

# Annotation Guidelines for Hyperpartisan and PRCT Detection in News Headlines

Media Cloud Dataset

*HYBRIDS Project*

Marie Skłodowska-Curie Actions

May 30, 2025

# Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
1.1	Annotation Plan . . . . .	3
<b>2</b>	<b>Hyperpartisan Detection</b>	<b>3</b>
2.1	Definition . . . . .	3
2.2	Labels . . . . .	4
2.3	Examples . . . . .	4
<b>3</b>	<b>Population Replacement Conspiracy Theories (PRCT) Detection</b>	<b>4</b>
3.1	Distinguishing PRCT from Racism . . . . .	4
3.2	Key PRCT Theories and Concepts . . . . .	4
3.3	Additional PRCT Indicators and Dog Whistles . . . . .	5
3.4	Labels . . . . .	5
3.5	Examples of PRCT Dog Whistles . . . . .	6
3.6	Borderline Cases with Explanations . . . . .	6
3.6.1	Examples of Racist Content (Label 0 for PRCT) . . . . .	6
3.6.2	Examples of Clear PRCT Content (Label 1) . . . . .	6
3.6.3	Borderline Cases . . . . .	7
<b>4</b>	<b>Rhetorical Bias Detection and Span Annotation</b>	<b>7</b>
4.1	General Principles . . . . .	7
4.2	Loaded Language Span Detection . . . . .	7
4.2.1	Definition . . . . .	7
4.2.2	Span Selection Criteria . . . . .	7
4.2.3	Exclusion Criteria . . . . .	8
4.3	Appeal to Fear Span Detection . . . . .	8
4.3.1	Definition . . . . .	8
4.3.2	Span Selection Criteria . . . . .	8
4.3.3	Specific Examples . . . . .	8
4.3.4	Exclusion Criteria . . . . .	9
4.4	Name Calling Span Detection . . . . .	9
4.4.1	Definition . . . . .	9
4.4.2	Span Selection Criteria . . . . .	9
4.4.3	Specific Examples . . . . .	9
4.4.4	Exclusion Criteria . . . . .	9
<b>5</b>	<b>Stance Detection</b>	<b>9</b>
<b>6</b>	<b>Annotation Process</b>	<b>10</b>
<b>7</b>	<b>Additional Notes</b>	<b>10</b>
7.1	Language-Specific Nuances . . . . .	10
7.2	Cultural Context . . . . .	10
7.3	Critical Mentions . . . . .	11
7.4	Consistency . . . . .	11
<b>8</b>	<b>Frequently Asked Questions (FAQ)</b>	<b>11</b>

8.1	Can a text be both hyperpartisan and contain PRCT elements? . . . . .	11
8.2	What if a headline uses neutral language but the newspaper is known for spreading hyperpartisan news or promoting PRCTs? . . . . .	11
8.3	How should we handle sarcasm or irony in headlines? . . . . .	11
8.4	What if a headline is ambiguous or unclear? . . . . .	11
8.5	How should we handle headlines that discuss immigration but do not explicitly mention PRCT or hyperpartisanship? . . . . .	12
8.6	What if a headline is written in a sensationalist style but does not explicitly promote a conspiracy theory? . . . . .	12
8.7	What if a headline contains quotes? . . . . .	12
8.8	Should we leverage external knowledge to detect hyperpartisanship when the headline appears neutral? . . . . .	12
8.9	How should quotes be considered in stance detection? . . . . .	13

# 1 Introduction

These guidelines are designed to assist annotators in classifying news headlines along six dimensions: **Hyperpartisan Detection**, **Population Replacement Conspiracy Theory (PRCT) Detection**, **Stance Detection**, and three rhetorical biases: **Loaded Language**, **Appeal to Fear**, and **Name Calling**.

The dataset consists of migration-related headlines in **Italian**, **Portuguese**, and **Spanish**, collected using the Media Cloud news aggregator. The goal is to ensure consistency and accuracy in labeling, which is crucial for evaluating the performance of Large Language Models (LLMs) in these tasks.

## 1.1 Annotation Plan

The annotation process will consist of four rounds: 100, 100, 150, and 150 headlines respectively. The expected duration is four weeks, with weekly meetings scheduled and training provided by in-domain experts.

Annotators will be asked to complete the COMPASS test (<https://www.politicalcompass.org/>) to include ideological orientation in the annotation process. Socio-demographic information will also be collected, anonymized, and not disclosed. This information will be useful for training models using dimensions such as annotator disagreement, gender, age, background, and political bias, leading to meaningful insights on model explainability.

