidad computacional y plantea la necesidad de desarrollar versiones optimizadas que mantengan la calidad de los resultados sin requerir una infraestructura de hardware avanzada.

En relación al flujo de trabajo de diseño y arquitectura, *D5-Hi* se integra principalmente durante las etapas de previsualización y evaluación de propuestas visuales. Al generar imágenes en tiempo real, la plataforma permite a los diseñadores probar rápidamente diferentes configuraciones visuales, lo que facilita la toma de decisiones y mejora la comunicación con clientes y equipos de trabajo. Esta capacidad de adaptación es especialmente útil en entornos colaborativos, donde las propuestas visuales deben revisarse y ajustarse con frecuencia para alinearse con los requisitos del cliente (Iglesias Yáñez, 2021). Además, *D5-Hi* puede conectarse con otras herramientas de diseño y modelado 3D, permitiendo que los modelos generados en plataformas como *AutoCAD* o *SketchUp* se importen para ser visualizados con los efectos de iluminación y materiales de *D5-Hi* . Esto ayuda a los profesionales a evaluar sus diseños en un entorno más realista sin necesidad de cambiar de plataforma, mejorando así la eficiencia del proceso de diseño.

Figura 20. Esquema funcional de *D5-Hi*



Fuente: Elaborado por el autor.

Cabe acotar que *D5-Hi* ofrece varias ventajas en comparación con otras herramientas de renderizado, como la capacidad de generar visualizaciones en tiempo real y el uso de inteligencia artificial para optimizar el proceso (D5-Hi, 2024). Una de las ventajas más notables es su enfoque en el procesamiento de alta calidad visual sin requerir una infraestructura de *hardware* intensiva, lo cual permite a los diseñadores trabajar en proyectos visualmente complejos de manera ágil. Sin embargo, una limitación de *D5-Hi* es que su desempeño óptimo aún depende de

hardware avanzado para mantener una calidad de imagen superior en configuraciones gráficas exigentes. A diferencia de plataformas más tradicionales que se centran en el renderizado fotorrealista a través de métodos de procesamiento fuera de línea, *D5-Hi* se enfoca en el desarrollo de imágenes en tiempo real, lo que lo hace especialmente útil para flujos de trabajo que requieren una revisión continua y ajustes rápidos. Sin embargo, su capacidad para alcanzar niveles de detalle extremadamente altos puede estar limitada en comparación con herramientas de renderizado fotográfico, que aún son preferidas en casos donde la precisión visual es una prioridad.

