

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
1.7. Revision de literatura	2
2. Metodología	3
2.1. Revisión sistemática	3
2.1.1. Ejecución del estudio de la literatura	6
2.2. Enfoque Design Science Research (DSR)	9
2.3. Diseño y desarrollo del artefacto	12
3. Pruebas, Resultados, Conclusiones y Recomendaciones	13
3.1. Demostracion	13
3.2. Evaluacion del desempeño	13

3.3. Resultados	13
3.4. Conclusiones	13
3.5. 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

1.7. Revision de literatura

Capítulo 2

Metodología

2.1. Revisión sistemática

La Umbrella Review, según los lineamientos del Instituto Joanna Briggs (JBI), es un tipo de revisión sistemática que recopila y analiza evidencia secundaria, es decir, revisiones sistemáticas y metaanálisis ya publicados. Su propósito es consolidar el conocimiento disponible, identificar coincidencias y contradicciones en la literatura existente, así como señalar vacíos de evidencia. Para ello, requiere la elaboración de un protocolo previo que establezca criterios de inclusión y exclusión, estrategias de búsqueda y métodos de síntesis, garantizando un proceso transparente y riguroso.

Por otra parte, la estrategia de propagación de citas (Back-and-Forward Citation Propagating) complementa este enfoque al permitir encontrar dinámicamente la literatura. A través de la propagación de citas se amplía y actualiza la literatura encontrada en las bases de datos tradicionales. De este modo, se superan limitaciones como la indexación incompleta, las variaciones en el uso de palabras clave o la exclusión de ciertas publicaciones.

Metodología: Umbrella Review con Propagación de Citaciones

Como parte de la metodología Umbrella Review es necesario establecer un protocolo para ejecutar la revisión. Se han considerado las siguientes fases para dicho protocolo:

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 definen de manera general como la incorporación de revisiones y metaanálisis que sean pertinentes, de calidad y relacionados con el tema de estudio, y la exclusión de aquellos trabajos que no cumplan con estos requisitos de relevancia.

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 serán incorporadas al análisis.

6. Valoración de la calidad de la evidencia

De los estudios seleccionados se procederá a excluir aquellos artículos que no cumplan con los criterios de inclusión previamente establecidos o que presenten un nivel de profundidad insuficiente para los objetivos de la revisión. La selección final de los estudios se llevará a cabo en consenso con expertos en el área, lo que asegura la validez y relevancia de la evidencia incluida. Para la

organización, codificación y síntesis de la información se empleará una herramienta de apoyo, como ATLAS.ti, que facilitará la estructuración sistemática de los hallazgos.

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 dos niveles complementarios:

- **Tabular:** tablas comparativas de las revisiones incluidas. (tal vez)
- **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.

2.1.1. Ejecución del estudio de la literatura

Propósito y objetivos de la revisión

El propósito de esta revisión es consolidar la información disponible sobre los RAG, abordando su estudio desde los fundamentos hasta las fases de desarrollo. Se inicia con su definición y arquitectura, para luego profundizar en las etapas clave del proceso: Extracción del corpus, preprocesamiento, vectorización, recuperación de información, evaluación, almacenamiento en bases vectoriales y generación de resultados. Asimismo, se examinan los paradigmas, las métricas de evaluación y el futuro de RAG. Durante esta revisión se busca lograr el objetivo general de proporcionar un panorama global y actualizado sobre los RAG, exponiendo sus fundamentos, desarrollo y aplicación.

Criterios de inclusión y exclusión

Se incluyen únicamente revisiones sistemáticas y metaanálisis publicados entre 2018 y 2025, en inglés o español, dado que la producción científica en el área comienza a incrementarse a partir de 2018, con la popularización de los modelos de lenguaje basados en transformers (p. ej., BERT, GPT-2, T5). Los estudios deben provenir de fuentes confiables los cuales son, a su vez, revisados por un experto. Se da preferencia a aquellos que presenten una cobertura de los temas más relevantes para el objeto de estudio. Se excluyen las revisiones narrativas, los documentos que carezcan de transparencia en sus métodos de búsqueda o síntesis, así como las publicaciones que no estén directamente relacionadas con el objeto de estudio delimitado.

