

# Asistente PLN para Medios de Comunicación


# ¿Qué problema buscamos resolver?

## El problema:

- En la actualidad, los medios digitales argentinos publican gran cantidad de artículos por día.
- Esta sobrecarga informativa dificulta identificar rápidamente el enfoque emocional o ideológico de una noticia.
- Además, los lectores (y los propios periodistas) enfrentan el desafío de resumir, entender y comunicar eficazmente lo que sucede.
- También crece la demanda de generar contenido dinámico (tweets, titulares, resúmenes) en poco tiempo y con estilo local.

## Nuestro desafío es doble:

- Analizar automáticamente el contenido noticioso para extraer lo más importante.
- Generar contenido breve y contextualizado para medios y redes.



# Nuestra solución: un asistente inteligente de noticias

## **Nuestra propuesta:**

Desarrollar un asistente basado en procesamiento de lenguaje natural (PLN) orientado a medios de comunicación digitales.

El asistente se enfoca en analizar noticias actuales de medios argentinos y generar contenido útil a partir de ellas.

## **¿Qué hace el asistente?**

Extrae automáticamente artículos de medios digitales (como Infobae/La Nación).

Limpia y analiza el texto para identificar entidades clave (personas, organizaciones, lugares).

Evalúa el tono emocional del conjunto de notas (positivo, negativo o neutro).

Genera resúmenes automáticos y frases tipo tweet con estilo local.



# Pipeline: del texto crudo a la generación automática

1. Web Scraping: El asistente comienza extrayendo artículos completos desde sitios de medios digitales argentinos. Se recopilan entre 10 y 20 notas recientes sobre economía, con el cuerpo completo del texto.
2. Preprocesamiento: Se limpia el texto: se eliminan HTMLs, símbolos raros y se normaliza el lenguaje. Se tokeniza, lematiza y filtra usando spaCy (en español) para quedarnos con palabras clave.
3. Análisis automático: NER (Named Entity Recognition): se identifican entidades importantes como personas, lugares y organizaciones. Análisis de sentimiento: se evalúa si el tono del conjunto de notas es positivo, negativo o neutro.
4. Generación de contenido: Se genera un resumen automático del corpus total, una nube de palabras, se muestran las entidades más relevantes, se analiza el sentimiento y se genera una nota de opinión.
5. Visualizaciones e Interfaz: A través de una interfaz interactiva de Gradio se muestran los resultados obtenidos.



# Impacto y beneficios

Ahorra tiempo en el análisis de noticias y generación de contenido.

Detecta patrones de enfoque emocional o sesgo en medios.

Extrae lo más importante de forma estructurada y rápida.

Genera textos útiles (resúmenes, tweets) con estilo local.

Escalable y adaptable a distintos temas y medios.

Una herramienta pensada para periodistas, editores y usuarios que quieren entender más y mejor en menos tiempo.