2.5.2.2. Prome-AI

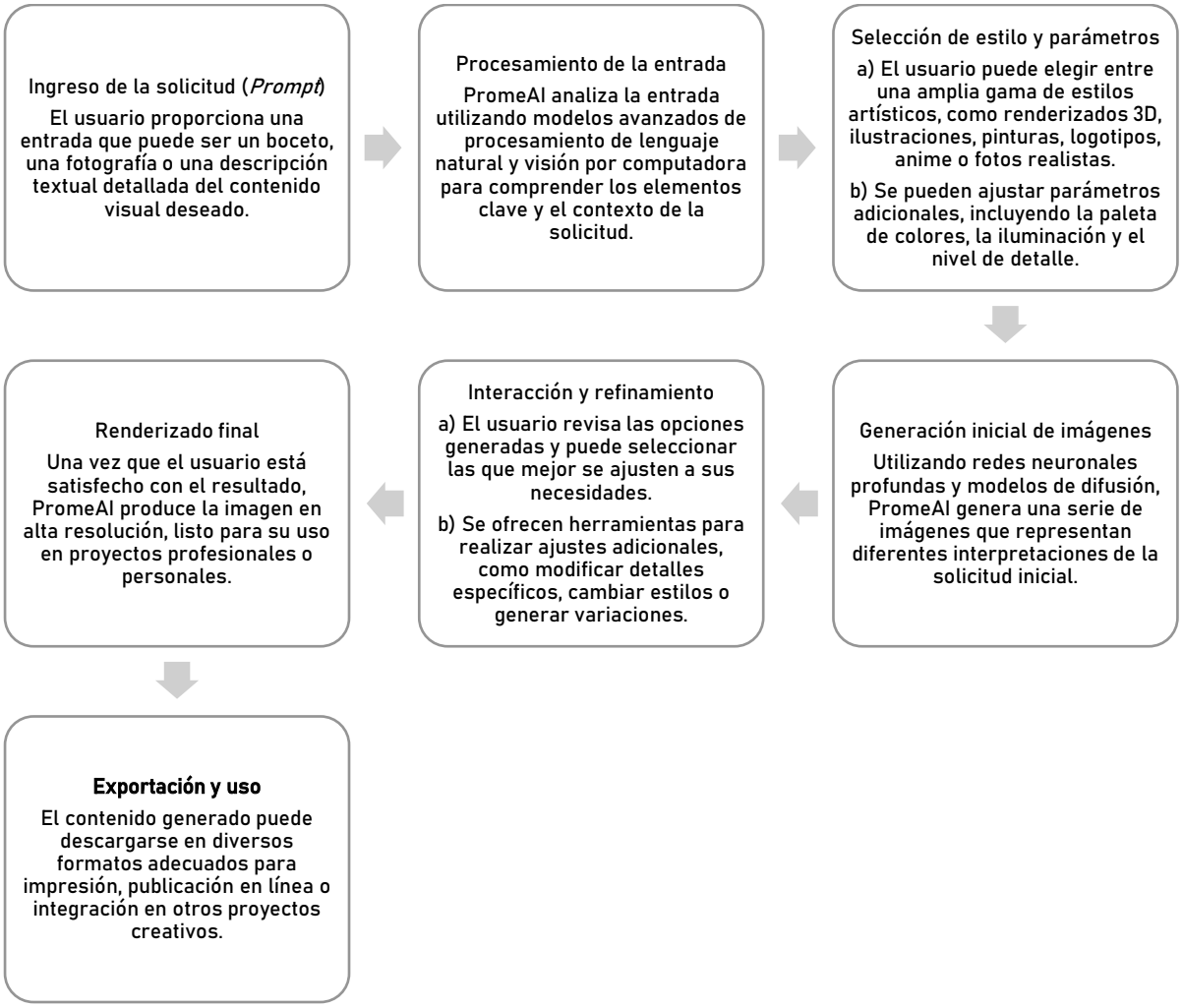
Prome-AI es una plataforma que integra herramientas de inteligencia artificial orientadas a la generación y transformación de contenido visual. Su propósito principal radica en facilitar el desarrollo de arte digital e imágenes mediante asistentes de diseño impulsados por IA y una amplia biblioteca de estilos de modelos controlables. Esta plataforma está diseñada para ser una herramienta esencial para aficionados, arquitectos, diseñadores de interiores, diseñadores de productos y diseñadores de juegos o animaciones (PromeAI, 2024). Cabe acotar que *Prome-AI* emplea modelos de inteligencia artificial generativa, específicamente modelos de difusión controlables (C-AIGC), que permiten a los usuarios generar contenido visual a partir de entradas textuales o bocetos. Estos modelos se basan en redes neuronales profundas entrenadas en vastos conjuntos de datos de imágenes, lo que les permite aprender patrones y estilos visuales diversos. La integración de estas tecnologías permite a *Prome-AI* ofrecer herramientas como el renderizado de bocetos, generador de imágenes IA, renderizado de consistencia y otras funcionalidades que facilitan el desarrollo y modificación de contenido visual de alta calidad (PromeAI, 2024).

Prome-AI encuentra aplicaciones en múltiples disciplinas creativas. En arquitectura y diseño de interiores, permite transformar bocetos en representaciones fotorrealistas, facilitando la visualización de proyectos antes de su ejecución. En el ámbito del diseño de productos, ayuda a generar prototipos visuales que pueden ser evaluados y modificados rápidamente. Además, en la industria de los videojuegos y

