

MultiCaption



UK Research
and Innovation



Funded by
the European Union

Usage/License



MultiCaption is released under the **Creative Commons Attribution 4.0 International License (CC BY 4.0)**. Under this license, users are permitted to share and adapt the material provided appropriate credit is given to the original authors. However, access to the dataset is **restricted to research purposes**. Users must comply with the research-only access conditions specified by the dataset providers.

Download



The MultiCaption datasets are available for download via Zenodo:
<https://zenodo.org/records/18230659>

Please refer to the repository for detailed documentation, dataset structure, and access instructions. Access to the full dataset requires prior agreement to specific terms and conditions.

Researchers



- Rafael Martins, Frade
- Rrubaa Panchendrarajan
- Arkaitz Zubiaga

These researchers are affiliated with Newtral Media Audiovisual (Spain) and Queen Mary University of London and, collaborating on multilingual automated fact-checking research.

Summary

MultiCaption is a multilingual dataset designed to support misinformation detection by identifying **contradicting captions associated with the same visual content**. The dataset contains **11,088 visual claims across 64 languages**.

Each instance consists of two captions describing an image or video. Pairs of claims referring to the same image or video were labeled through multiple strategies to determine whether they contradict each other. Each pair is assigned a binary label indicating whether the captions are contradictory (1) or not contradictory (0).

What We Offer

- A Structured resource for detecting contradictory visual claims.
- Multilingual caption pairs suitable for text-only and multimodal modeling.
- A Scalable Dataset supporting cross-lingual misinformation research.

Key Features

- 11,088 Labeled Caption Pairs linked to shared visual content.
- 64-Language Coverage enabling multilingual and cross-lingual analysis.
- Binary Contradiction Labels for supervised learning.
- Designed for Misinformation Detection focusing on conflicting narratives.
- Multimodal-Ready Structure through image/video references.

Collaboration Objectives

The MultiCaption project aims to:

- Advance research in multilingual and multimodal contradiction detection.
- Facilitate automated identification of misleading visual narratives.
- Encourage robust cross-lingual misinformation analysis.