

Proyecto final – Unidad III: Proyecto de TAL

Entrega: 12 de diciembre de 2025

Trabajo en grupos: max. 2 estudiantes

Contexto

El proyecto se enmarca en *Sophia Search*, un sistema que recopila y organiza noticias de prensa chilenas. A partir de un **conjunto de noticias recientes (Chile, septiembre 2025)**, cada grupo deberá diseñar e implementar un **prototipo RAG (Retrieval-Augmented Generation)** que permita responder preguntas en lenguaje natural sobre los contenidos noticiosos.

El objetivo es combinar técnicas de procesamiento del lenguaje, recuperación semántica, agentes inteligentes y generación de texto, considerando además **metadatos de tiempo y lugar** para contextualizar correctamente las respuestas.

Objetivos específicos

- Aplicar técnicas de TAL.
- Diseñar e implementar un pipeline RAG funcional.
- Incorporar metadatos temporales y geográficos en el proceso de análisis.
- Sintetizar y comunicar los resultados del sistema en un contexto de innovación y análisis social computacional.
- Reflexionar sobre las implicancias éticas y sociales del uso de modelos de lenguaje.

Ejemplos de consultas posibles al sistema RAG

- “¿Qué dijo el Presidente de Chile la semana pasada?”
- “¿Qué noticias recientes se publicaron en la Región de Los Lagos sobre energía?”
- “¿Cómo cubrieron los medios el tema de la inflación durante octubre?”

Entregables

1. **Informe técnico (máx. 6 páginas):** Objetivos y contexto, Arquitectura del sistema, Técnicas TAL utilizadas, Resultados y ejemplos, Discusión ética y social, Conclusiones
2. **Código funcional**

3. Reunion final presencial u online (10 minutos) entre el 15 y 19 de diciembre

- Demostración del sistema
- Síntesis del enfoque y de los aprendizajes obtenidos