la animación, *Prome-AI* se utiliza para desarrollar personajes, escenarios y elementos visuales, agilizando el proceso de desarrollo y reduciendo los costos asociados a la producción de contenido visual (PromeAI, 2024). Algo a considerar de esta herramienta es que el uso de *Prome-AI* plantea desafíos éticos relacionados con la propiedad intelectual y la originalidad del contenido generado. Dado que la plataforma se basa en modelos entrenados con grandes conjuntos de datos, existe el riesgo de replicar estilos o elementos de obras protegidas por derechos de autor, lo que podría derivar en disputas legales. Técnicamente, la generación de contenido de alta calidad requiere recursos computacionales significativos, lo que puede limitar su accesibilidad para usuarios con hardware menos potente. Además, la dependencia de modelos entrenados puede introducir sesgos inherentes en el contenido generado, reflejando prejuicios presentes en los datos de entrenamiento (Bender et al., 2021).

En relación a la integración al flujo de trabajo creativo, *Prome-AI* ofrece herramientas que permiten la rápida generación y modificación de contenido visual. Los profesionales pueden utilizar la plataforma para esbozar ideas iniciales, experimentar con diferentes estilos y realizar ajustes en tiempo real, lo que facilita la iteración y refinamiento de conceptos. Esta integración permite una mayor eficiencia en el proceso creativo, reduciendo el tiempo entre la concepción de una idea y su materialización visual, y fomentando la exploración de múltiples enfoques antes de decidirse por una dirección específica (PromeAI, 2024). Finalmente, cabe mencionar que una ventaja notable de Prome-AI radica en su enfoque en la generación de contenido visual controlable, lo que permite a los usuarios influir en el estilo y características del resultado final de manera más precisa que otras herramientas generativas (PromeAI, 2024). Además, su amplia biblioteca de estilos y modelos ofrece una versatilidad que puede adaptarse a diversas necesidades creativas. Sin embargo, una limitación es la dependencia de recursos computacionales robustos para generar contenido de alta calidad, lo que puede restringir su uso en dispositivos con menor capacidad de procesamiento. Asimismo, la calidad del contenido generado está intrínsecamente ligada a la calidad y diversidad de los datos de entrenamiento, lo que implica que en áreas con datos limitados, los resultados pueden ser menos satisfactorios (PromeAI, 2024).

Figura 21. Esquema funcional de *PromeAI*



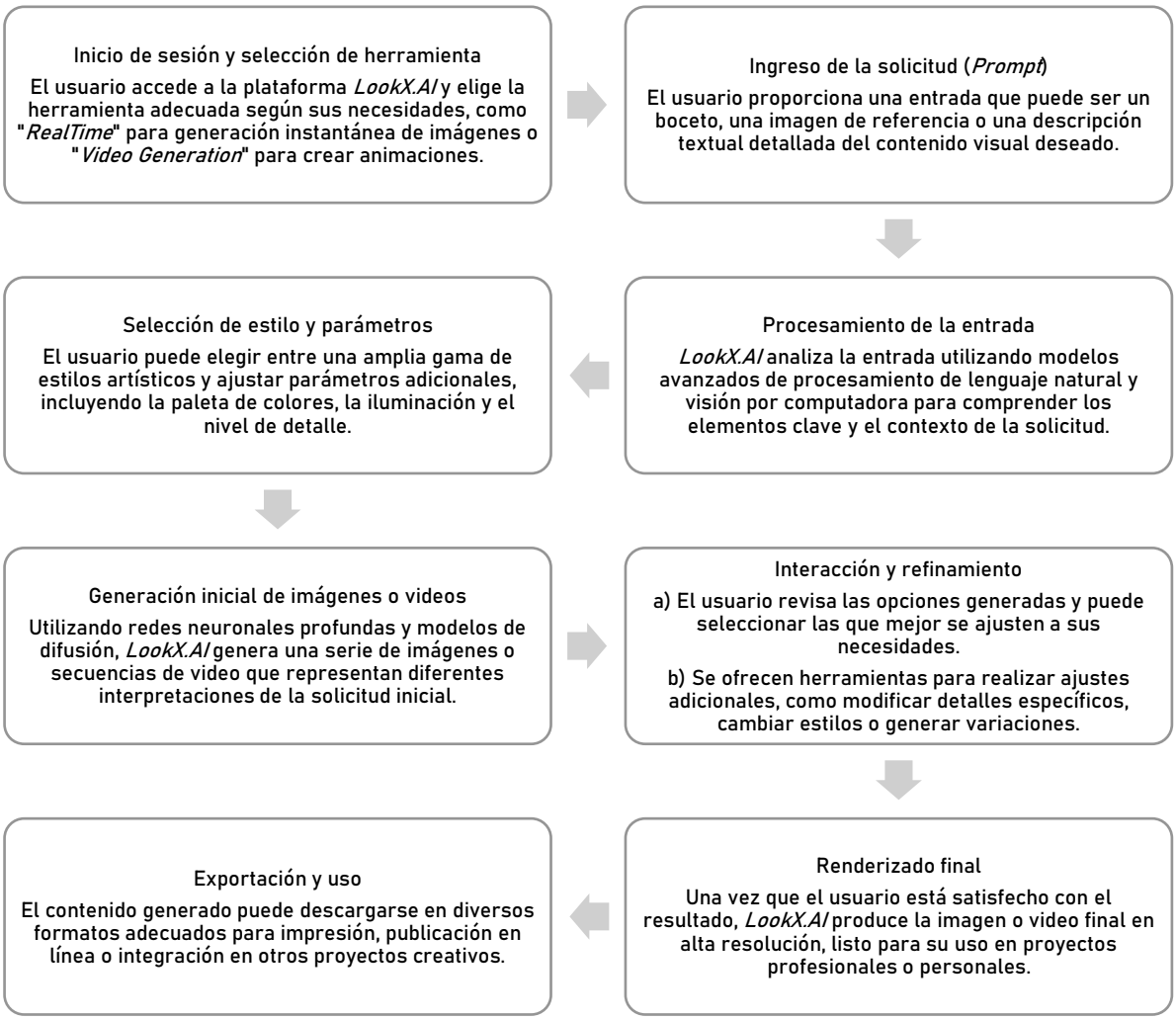
Fuente: Elaborado por el autor.