We recognize that annotation is a time-consuming and exhausting practice, and we appreciate your effort and contribution to this research.

# 2 Hyperpartisan Detection

## 2.1 Definition

Hyperpartisan content is characterized by strong ideological bias, explicit references to extremist political positions, or a subjective style intended to influence reader opinion. It often uses emotionally charged language, exaggerations, or one-sided arguments. These traits can be implicit or explicit.

Hyperpartisan language leverages the adoption of specific rhetorical biases, particularly:

- **Appeal to Fear:** This technique aims at promoting or rejecting an idea through the repulsion or fear of the audience towards this idea or its alternative. The alternative could be the status quo, described in a frightening way with loaded language.
- **Loaded Language:** Use of specific words and phrases with strong emotional implications (either positive or negative) to influence and convince the audience that an argument is valid.
- **Name Calling:** A form of argument in which loaded labels are directed at an individual or group, typically in an insulting or demeaning way.

**Note:** The biases described above can also be applied to PRCT news headlines.

## 2.2 Labels

- **Label 0 (Non-Hyperpartisan):** Neutral content that presents information objectively without ideological bias or attempts to influence opinion.
- **Label 1 (Hyperpartisan):** Content that exhibits clear ideological bias, uses emotionally charged language, or promotes a specific political agenda.

## 2.3 Examples

Language	Headline	Label	Explanation
Italian	"Ripartizione dei migranti, cosa prevede Dublino"	0	Neutral reporting on a policy topic
Portuguese	"Imigração: desafios e oportunidades para o mercado de trabalho"	0	Neutral discussion of immigration and labor
Spanish	"La UE debate nuevas políticas migratorias para 2024"	0	Informative and neutral headline
Italian	"Migranti, sbarchi senza fine: Lampedusa in tilt mentre la Germania ci rifila l'ennesima fregatura"	1	Emotional language ("tilt", "ennesima fregatura") and ideological bias
Portuguese	"Governo de esquerda incentiva imigração ilegal para destruir a nação!"	1	Extreme accusation ("destruir a nação") and polarized language
Spanish	"La derecha fascista cierra fronteras mientras los migrantes mueren"	1	Loaded language ("fascista", "mueren") and ideological bias

## 3 Population Replacement Conspiracy Theories (PRCT) Detection

### 3.1 Distinguishing PRCT from Racism

PRCT represents a specific family of conspiracy theories that claim the existence of a **deliberate**, **orchestrated**, and **coordinated** plan to replace the "native" population through immigration. It is important to distinguish PRCT from general racism or anti-immigration sentiment:

### 3.2 Key PRCT Theories and Concepts

PRCT encompasses several related conspiracy theories, including:

1. **The Great Replacement Theory:** Claims that Western populations are being deliberately replaced through mass migration
2. **The Kalergi Plan:** Alleges a conspiracy to mix European races with African and Asian races through migration

Characteristic	PRCT	General Racism
Conspiracy	Asserts the existence of a <b>coordinated plan</b>	Does not necessarily imply conspiracy
Orchestrators	Identifies specific <b>orchestrators</b> (elites, governments, NGOs, etc.)	May not identify specific orchestrators
Intentionality	Emphasizes the <b>intentionality</b> of replacement	May express prejudice without implying intentional planning
Language	Uses terms like "invasion," "replacement," "plan"	May use stereotypes and insults without replacement concept
Demographic vision	Presents an <b>apocalyptic view</b> of demographic changes	May express hostility without demographic predictions

3. **Eurabia:** Claims a conspiracy to Islamize Europe through demographic change
4. **White Genocide Theory:** Asserts that immigration is being used to eliminate white populations

### 3.3 Additional PRCT Indicators and Dog Whistles

Beyond those already mentioned, watch for:

- References to "demographic suicide," "demographic winter," or "demographic collapse"
- Phrases like "they're not refugees, they're an army" or "sleeping jihadists"
- Claims about "population engineering" or "demographic engineering"
- Suggestions of "breeding campaigns" or references to migrant birth rates as weapons
- Terminology like "imported voters" or "imported electorate"
- References to "global elites" or "globalists" orchestrating migration flows
- Mentions of migration as a "trojan horse" for cultural replacement

