** F\_AA\_234A**

**VICERRECTORADO DE DOCENCIA**

**PLAN DE TRABAJO DE INTEGRACIÓN CURRICULAR**

|  |  |  |
| --- | --- | --- |
| **I. INFORMACIÓN BÁSICA** | | |
| **Unidad Académica:** | Facultad de Ingeniería de Sistemas | |
| **Carrera:** | **Ingeniería en ciencias de la computación** | |
| **Proyecto:** | Proyecto Centinela Avanzado V2: Búsqueda Optimizada y Fortalecimiento en Colaboración Académica | |
| **Componente:** | Gestión y Personalización de Feeds Académicos | |
| **Línea de investigación:** | **Ciencias de la Computación e Inteligencia Artificial - Tópico: Inteligencia artificial**  **Ingeniería de Software - Tópico: Motores de Búsqueda y Recuperación de Información** | |
| **Nombres y apellidos del estudiante:** | **Alejandro Sebastian Chavez Vega** | |
| **Nombres y apellidos del Profesor:** | **Ing. Gabriela Suntaxi, Ph.D.** | |
| **II. DESCRIPCIÓN DEL COMPONENTE** | | |
| Con la popularizacion de los LLM, los sistemas de recuperacion de informacion se adapataron a busquedas con lenguaje natural y tener un contexto semantico de lo que requiere el usuario. Lo que principalmente se ajusta a lo que necesita centinela. Centinela es una plataforma que busca fortalecer la investigacion científica en Ecuador. Su objetivo principal es mejorar el acceso a la información científica y fomentar la colaboración entre investigadores. Incorporar RAG en Centinela transforma la plataforma en un agente capaz de interpretar el contexto del usuario y sus intereses para generar respuestas coherentes y utiles. A diferencia con el buscador tradicional actual que simplente devuelve una lista de articulos, RAG primero recupera los documentos mas relevantes y luego genera un respuesta en lenguaje natural, basada en los articulosa académicos almacenados en las bases de datos del sistema. Esto permite que el investigador tenga una sintesis clara incluso si no se usa terminos tecnicos exactos, el sistema extrae un resumen de los resultados obtenidos y los autores, todo a partir de los articulos disponibles.  Como se observa en la figura, la plataforma se divide en tres módulos principales:  • **Recuperador de Información:** Este módulo gestiona la persistencia de datos provenientes de Scopus, permitiendo búsquedas actualizadas por perfiles, autores y artículos relevantes. Aquí se propone integrar el componente RAG como una capa adicional sobre el motor de búsqueda actual. Este componente sería responsable de interpretar consultas en lenguaje natural, recuperar artículos relevantes y generar respuestas comprensibles basadas en el contenido de los documentos. Esta capa puede alojarse dentro del backend del Recuperador, conectándose tanto al índice como a los modelos de generación.  • **Servicio de Analítica Descriptiva:** A través de este servicio, Centinela ofrece información para la toma de decisiones a través de visualizaciones de datos. Este módulo se complementa con un 'Data Lake' que proporciona los recursos necesarios para la analítica y la generación de dashboards interactivos.  • **Interacción Académica y Consenso:** Este módulo promueve la colaboración entre investigadores mediante la creación de grupos, votaciones y talleres de consenso. El sistema RAG puede extenderse también como asistente dentro de estos espacios, sugiriendo artículos relevantes para debates, generando resúmenes temáticos o respondiendo preguntas específicas durante las sesiones de discusión académica. | | |
| **III. OBJETIVOS** | | |
| **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.  **Objetivos Específicos:**   * Realizar una revisión sistemática de la literatura sobre metodologías y/o frameworks para la implementacion de RAG. * Diseñar e implementar la arquitectura técnica del sistema RAG utilizando modelos de recuperación y generación de texto. * Evaluar el sistema RAG desarollado en comparación con el buscador tradicional de Centinela mediante métricas estándar y pruebas de usuario. | | |