2.5.2.3. Lookx.ai

LookX.ai es una plataforma que utiliza inteligencia artificial para generar y transformar contenido visual de manera rápida y eficiente. Su propósito principal radica en permitir a los usuarios, especialmente profesionales en arquitectura y diseño, desarrollar imágenes, gráficos y conceptos visuales mediante la automatización de procesos creativos. *LookX.AI* busca democratizar el acceso a herramientas avanzadas de diseño, permitiendo que tanto expertos como no expertos en diseño accedan a un conjunto de herramientas optimizadas para el desarrollo visual, sin necesidad de un conocimiento profundo de técnicas de diseño (Lookx.ai, 2024). Cabe acotar que la tecnología de *LookX.ai* se basa en modelos generativos de redes neuronales profundas, combinando técnicas de aprendizaje automático y procesamiento de lenguaje natural para interpretar comandos

textuales y transformarlos en contenido visual. Estos algoritmos son capaces de comprender el contexto, los estilos y las tendencias visuales, permitiendo a la plataforma ofrecer resultados personalizados (Lookx.ai, 2024). Entre las tecnologías destacadas utilizadas se encuentran los modelos de aprendizaje por transferencia, que permiten adaptar imágenes a estilos específicos, y algoritmos de síntesis de imágenes que generan contenido visual coherente y atractivo en función de las necesidades del usuario (Goodfellow et al., 2014). Además, *LookX.AI* optimiza el proceso mediante el uso de bibliotecas preentrenadas que enriquecen las capacidades de personalización y ofrecen una gama diversa de estilos visuales. La capacidad de esta tecnología para ajustar parámetros y ofrecer resultados consistentes es clave para su aplicación en el campo del diseño y la arquitectura.

Figura 22. Esquema funcional de *Lookx.ai*



Fuente: Elaborado por el autor.