### 3.4 Labels

#### Label 0 (Non-PRCT):

- Content that discusses immigration in neutral terms
- Criticisms of immigration policies without conspiracy elements
- Discussion of demographic changes without suggesting orchestrated plans
- Reports of criminal activities or problems without suggesting conspiratorial schemes
- Use of terms like "crisis" or "emergency" without implying orchestrated replacement

#### Label 1 (PRCT):

- Content that explicitly mentions replacement theories
- Use of conspiracy-related language suggesting coordinated plans
- Portrayal of immigration as part of a deliberate replacement strategy
- Claims about migrants or their children being agents of intentional societal destruction
- Suggestions of hidden agendas by political parties or organizations to alter demographics through migration

### 3.5 Examples of PRCT Dog Whistles

1. **Explicit:** "The Kalergi Plan in action: ethnic replacement through migration"
2. **Military metaphors:** "Army of sleeping jihadists: migrants ready to strike"
3. **Hidden agenda:** "Left-wing parties pushing for citizenship to secure migrant votes"
4. **Destruction narrative:** "Their children hate us and will destroy us from within"
5. **Coordinated plan:** "NGOs caught in the act: planned operation to bring migrants only to our country"

### 3.6 Borderline Cases with Explanations

#### 3.6.1 Examples of Racist Content (Label 0 for PRCT)

- "I migranti africani sono pigri e non vogliono lavorare" (African migrants are lazy and don't want to work)
- "Gli stranieri portano criminalità e degrado nei nostri quartieri" (Foreigners bring crime and degradation to our neighborhoods)
- "Los inmigrantes ilegales son una carga para nuestro sistema de salud" (Illegal immigrants are a burden on our health system)
- "La mayoría de los robos son cometidos por extranjeros" (Most thefts are committed by foreigners)

**Reason:** These headlines express racist prejudices but do not contain elements of organized conspiracy or deliberate replacement plan.

#### 3.6.2 Examples of Clear PRCT Content (Label 1)

- "Piano Kalergi in azione: le élite europee orchestrano la sostituzione etnica" (Kalergi Plan in action: European elites orchestrate ethnic replacement)
- "Soros finanzia ONG per accelerare la Grande Sostituzione in Europa" (Soros finances NGOs to accelerate the Great Replacement in Europe)
- "L'islamizzazione dell'Italia è il vero obiettivo dietro l'accoglienza dei migranti" (The Islamization of Italy is the real objective behind migrant reception)
- "El Plan Kalergi avanza: la sustitución de los europeos es inminente" (The Kalergi Plan advances: the replacement of Europeans is imminent)

- "La bomba demográfica musulmana cambiará Europa para siempre según el plan establecido" (The Muslim demographic bomb will change Europe forever according to the established plan)
- "Eurabia: así se está consumando el reemplazo de la población española" (Eurabia: this is how the replacement of the Spanish population is being consummated)

These headlines contain explicit references to orchestrated plans, conspiracy, and intentionality of demographic replacement.

### 3.6.3 Borderline Cases

- "L'immigrazione sta cambiando il volto demografico dell'Europa" (Immigration is changing the demographic face of Europe)
- "La identidad europea en peligro por la inmigración masiva" (European identity in danger due to mass immigration)

**Decision:** Label 0 (non-PRCT) - Although they express concern about demographic changes, they don't contain the conspiratorial element of an orchestrated plan.

- "Bomba demografica: in 20 anni gli italiani saranno minoranza nel proprio paese" (Demographic bomb: in 20 years Italians will be a minority in their own country)
- "En 2050 los europeos serán minoría en su propio continente" (By 2050 Europeans will be a minority on their own continent)

**Decision:** Label 1 (PRCT) - Even without explicitly mentioning a plan, they use apocalyptic narrative typical of PRCT theories and imply an irreversible process of replacement.

## 4 Rhetorical Bias Detection and Span Annotation

### 4.1 General Principles

The goal is to identify the most precise textual elements that convey bias, focusing on specificity and contextual understanding.

### 4.2 Loaded Language Span Detection

#### 4.2.1 Definition

Loaded language involves words or phrases that have emotional or evaluative connotations beyond their literal meaning, designed to influence the audience's perception.

