

# IAT2BAS Dataset

From Inference Anchoring Theory-annotated conversations to Bipolar Argument Structures Dataset



Funded by  
the European Union



UK Research  
and Innovation

## Usage/License



This work uses publicly available datasets from the Argument Interchange Format Database (AIFdb: <https://corpora.aifdb.org/>). Our adapted datasets are released under the Creative Commons Attribution 4.0 License (CC BY 4.0). Users should check the original dataset licenses for any additional terms.

## Download



The dataset and its documentation can be found on our Github repository: [github.com/The-obsrvr/IAT-BAS-Data-Pipeline](https://github.com/The-obsrvr/IAT-BAS-Data-Pipeline).  
HYBRIDS ZENODO: <https://doi.org/10.5281/zenodo.19627671>

## Researchers



- Siddharth Bhargava, FBK / UDC
- Sara Tonelli, FBK
- Patricia Martín Rodilla, UDC
- Javier Parapar López, UDC

## Summary

The dataset comprises publicly available Inference Anchoring Theory (IAT)-annotated corpora restructured into Bipolar Argument Structures (BAS). It currently includes three diverse corpora, though the framework is extensible to all IAT corpora in the Argument Interchange Format (AIF).

These corpora enable benchmarking and computational modeling for argument structure prediction in multi-party dialogues, grounded in the IAT framework.

## What We Offer

- A large and diverse collection of IAT-annotated corpora re-adapted into Bipolar Argument Structures (BAS) for benchmarking and computational modeling.
- An end-to-end pipeline converting raw IAT annotations into structured argument representations, extendable to all AIF corpora.
- A framework for structured discourse analysis to study argument evolution in multi-party conversations.

## Key Features

- Structured argument representations derived from IAT-annotated conversations, consisting of span-level argument units connected through support and attack relations.
- Fine-grained reasoning at the argument-unit level, grounded in the original Inference Anchoring Theory (IAT) annotations.
- Diverse corpora encompassing offline and online interactions, including political and non-political discussions, and both persuasive and collaborative dialogue settings.

## Collaboration Objectives

- Facilitates development of downstream tasks such as stance detection, text summarization, and misinformation detection
- Promote interdisciplinary research across argumentation mining, computational linguistics, conversational modeling, and discourse analysis.
- Encourage the development of analytical tools and software for social media analysis, multi-party interaction analysis, and fostering critical thinking.