Identificación del estudio semilla y selección de revisiones relevantes

El proceso de búsqueda se inicia con la identificación de dos estudios semilla, extraídos de Google Scholar con los parámetros de búsqueda "retrieval informationz retrieval augmented generation", aplicando el filtro desde el año 2018 hasta 2025.

El primer estudio seleccionado fue Information Retrieval: Recent Advances and Beyond (Hambarde & Proença, 2023), publicado en IEEE Access. Este trabajo constituye una revisión exhaustiva de la recuperación de información, abarcando desde los métodos tradicionales hasta los enfoques basados en deep learning y transformers, por lo que resulta un punto de partida principal para explorar la literatura

reciente y relevante.

El segundo estudio semilla corresponde al artículo Retrieval-Augmented Generation for Large Language Models (Gao, Xiong, Gao, Jia, Pan, Bi, Dai, Sun & Wang, 2023), publicado en arXiv, el cual presenta un marco conceptual y aplicado sobre la integración de recuperación de información y modelos generativos de gran escala. Su incorporación permite establecer una base teórica para contextualizar el análisis de las revisiones seleccionadas.

A partir de estos dos estudios semilla, y aplicando los criterios de inclusión y exclusión previamente definidos, se identificaron 25 revisiones relevantes que cumplan con los criterios establecidos. Estas revisiones constituyen la base para el análisis y síntesis en el presente trabajo.

Valoracion de la evidencias y extracion de la infomacion

De los estudios seleccionados se procede a realizar un análisis, con el fin de excluir aquellos artículos que no cumplen con los criterios establecidos o que presentan un nivel de profundidad insuficiente para los objetivos de la revisión. La selección final de los estudios se realiza en consenso con expertos en el área, garantizando así la pertinencia y relevancia de la evidencia incluida.

para resuemir la informacion se uso (herramienta) atlas por ejemplo

Síntesis y representación de resultados

Con la literatura seleccionada se identificó la hoja de ruta que se presenta a continuación en la figura 2.1.