3. Conclusiones, reflexiones y recomendaciones

3.1. Conclusiones

El avance de la inteligencia artificial en los últimos años ha transformado de manera significativa los procesos creativos y analíticos en campos como la arquitectura, el diseño y la investigación científica. A través de este libro, se han explorado herramientas avanzadas de IA que destacan por su capacidad para generar texto, imágenes y modelos visuales de alta calidad, redefiniendo los límites de lo que es posible en estas disciplinas. En primer lugar, los modelos generativos de lenguaje, como *ChatGPT, Gemini* o *Copilot*, han demostrado ser herramientas versátiles para la síntesis de información, la automatización de procesos repetitivos y la generación de contenido contextualizado. Estas tecnologías han facilitado la comunicación y la producción intelectual, permitiendo a profesionales y académicos ahorrar tiempo y enfocarse en tareas de mayor complejidad creativa y estratégica.

A su vez, herramientas como *Midjourney, Stable Diffusion, Krea* o *Leonardo* han impulsado una nueva era en la creación visual, combinando precisión técnica con innovación artística. Estas plataformas han empoderado a diseñadores, arquitectos y artistas al proporcionarles una base tecnológica que potencia sus capacidades creativas, integrando estilos y conceptos imposibles de lograr manualmente en tiempos reducidos. No obstante, este libro también pone en evidencia algunos de los desafíos inherentes al uso de la IA. La dependencia de datos de entrenamiento, los posibles sesgos en los resultados y las cuestiones éticas sobre el uso de contenido protegido por derechos de autor son aspectos que deben abordarse con seriedad. Estas problemáticas reflejan la necesidad de desarrollar marcos regulatorios y prácticas éticas que guíen el uso responsable de estas tecnologías. En conclusión, la inteligencia artificial no solo es una herramienta de apoyo, sino una extensión de la creatividad y el ingenio humano. Su impacto continuará creciendo en los años venideros, y será responsabilidad de académicos y profesionales adaptarse a esta nueva realidad, aprovechando sus beneficios mientras mitigan sus inherentes riesgos.

3.2. Reflexiones

El uso de la inteligencia artificial en disciplinas creativas y analíticas plantea preguntas fundamentales sobre el rol de la tecnología en nuestra sociedad y la naturaleza misma de la creatividad. Este libro invita al lector a reflexionar sobre cómo estas herramientas están redefiniendo no solo los procesos, sino también las expectativas en torno al trabajo humano. En primer lugar, es necesario considerar cómo la automatización de tareas complejas podría desplazar ciertas competencias tradicionales, al tiempo que abre nuevas oportunidades para la especialización. Si bien la IA facilita tareas como la generación de texto y la visualización de conceptos, su dependencia podría llevar a una disminución en la práctica y perfeccionamiento de habilidades tradicionales. La reflexión crítica sobre el equilibrio entre lo manual y lo automatizado es, por lo tanto, fundamental.

Además, la creatividad humana está siendo complementada, y en algunos casos cuestionada, por la capacidad de la IA para generar resultados novedosos. Esto nos lleva a reflexionar sobre qué distingue la creatividad humana de la generada por máquinas. ¿Es la intencionalidad, la experiencia, o algo más intangible? Estas preguntas, lejos de ser abstractas, tienen implicaciones prácticas en cómo valoramos y remuneramos el trabajo creativo. Por último, la IA nos enfrenta a desafíos éticos y sociales que no pueden ser ignorados. Desde la propiedad intelectual hasta la representación justa de datos, estas herramientas requieren una supervisión consciente para garantizar que sus resultados sean inclusivos y éticamente sólidos.

3.3. Recomendaciones

El uso de herramientas de inteligencia artificial en los campos del diseño, la arquitectura y la investigación requiere un enfoque estratégico y ético para maximizar su potencial y minimizar posibles riesgos. En este sentido, es fundamental promover una alfabetización tecnológica integral entre profesionales y estudiantes. La comprensión profunda de las capacidades y limitaciones de estas herramientas no solo podría permitir un uso más eficaz, sino que también podría fomentar la innovación en la aplicación de estas tecnologías. Los programas educativos, tanto en instituciones académicas como en entornos profesionales, deberían incluir

formación sobre el funcionamiento técnico y práctico de las herramientas de inteligencia artificial, asegurando que los usuarios estén equipados para aprovechar sus beneficios.

