

ESCUELA POLITÉCNICA NACIONAL

FACULTAD DE CIENCIAS

**INTEGRACIÓN DE MODELOS GENERATIVOS PARA LA
RECUPERACIÓN ACADÉMICA**

**PROYECTO PREVIO A LA OBTENCIÓN DEL TÍTULO DE INGENIERO DE
CIENCIAS DE LA COMPUTACION**

ALEJANDRO SEBASTIAN CHAVEZ VEGA
chavezalejo85@gmail.com

Director: DRA. GABRIELA SUNTAXI
Gabriela.suntaxi@epn.edu.ec

QUITO, JULIO 2025

DECLARACIÓN

Yo ALEJANDRO SEBASTIAN CHAVEZ VEGA, declaro bajo juramento que el trabajo aquí escrito es de mi autoría; que no ha sido previamente presentado para ningún grado o calificación profesional; y que he consultado las referencias bibliográficas que se incluyen en este documento.

A través de la presente declaración cedo mis derechos de propiedad intelectual, correspondientes a este trabajo, a la Escuela Politécnica Nacional, según lo establecido por la Ley de Propiedad Intelectual, por su reglamento y por la normatividad institucional vigente.

Alejandro Sebastian Chavez Vega

CERTIFICACIÓN

Certifico que el presente trabajo fue desarrollado por ALEJANDRO SEBASTIAN CHAVEZ VEGA, bajo mi supervisión.

Dra. Gabriela Suntaxi
Director del Proyecto

AGRADECIMIENTOS

A todos.

DEDICATORIA

*A Georg Ferdinand Ludwig Philipp Cantor,
pues nadie nos expulsará del paraíso que creó para nosotros.*

Índice general

Resumen	VIII
Abstract	IX
1. Introducción	1
1.1. Planteamiento del problema	1
1.2. Justificación	1
1.3. Justificación Metodológica	1
1.4. Objetivos	1
1.4.1. Objetivo general	1
1.4.2. Objetivos específicos	1
1.5. Alcance	2
1.6. Marco Teórico	2
2. Metodología	3
2.1. Revisión sistemática	3
2.2. Enfoque Design Science Research (DSR)	6
3. Revisión de la Literatura	9
4. Diseño del Sistema RAG	10
5. Implementación y demostración	11
6. Evaluación	12

7. Conclusiones y recomendaciones	13
Bibliografía	14

Resumen

En el presente trabajo...

Abstract

In this paper...

Capítulo 1

Introducción

1.1. Planteamiento del problema

1.2. Justificación

1.3. Justificación Metodológica

1.4. Objetivos

1.4.1. Objetivo general

Desarrollar e implementar un sistema RAG que mejore el desempeño del buscador de la plataforma Centinela, permitiendo recuperar información científica relevante y generar respuestas automáticas de valor para el usuario.

1.4.2. Objetivos específicos

- Realizar una revisión sistemática de la literatura sobre metodologías y/o frameworks para la implementación de RAG.
- Diseñar e implementar la arquitectura técnica del sistema RAG utilizando modelos de recuperación y generación de texto.
- Evaluar el sistema RAG desarrollado mediante métricas estándar.

1.5. Alcance

1.6. Marco Teórico

Capítulo 2

Metodología

2.1. Revisión sistemática

La presente revisión se fundamenta en una metodología combinada que integra la estructura de Umbrella Review, según los lineamientos del Instituto Joanna Briggs (JBI), con la estrategia de búsqueda Back-and-Forward Citation Propagating. Esta combinación metodológica permite consolidar y actualizar la evidencia secundaria (revisiones sistemáticas y metaanálisis) de manera rigurosa y dinámica, garantizando tanto la planificación estructurada como la ejecución sistemática del proceso.

Umbrella Review, de acuerdo con JBI, es un tipo de revisión sistemática que sintetiza evidencia secundaria, es decir, revisiones sistemáticas y metaanálisis ya publicados, y no estudios primarios. Su objetivo es consolidar conocimientos, identificar áreas con consenso, detectar contradicciones entre estudios secundarios y revelar vacíos en la evidencia. Este enfoque requiere la elaboración de un protocolo previo, que detalle aspectos metodológicos esenciales, como los criterios de inclusión y exclusión, las estrategias de búsqueda, el método de síntesis y otros elementos clave, aunque JBI no establece un formato único obligatorio para dicho protocolo.