#### 4.2.2 Span Selection Criteria

##### 1. Single Words

- Emotionally charged verbs: "indoctrinate", "manipulate", "brainwash"
- Strongly negative or positive adjectives: "stupid", "evil", "heroic"



- Inflammatory nouns: "traitor", "fanatic", "zealot"

## 2. Phrases

- Contextually charged expressions where the bias is not in a single word
- Example: "Gay people tired us with so many celebrations for their rights"
- Span: The entire phrase conveys a loaded, dismissive attitude

## 3. Multiple Words

- When multiple words together create a biased sentiment
- Example: "Why I hate the democristians who eat babies"
- Spans: "hate", "eat babies"

### 4.2.3 Exclusion Criteria

- Neutral language without emotional charge
- Literal descriptions without evaluative intent
- Generic statements that don't carry inherent bias

## 4.3 Appeal to Fear Span Detection

### 4.3.1 Definition

Appeal to fear involves language that attempts to create or exploit fear to promote or reject an idea, manipulating audience emotions.

### 4.3.2 Span Selection Criteria

#### 1. Direct Fear Promotion

- Phrases that explicitly connect an idea with a threatening outcome
- Words that suggest imminent danger or catastrophic consequences

#### 2. Emotional Manipulation

- Language that generates anxiety or dread about a specific concept
- Spans that project worst-case scenarios

### 4.3.3 Specific Examples

- "Why I'm scared of the communists" - No proper span, as it doesn't promote a specific fear
- "If we don't stop immigration, our country will be destroyed" - Span: "will be destroyed"
- "Radical environmentalists want to take away your freedom and prosperity" - Spans: "take away your freedom", "prosperity"

#### 4.3.4 Exclusion Criteria

- Mere mention of fear without manipulative intent
- Personal expressions of fear without broader emotional manipulation
- Neutral discussions of potential risks

### 4.4 Name Calling Span Detection

#### 4.4.1 Definition

Name calling involves using language that demeans, labels, or categorizes individuals or groups in a derogatory manner.

#### 4.4.2 Span Selection Criteria

##### 1. Nouns

- Derogatory group labels
- Pejorative terms describing individuals
- Examples: "snowflake", "fascist", "bigot"

##### 2. Adjectives

- Descriptive terms used to belittle or criticize
- Examples: "lazy", "stupid", "radical"

#### 4.4.3 Specific Examples

- "Those lazy liberals always want handouts" - Spans: "lazy", "liberals"
- "Right-wing extremists are destroying our democracy" - Spans: "right-wing", "extremists"

#### 4.4.4 Exclusion Criteria

- Neutral descriptive terms
- Academic or statistical categorizations
- Terms used without clear derogatory intent

## 5 Stance Detection

Since the dataset focuses on one topic and includes headlines from more than 200 media outlets, stance detection will be performed using the following labels towards immigration policies: **pro**, **against**, **neutral**.

**Warning:** While PRCT content is often against immigration policies due to the nature of the phenomenon investigated, hyperpartisan news can be both right-wing or left-wing.

## 6 Annotation Process

1. **Read the Headline Carefully:** Ensure you understand the context and wording of the headline.
2. **Evaluate for Hyperpartisanship:** Check for ideological content and explicit mentions against entities/institutions. Pay attention to emotionally charged language, ideological bias, or attempts to influence opinion. Assign **Label 0** if the content is neutral and **Label 1** if it is hyperpartisan.
3. **Evaluate for PRCT:** Look for explicit mentions of replacement theories, conspiracy-related terms, or suggestions of deliberate demographic changes. Assign **Label 0** if the content is neutral and **Label 1** if it contains PRCT elements.
4. **Evaluate for Loaded Language:** Detect explicit words/adjectives/adverbs that are sentimentally loaded/exaggerated. Select the words affected by loaded language. Assign **Label 0** if the content is neutral and **Label 1** if it contains loaded language elements.
5. **Evaluate for Appeal to Fear:** Look for explicit or implicit messages promoting or rejecting an idea by exploiting the audience's revulsion or fear. Select the words affected by appeal to fear. Assign **Label 0** if the content is neutral and **Label 1** if it contains appeal to fear elements.
6. **Evaluate for Name Calling:** Identify when names or adjectives are given to an individual, institution, or group with the intent to denigrate or question their authority. Select the words affected by name calling. Assign **Label 0** if the content is neutral and **Label 1** if it contains name calling elements.
7. **Evaluate for Stance Detection:** Based on the content of the headline and the previously assigned biases, assign one of the following labels: **pro**, **against**, or **neutral** towards immigration policies.
8. **Double-Check:** Ensure that the labels align with the definitions and examples provided.

