

# PARTISANLENS

A Multilingual Dataset of Hyperpartisan and Conspiratorial Immigration Narratives in European Media



Funded by  
the European Union



UK Research  
and Innovation

## Usage/License



This dataset is released under the Apache License 2.0, allowing reuse, modification, and distribution under its terms

## Download



GitHub Repo: <https://github.com/MichJoM/PartisanLens>

## Researchers



- Michele Joshua Maggini
- Paloma Piot
- Anxo Pérez
- Erik Bran Marino
- Lúa Santamaría Montesinos
- Ana Lisboa
- Marta Vázquez Abuín
- Javier Parapar
- Pablo Gamallo

Affiliated with CiTIUS (USC), IRLab-CITIC (UDC), UEVORA, UNIRIOJA, and GESIS

## Summary

PARTISANLENS is the first multilingual dataset of 1,617 hyperpartisan news headlines in Spanish, Italian, and Portuguese, collected from 329 outlets between 2020 and 2024. Each headline is annotated across six dimensions: hyperpartisanship, Population Replacement Conspiracy Theories (PRCT), stance toward immigration, and three rhetorical bias types (loaded language, appeal to fear, name-calling). The dataset establishes robust LLM baselines and explores ideology-conditioned persona simulation for annotation.

## Key Features

- Multilingual coverage of three underrepresented Romance languages (ES, IT, PT)
- Six-dimensional annotation capturing hyperpartisanship, PRCT, stance, and rhetorical bias
- Strong inter-annotator agreement (Fleiss'  $\kappa$ : 0.70–0.89)
- Benchmark experiments: zero-/few-shot and fine-tuning with mBERT and LLMs (Llama, Mistral)
- Novel persona simulation using annotator socio-demographic and ideological profiles

## What We Offer

- A curated, balanced dataset of 1,617 immigration-related headlines with both individual and majority-vote labels
- Full annotation guidelines, prompts, and a custom web annotation interface
- Rationale-enhanced annotations linking rhetorical spans to classification decisions
- All data, code, and prompts released publicly for reproducibility

## Collaboration Objectives

- Address gaps in multilingual European misinformation research with a focus on immigration discourse
- Enable joint analysis of hyperpartisanship, conspiracy, and stance in political discourse
- Study LLMs as surrogate annotators and the impact of ideological profiles
- Support cross-lingual transfer and discourse-aware modeling for partisan detection