En este punto, la estrategia Back-and-Forward Citation Propagating aporta una perspectiva metodológica complementaria para la búsqueda y selección de literatura, ya que, aplicada en el contexto de las revisiones tipo umbrellla, permite construir dinámicamente el corpus de evidencia mediante la exploración de redes de citación en dos direcciones:

- **Propagación hacia atrás (backward citation):** revisión de las referencias citadas en cada estudio seleccionado, lo que permite identificar trabajos previos

relevantes no detectados en la búsqueda inicial.

- **Propagación hacia adelante (forward citation):** identificación de investigaciones más recientes que han citado los trabajos clave, lo cual enriquece y actualiza el conjunto de evidencia.

A través de la propagación de citas es posible identificar literatura relevante que no siempre aparece en las búsquedas tradicionales de bases de datos como Scopus, Web of Science, etc. Estas fuentes pueden presentar limitaciones derivadas de la indexación incompleta de ciertos artículos, variaciones en el uso de palabras clave, cobertura parcial de revistas o la exclusión de publicaciones en determinados idiomas y regiones.

Metodología: Umbrella Review con Propagación de Citaciones

1. Propósito de la revisión

La revisión se justifica en la necesidad de consolidar evidencia secundaria de calidad, aprovechando el enfoque de propagación de citas para garantizar una búsqueda amplia, estructurada y actualizada.

2. Objetivos específicos

Se definen los objetivos generales y específicos que guiarán la identificación de literatura mediante la propagación de citas, así como el proceso de síntesis de resultados.

3. Criterios de inclusión y exclusión

Se incluirán únicamente revisiones sistemáticas y metaanálisis publicados entre 2019 y 2025, en inglés o español, que provengan de fuentes confiables y revisadas por pares. Se dará preferencia a aquellos trabajos que presenten de manera integral los temas más relevantes para el objeto de estudio. Se excluirán revisiones narrativas, documentos sin transparencia en los métodos de búsqueda o síntesis, y publicaciones que no se enfoquen en el objeto de estudio delimitado.

4. Identificación del estudio semilla y propagación de citas

La búsqueda se inicia en bases de datos académicas como *Scopus*, *Web of Science*, *IEEE Xplore* o *Google Scholar*, a fin de localizar un **estudio semilla** (revisión o resumen amplio) que ofrezca una cobertura representativa del tema. A partir

de este estudio, se aplica la estrategia de **Back-and-Forward Citation Propagation**, que combina:

- *Backward citation*: revisión de las referencias citadas en el estudio semilla.
- *Forward citation*: identificación de trabajos más recientes que citan al estudio semilla.

De este modo, el corpus de literatura se amplía progresivamente hasta alcanzar un punto de saturación en el que la propagación deja de aportar nueva evidencia relevante.

5. Selección de revisiones relevantes

A partir de la propagación de citas, se aplican los criterios de inclusión/exclusión para determinar qué revisiones sistemáticas y metaanálisis se incorporan al análisis final.

6. Valoración de la calidad de la evidencia

La calidad de las revisiones se considera a partir de la consistencia de los hallazgos entre distintos estudios, la claridad en la presentación de los métodos, el alcance temporal de la búsqueda (es decir, hasta qué año llega la evidencia) y las limitaciones reconocidas por los autores. Este enfoque permite diferenciar áreas con evidencia sólida de aquellas en las que persisten vacíos temáticos.

7. Extracción de información clave

De cada revisión seleccionada se extraerán datos esenciales, organizados en una tabla de extracción que incluirá:

- Autor y año de publicación
- Objetivo del estudio
- Tipo de revisión
- Número de estudios primarios incluidos
- Principales hallazgos
- Conclusiones generales
- Limitaciones reportadas

8. Síntesis y representación de resultados

Los hallazgos se organizarán en tres niveles complementarios:

- **Tabular:** tablas comparativas de las revisiones incluidas.
- **Narrativo:** síntesis descriptiva de los principales hallazgos.
- **Temático y visual:** mapas de evidencia, matrices comparativas y esquemas que reflejen la propagación de citas, mostrando la evolución temporal y las conexiones entre estudios clave.