| **IV. ALCANCE DEL COMPONENTE** | | |
| Este componente del proyecto se enfocará en el desarrollo e integración de un sistema RAG en el módulo de búsqueda de Centinela en los articulos relvantes. Para llevar a cabo, se usara umbrella slr, es un tipo de revisión sistemática que sintetiza evidencia secundaria, es decir, revisiones sistemáticas y metaanálisis ya publicados, y no estudios primarios, y prisma aporta un marco estandarizado y ampliamente validado para el desarrollo del protocolo. Este enfoque combinado utiliza 2 fases, en la primera se actua principalmente donde permite definir de forma clara y anticipada todos los componentes del protocolo necesarios para ejecutar la revisión y en la segunda fase se ejecuta la búsqueda definida, se seleccionan las revisiones sistemáticas relevantes, se evalúa su calidad metodológica con las herramientas del JBI.  Por otra parte, se usa DSR para crear una solucion al probema que enfrenta Centinela y, al mismo tiempo, generar conocimiento útil y aplicable sobre cómo diseñar estas soluciones.  Las etapas que se aplicarán en el presente trabajo son las siguientes:   * Definir los objetivos para la solución * Diseño y desarrollo del artefacto * Demostración del uso del artefacto para resolver el problema • Evaluación del desempeño del artefacto * Comunicación de los resultados | | |
| **V. TAREAS ESPECÍFICAS** | | |
| Con base en las actividades establecidas en el proyecto y asignadas al componente, se plantea las tareas que permiten cumplir con los objetivos del proyecto, y que, posteriormente, permiten la evaluación de su cumplimiento a través de los resultados esperados.   |  |  |  |  | | --- | --- | --- | --- | | **Semana referencial /**  **Etapas** | **Fecha inicio-fin (si aplica)** | **Tareas específicas** | **Resultado esperado (si aplica)** | | **1** |  | Familiarización con Centinela V2: arquitectura y estado actual |  | | **2** |  | Deployment de versión de pruebas del sistema | Versión de pruebas (funcional) del sistema | | **3** |  | Revisión de literatura |  | | **4** |  | **Levantamiento de requisitos específicos del componente** | **Detalle de requisitos** | | **5** |  | Investigación de herramientas y modelos aplicables a la arquitectura RAG. |  | | **6** |  | Desarrollo del módulo de recuperación semántica (Retriever). |  | | **7** |  | Implementación del módulo de enriquecimiento y expansión de consultas (Augmentation). |  | | **8** |  | Integración de un modelo generativo de texto (Generator). |  | | **9** |  | Componente integrado en Centinela |  | | **10** |  | **Difusión de proyecto y resultados** | **Reporte detallado documentando las recomendaciones específicas para mejorar la calidad de la busqueda.** |   **NOTAS:**   * **El estudiante, en la fecha establecida en el Calendario Académico (aproximadamente cuatro semanas antes del cierre del SAI), remite el documento final y/o el producto al Director del TIC.** * **El Director, en la fecha establecida en el Calendario Académico, remite el documento final y/o el producto para revisión, al Decano o Director de la ESFOT.** * **Los revisores remiten la calificación al Decano o Director de la ESFOT con copia al Director del TIC, en la fecha establecida en el Calendario Académico. Los revisores no podrán solicitar correcciones.** | | |
| **VI. BIBLIOGRAFÍA** | | |
| **[1] 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.**  **[2] L. Recalde, G. Suntaxi, D. Martinez-Mosquera, R. Masabanda, and D. Cabrera, "Centinela: An Intelligent System Based on an Integrated Architecture for Supporting Scholars," in Proc. Conf. Information and Communication Technologies of Ecuador, Cham, Switzerland: Springer Nature Switzerland, Oct. 2024, pp. 160–177** | | |
| **VII. FIRMAS DE RESPONSABILIDAD** | | |
| **Desarrollado por:** | | **Alejandro Sebastian Chavez Vega** |
| **Aprobado por:** | | **Ing. Gabriela Suntaxi, Ph.D.** |