Además, es imprescindible fomentar un uso ético de estas tecnologías. Las cuestiones relacionadas con la propiedad intelectual, la privacidad de los datos y la mitigación de sesgos en los modelos de IA son desafíos éticos clave que deberían abordarse con rigor. Las instituciones académicas y profesionales tienen la responsabilidad de desarrollar guías y marcos regulatorios claros que orienten el uso responsable de estas herramientas. Esto incluye garantizar que los modelos utilizados en los proyectos sean inclusivos y representativos, evitando la perpetuación de prejuicios o desigualdades. A su vez, la transparencia en el diseño y la implementación de estas herramientas debería ser una prioridad para mantener la confianza y la integridad en su aplicación.

Por otra parte, la personalización de estas herramientas sería una recomendación clave para optimizar su impacto. Muchas herramientas de inteligencia artificial están diseñadas para ser generales, pero los profesionales deberían explorar la posibilidad de entrenar modelos específicos utilizando conjuntos de datos personalizados. Este enfoque podría maximizar la relevancia y precisión de los resultados, especialmente en proyectos con requerimientos únicos o específicos. En ese sentido, la personalización permitiría no solo adaptarse a las necesidades del proyecto, sino también explorar nuevas posibilidades creativas y conceptuales que las herramientas generales no ofrecen de manera predeterminada.

La integración de la inteligencia artificial en los currículos académicos es otro aspecto crucial. La enseñanza de estas tecnologías debería convertirse en un componente central de los programas educativos en diseño, arquitectura e investigación científica. En ese sentido, preparar a las futuras generaciones para trabajar con estas herramientas podría garantizar no solo la competitividad en el mercado laboral, sino también una adopción más amplia y efectiva de estas tecnologías en diferentes industrias. Bajo ese contexto, los proyectos educativos podrían incluir el uso práctico de herramientas de inteligencia artificial, brindando a

los estudiantes la oportunidad de experimentar y comprender su funcionamiento en contextos reales.

Cabe acotar que la inteligencia artificial tiene el potencial de fomentar una colaboración interdisciplinaria más profunda. Su naturaleza transversal permite unir disciplinas tan diversas como la ingeniería, las humanidades, el diseño y la investigación científica. Bajo ese paradigma, Sería recomendable promover proyectos colaborativos que aprovechen estas herramientas para abordar problemas complejos desde múltiples perspectivas. Así, estas iniciativas podrían conducir a soluciones más efectivas, enriqueciendo no solo el producto final, sino también el proceso creativo y analítico en sí mismo. Finalmente, cabe recalcar que la inteligencia artificial representa una oportunidad sin precedentes para transformar procesos creativos y analíticos. Sin embargo, su implementación requiere un enfoque reflexivo y estratégico que integre conocimientos técnicos, consideraciones éticas y una mentalidad colaborativa.

Referencias bibliográficas

Abdul, H., Mathew, S., Ahmad, F., & Alqahtani, S. (2021). Artificial intelligence-assisted tools for redefining the communication landscape of the scholarly world. *Science editing*, 8(2), 134-144. https://doi.org/10.6087/kcse.244

Alpaydin, E. (2020). *Introduction to Machine Learning*. MIT Press. https://mitpress.mit.edu/9780262043793/introduction-to-machine-learning/

Bender, E., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, (pp. 610–623). https://doi.org/10.1145/3442188.3445922

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., & Amodei, D. (2020). Language models are few-shot learners. En *Advances in Neural Information Processing Systems*, (pp. 1877-1901). https://proceedings.neurips.cc/paper/2020/file/1457c0d6bfcb4967418bfb8ac142f64a-Paper.pdf

Castillo, P., & Beltrán, A. (2024). Explorando Sesgos Culturales en la Inteligencia Artificial Generativa (IAG): Un Enfoque en la Arquitectura y la silla Latinoamericanas. *Cuadernos del Centro de Estudios de Diseño y Comunicación*, (225). https://doi.org/10.18682/cdc.vi225.11223

Connected Papers. (1 de octubre de 2024). Explore academic papers in a visual graph. https://www.connectedpapers.com/

Consensus. (1 de octubre de 2024). Motor de búsqueda de inteligencia artificial para investigación. https://consensus.app/