9. Discusión y conclusiones

Los resultados se interpretan desde una perspectiva crítica, destacando fortalezas, limitaciones y la evolución de la evidencia en el tiempo. Se identifican coincidencias y divergencias entre revisiones, así como vacíos de conocimiento, y se proponen líneas de investigación futura.

En esta metodología, el Umbrella Review se utiliza como marco general para sintetizar evidencia secundaria a partir de revisiones sistemáticas y metaanálisis, complementándose con la propagación de citas (backward y forward) para ampliar perspectivas, integrar aportes recientes y reflejar la evolución del conocimiento disponible.

2.2. Enfoque Design Science Research (DSR)

De acuerdo con vom Brocke et al. [4], Design Science Research, desarrollada en 1969, es un paradigma de resolución de problemas que busca mejorar el conocimiento humano mediante la creación de artefactos innovadores. En otras palabras, es una metodología que crea soluciones a problemas reales y, al mismo tiempo, genera conocimiento útil y aplicable sobre cómo diseñar estas soluciones. Las etapas que se aplicarán en el presente trabajo son las siguientes:

- **Identificación del problema y motivación** En esta etapa se precisa el problema y se justifica por qué es necesaria una solución. De acuerdo con Peffers et al. (2008), esta etapa exige analizar el problema en detalle, descomponiéndolo en sus partes clave para identificar sus causas, efectos y alcance. Además, es crucial justificar la relevancia del problema, tanto desde una perspectiva teórica (es decir, cómo contribuye al conocimiento académico) como desde una perspectiva práctica (cómo afecta a organizaciones, usuarios o sistemas reales). También implica explorar la literatura para verificar que el problema es relevante, desafiante y nuevo, lo que permite definir los límites del proyecto de investigación.

- **Definir los objetivos para la solución** Se plantean los criterios que debe cumplir una solución exitosa basándose en el conocimiento existente y en la factibilidad técnica y organizacional. Los objetivos deberán permitir construir algo efectivo y deseable, no solamente desde el ámbito académico sino también en el entorno en que se aplicará. Estos pueden expresarse en términos cualitativos o cuantitativos; el investigador establece aquí la meta hacia donde se dirigirá el artefacto.
- **Diseño y desarrollo del artefacto** En esta etapa se construye una solución concreta, como un modelo, software o sistema, que responde directamente a los objetivos planteados. Para ello, se utiliza el conocimiento existente que fundamenta las decisiones del diseño y la estructura del artefacto. No solo se trata de crear algo, sino de asegurar que pueda ser comprendido, evaluado y replicado por otros.
- **Demostración del uso del artefacto para resolver el problema** Se muestra cómo se usa el artefacto en un escenario real o simulado. Esta demostración no valida científicamente su efectividad, sino que muestra su aplicabilidad, evidenciando que el artefacto propuesto puede operar de forma efectiva. Por su parte, vom Brocke et al. (2020) destacan que esta etapa es fundamental para conectar el diseño teórico con la realidad del usuario o del entorno organizacional, permitiendo detectar oportunidades de mejora antes de una evaluación rigurosa.
- **Evaluación del desempeño del artefacto** Se busca medir su efectividad, eficiencia e impacto, aportando evidencia que justifique su valor y utilidad. Además, según vom Brocke et al. (2020), esta puede asumirse de forma continua mediante una evaluación formativa que permita ciclos iterativos de rediseño y mejora a lo largo del proceso de investigación.
- **Comunicación de los resultados al público académico y profesional** Finalmente, esta etapa consiste en difundir de forma clara los resultados del diseño y de la investigación realizada.

Estos pasos están basados en el modelo clásico de DSR de Peffers (2008), que vom Brocke adapta y expande en su guía.