## 7 Additional Notes

### 7.1 Language-Specific Nuances

- In **Italian**, terms like "invasione" or "sostituzione etnica" are strong indicators of PRCT.
- In **Portuguese**, expressions like "Grande Substituição" or "destruir a nação" are clear signals.
- In **Spanish**, words like "invasión" or "reemplazo" should be carefully evaluated.

### 7.2 Cultural Context

Some conspiracy theories (e.g., "Kalergi Plan") are more common in certain linguistic contexts. Annotators should be aware of these differences.

## 7.3 Critical Mentions

Even if a headline critically discusses PRCT, it should still be labeled as **1** for PRCT.

## 7.4 Consistency

Ensure that similar headlines are labeled consistently across the dataset.

# 8 Frequently Asked Questions (FAQ)

## 8.1 Can a text be both hyperpartisan and contain PRCT elements?

**Answer:** Yes. A headline can be labeled as **1** for both Hyperpartisan and PRCT if it exhibits ideological bias (Hyperpartisan) **and** includes explicit mentions of population replacement conspiracy theories (PRCT).

**Example:** "La sinistra sta usando l'immigrazione per sostituire la popolazione: è il piano Kalergi!"

- Hyperpartisan: 1
- PRCT: 1

## 8.2 What if a headline uses neutral language but the newspaper is known for spreading hyperpartisan news or promoting PRCTs?

**Answer:** The label should be based solely on the content of the headline itself, not on external sources or context. To prevent this kind of bias, annotators will be provided only with the headlines.

## 8.3 How should we handle sarcasm or irony in headlines?

**Answer:** Sarcasm or irony can be challenging. If the sarcasm or irony is used to promote a hyperpartisan or PRCT narrative, label it accordingly. If it is neutral or critical of such narratives, label it as **0**.

**Example:** "Claro, a imigração é ótima... para quem quer ver o país destruído!"

- Hyperpartisan: 1
- PRCT: 1

**Explanation:** The sarcastic tone promotes a hyperpartisan and PRCT narrative.

## 8.4 What if a headline is ambiguous or unclear?

**Answer:** When in doubt, default to **Label 0** for both tasks. Only assign **Label 1** if there is clear evidence of hyperpartisanship or PRCT elements.

**Example:** "Immigrazione: un problema complesso che richiede soluzioni urgenti."

- Hyperpartisan: 0
- PRCT: 0

**Explanation:** The headline is neutral and does not explicitly show bias or conspiracy elements.

## 8.5 How should we handle headlines that discuss immigration but do not explicitly mention PRCT or hyperpartisanship?

**Answer:** If the headline discusses immigration in a neutral or policy-focused way, label it as **0** for both tasks.

**Example:** "Migranti, 2.500 sbarchi e 105mila arrivi nel 2022."

- Hyperpartisan: 0
- PRCT: 0

**Explanation:** The headline provides factual information without bias or conspiracy elements.

## 8.6 What if a headline is written in a sensationalist style but does not explicitly promote a conspiracy theory?

**Answer:** If the sensationalism is used to promote a specific ideological agenda, label it as **1** for Hyperpartisan. If it does not mention PRCT elements, label it as **0** for PRCT.

**Example:** "¡Crisis migratoria fuera de control! ¿Hasta cuándo lo permitiremos?"

- Hyperpartisan: 1
- PRCT: 0

**Explanation:** The headline uses sensationalist language to promote a specific agenda but does not mention PRCT.

## 8.7 What if a headline contains quotes?

**Answer:** Quotes inside headlines are considered as headlines. Thus, if they contain hyperpartisan or PRCT elements, label them as 1, otherwise 0.

## 8.8 Should we leverage external knowledge to detect hyperpartisanship when the headline appears neutral?

**Answer:** The judgment of a headline as hyperpartisan depends on the subject's political leaning and knowledge of the facts.

**Example:** "Robles responde al PP tras su propuesta sobre inmigración: 'Respeten a las Fuerzas Armadas'"

This headline appears neutral if you don't know the context. How you apply your external knowledge to detect hyperpartisanship framing depends on your political bias, which is a

variable we want to analyze during this experiment. We are keeping the disagreement for this reason as well.

To summarize: style (loaded adjectives/verbs) is not the only means to communicate hyperpartisanship.

## 8.9 How should quotes be considered in stance detection?

**Answer:** Like in hyperpartisan detection, if a headline contains quotes, we shift our attention to the content of the reported speech. Usually quotes occupy the largest part in headlines, becoming the central argumentation.