D5-Hi. (1 de octubre de 2024). Convierte tus ideas en imágenes. https://d5next.ai/

DeepMind. (1 de octubre de 2024). Gemini: A Multimodal Generative AI Model. https://deepmind.com/gemini

Devlin, J., Chang, MW, Lee, K. y Toutanova, K. (2019). BERT: preentrenamiento de transformadores bidireccionales profundos para la comprensión del lenguaje. En *Actas de NAACL-HLT*, (pp. 4171-4186). https://aclanthology.org/N19-1423/

Dong, S., Wang, P., & Abbas, K. (2021). A survey on deep learning and its applications. *Computer Science Review*, *40*, 100379. https://doi.org/10.1016/j.cosrev.2021.100379

Elicit. (1 de octubre de 2024). Analizar artículos de investigación a una velocidad sobrehumana. https://elicit.com/

Floridi, L., & Cowls, J. (2019). A unified framework of five principles for AI in society. *Harvard Data Science Review*, 1(1). https://doi.org/10.1162/99608f92.8cd550d1

García, M. (2022). Sobre la autonomía, la creatividad y las consideraciones éticas de la inteligencia artificial en el arte contemporáneo. *H-ART. Revista de historia, teoría y crítica de arte*, (12), 71-96. https://doi.org/10.25025/hart12.2022.04

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press. https://goo.su/g7Kbgx

Google. (1 de octubre de 2024). Trabaja de forma más inteligente en lugar de trabajar más. https://notebooklm.google/

Iglesias Yáñez, S. (2021). El renderizado en tiempo real: integración y posibilidades en el ámbito académico. [Tesis de grado, Universidade Da Coruña]. Repositorio de la Universidade Da Coruña. https://ruc.udc.es/dspace/handle/2183/28843

Iris. (1 de octubre de 2024). Investigar de forma más inteligente. Innovar más rápido. https://iris.ai/

Khan, A. (2024). Copilot, Generative AI, and Future of Work. In *Introducing Microsoft Copilot for Managers: Enhance Your Team's Productivity and Creativity with Generative AI-Powered Assistant* (pp. 695-696). Berkeley, CA: Apress. https://link.springer.com/chapter/10.1007/979-8-8688-0419-9_18

Krea. (1 de octubre de 2024). La forma más sencilla de generar con IA. https://www.krea.ai/home

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444. https://www.nature.com/articles/nature14539

Leonardo. (1 de octubre de 2024). Creatividad, desatada. Aproveche la IA generativa con un conjunto único de herramientas para transmitir sus ideas al mundo. https://leonardo.ai/

Litmaps. (1 de octubre de 2024). Descubra el mundo de la literatura científica. https://www.litmaps.com/

Lookx.ai. (1 de octubre de 2024). Next Generation AI platform for architects & designers. https://www.lookx.ai/

Microsoft. (1 de octubre de 2024). Introducing Microsoft 365 Copilot—Your Copilot for Work. https://www.microsoft.com

Midjourney. (1 de octubre de 2024). Explora las fronteras de la imaginación. https://www.midjourney.com/home

Mikolov, T., Karafiát, M., Burget, L., Cernocký, J., & Khudanpur, S. (2013). Recurrent neural network-based language model. *Interspeech* 2010, (pp. 1045-1048). https://www.fit.vut.cz/research/group/speech/public/publi/2010/mikolov_interspeech2010_IS100722.pdf.cs

Murphy, K. (2012). *Machine Learning: A Probabilistic Perspective*. MIT Press. https://mitpress.mit.edu/9780262018029/machine-learning/

Ordelín, J. (2023). La piratería de la inteligencia artificial: el uso de las obras en el entrenamiento de modelos de IA generativos. *Actas de derecho industrial y derecho de autor*, (43), 185-205. https://dialnet.unirioja.es/servlet/articulo?codigo=9149901

Otter, D., Medina, J., & Kalita, J. (2020). A survey of the usages of deep learning for natural language processing. *IEEE transactions on neural networks and learning systems*, *32*(2), 604-624. https://ieeexplore.ieee.org/document/9075398

Perplexity AI. (1 de octubre de 2024). How Perplexity AI Works: Real-time Conversational Search. https://www.perplexity.ai