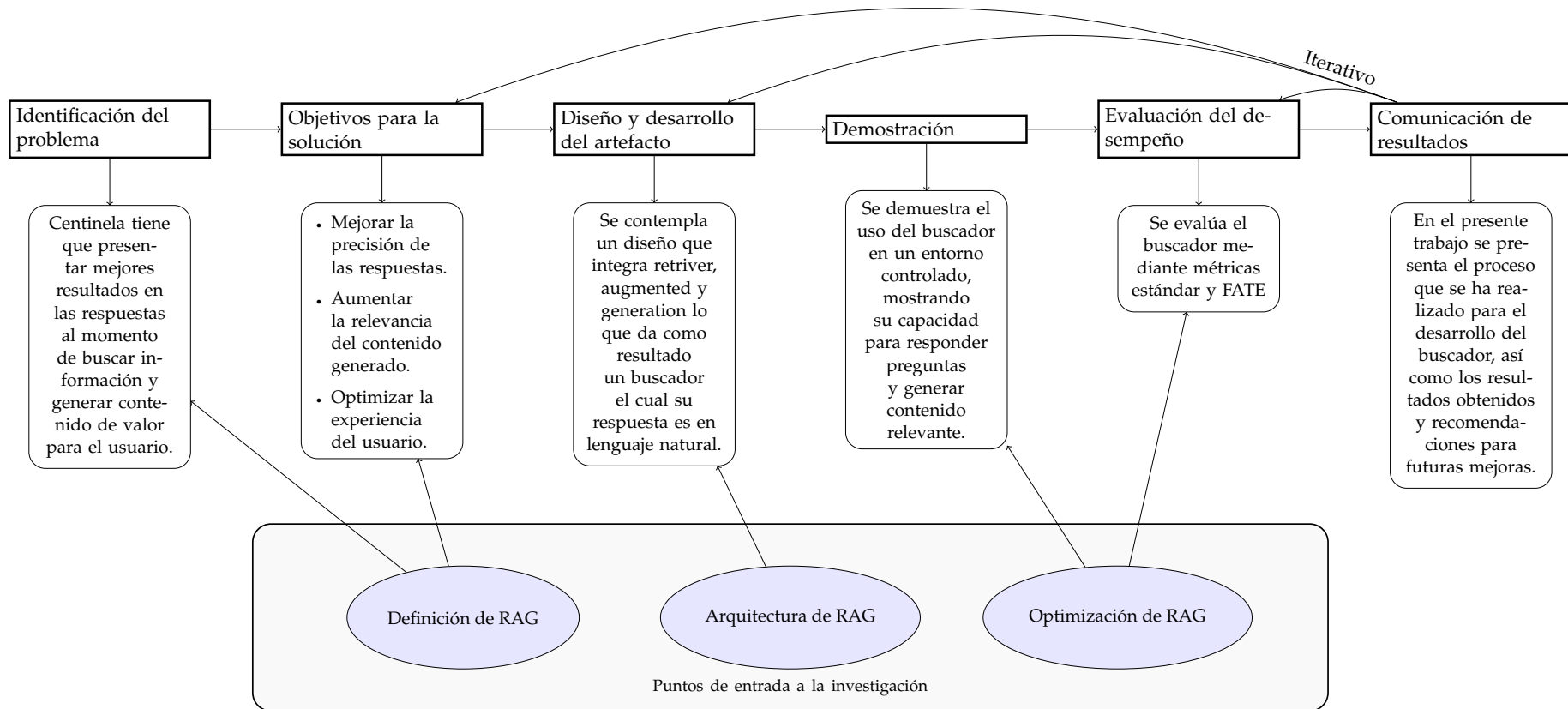


Figura 2.1: Proceso de Diseño de Investigación para el desarrollo de RAG en Centinela

Capítulo 3

Revisión de la Literatura

Capítulo 4

Diseño del Sistema RAG

Capítulo 5

Implementación y demostración

Capítulo 6

Evaluación

Capítulo 7

Conclusiones y recomendaciones

Bibliografía

- [1] Aromataris, E., Fernandez, R. S., Godfrey, C., Holly, C., Khalil, H., y Tungpun-
kom, P. *Methodology for JBI Umbrella Reviews*. Joanna Briggs Institute, 2014. Dis-
ponible en Research Online de la University of Wollongong.
- [2] Papatheodorou, S. Umbrella reviews: what they are and why we need them.
European Journal of Epidemiology, 34(6):543–546, 2019.
- [3] Peffers, K., Tuunanen, T., Rothenberger, M. A., y Chatterjee, S. A design science
research methodology for information systems research. *Journal of Management
Information Systems*, 24(3):45–77, 2008.
- [4] vom Brocke, J., Hevner, A., y Maedche, A. *Design Science Research*. Springer,
Cham, 2020.