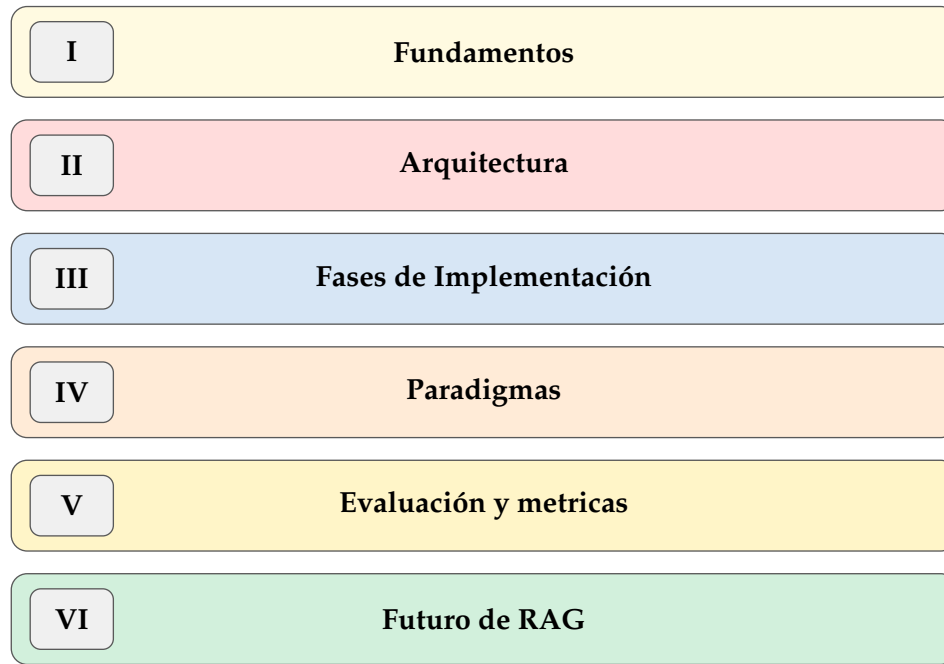


Figura 2.1: Resumen esquemático de RAG

A partir de esta hoja de ruta se desarrolla un esquema más detallado, en el que se expone primero exploraremos su teoría, características y aplicaciones, como se muestra en la Fig 2.2. Luego profundizamos en su arquitectura en la cual se describe cada uno de los componentes que lo conforman (retriever, augmented y generation) y las variantes y mejoras que existen de cada uno. Posteriormente, se detalla el proceso de implementación, desde la preparación de datos hasta el componente de generación, incluyendo las técnicas y herramientas más relevantes. A continuación en la Fig , se examinan los paradigmas de RAG, dando a conocer los tipos de paradigmas y sus clases, para luego en la Fig , se presentan las métricas y evaluadores automáticos utilizados en la evaluación de sistemas RAG, así como las consideraciones éticas y de equidad que deben tenerse en cuenta. Finalmente, se discuten las tendencias emergentes, los desafíos que actualmente se tienen y futuras direcciones que podrían tomar los sistemas RAG.

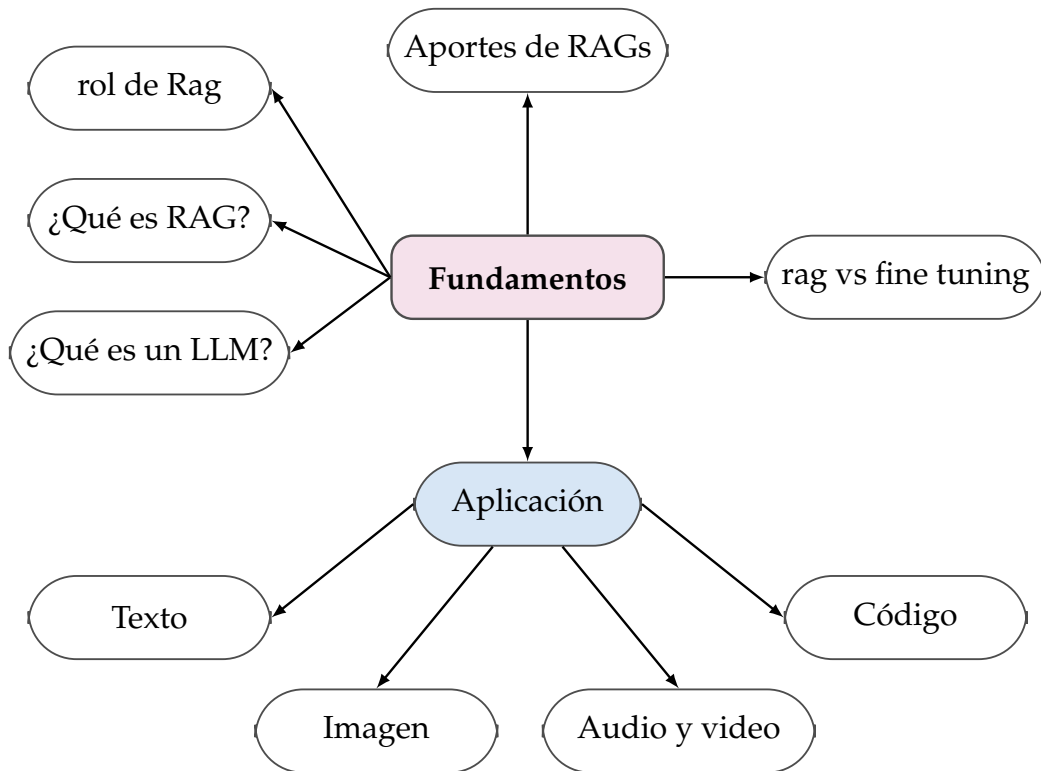


Figura 2.2: Fundamentos de RAG

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 rele-

vante, 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.