Piñón, M. (2024). Aproximaciones sobre la inteligencia artificial. *Nextia*, (20), 10-30. https://revistas.uvp.mx/index.php/nextia/article/view/311

PromeAI. (1 de octubre de 2024). El definitivo generador de arte con IA para dar vida a tu creatividad. https://www.promeai.pro/es

Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). Language models are unsupervised multitask learners. *OpenAI blog* 1(8), 9-10. https://cdn.openai.com/better-language-models/language_models_are_unsupervised_multitask_learners.pdf

ResearchRabbit. (1 de octubre de 2024). Reimaginar la investigación. https://www.researchrabbit.ai/

Remix. (1 de octubre de 2024). Remix. https://magic.remix.ai/text-to-image

Román Acosta. (2023). Más allá de las palabras: inteligencia artificial en la escritura académica. *Escritura Creativa*, *4*(2), 37-58. https://portal.amelica.org/ameli/journal/665/6654810004/6654810004.pdf

Scispace. (1 de octubre de 2024). La plataforma de investigación más rápida de la historia. https://scispace.com/

Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural Networks*, 61, 85-117. https://doi.org/10.1016/j.neunet.2014.09.003

SeaArt. (1 de octubre de 2024). Generador de arte AI gratuito y herramientas AI rápidas. https://www.seaart.ai/es

Segura-Bedmar, I., Ruz, L., & Guerrero-Aspizua, S. (2021). Evaluación de un modelo transformador aplicado a la tarea de generación de resúmenes en distintos dominios. *Procesamiento del Lenguaje Natural*, *66*, 27-39. http://journal.sepln.org/sepln/ojs/ojs/index.php/pln/article/view/6320

Semantic Scholar. (1 de octubre de 2024). Una herramienta de investigación gratuita basada en inteligencia artificial para literatura científica. https://www.semanticscholar.org/

Soliman, M., Fatnassi, T., Elgammal, I., & Figueiredo, R. (2023). Exploring the major trends and emerging themes of artificial intelligence in the scientific leading journals amidst the COVID-19 era. *Big Data and Cognitive Computing*, *7*(1), 12. https://www.mdpi.com/2504-2289/7/1/12

Stable Diffusion. (1 de octubre de 2024). Asistente estable. https://stability.ai/stable-assistant

Strubell, E., Ganesh, A., & McCallum, A. (2019). Energy and policy considerations for deep learning in PLN. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, (pp. 3645-3650). https://doi.org/10.18653/v1/P19-1355

Sutton, R., & Barto, A. (2018). *Reinforcement Learning: An Introduction*. A Bradford Book. https://web.stanford.edu/class/psych209/Readings/SuttonBartoIPRLBook2ndEd.pdf

TensorArt. (1 de octubre de 2024). TensorArt. https://tensor.art/

Tokayev, K. (2023). Ethical implications of large language models a multidimensional exploration of societal, economic, and technical concerns. *International Journal of Social Analytics*, *8*(9), 17-33. https://norislab.com/index.php/ijsa/article/view/42

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A., & Polosukhin, I. (2017). Atención es todo lo que necesitas. En *Avances en sistemas de procesamiento de información neuronal*, (pp. 5998-6008). https://papers.nips.cc/paper_files/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html

Vizcom. (1 de octubre de 2024). Hazlo realidad. Transforma tus bocetos en renders y modelos 3D en segundos. https://www.vizcom.ai/

Wang, K., Chandrasekhar, Y., Reas, R., Yang, J., Eide, D., Funk, K., Kinney, R., Liu, Z., Merrill, W., Mooney, P., Murdick, D., Rishi, D., Sheehan, J., Shen, Z., Stilson, B., Wade, AD, Wang, K., Wilhelm, C., Xie, B. & Weld, D (2020). CORD-19: El conjunto de datos de investigación abierta sobre la COVID-19. *ArXiv*. https://doi.org/10.48550/arXiv.2004.10706

-Editorial-

CILADI

Centro de Investigación Latinoamericano
para el Desarrollo e Innovación

ISBN: 978-9942-7292-6-2