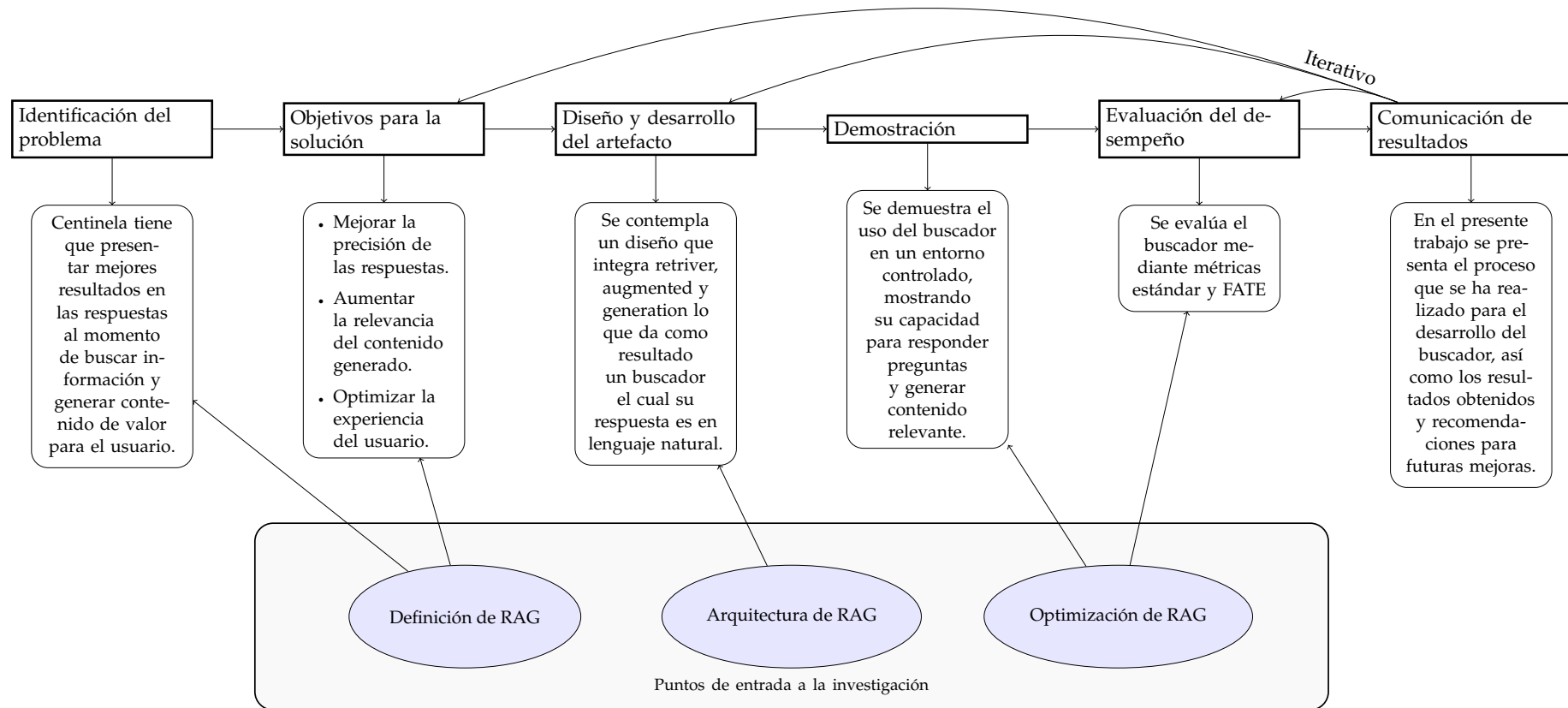


Figura 2.3: Proceso de Diseño de Investigación para el desarrollo de RAG en Centinela

2.3. Diseño y desarrollo del artefacto

Capítulo 3

Pruebas, Resultados, Conclusiones y Recomendaciones

3.1. Demostracion

3.2. Evaluacion del desempeño

3.3. Resultados

3.4. Conclusiones

3.5. 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.