

Augmented OUTFOX Dataset



Funded by
the European Union



UK Research
and Innovation

Usage/License



The original OUTFOX dataset is used under its Apache 2.0 license. Aeon-essays is used under its MIT license. Open models used for augmentation are available under their respective licenses: Gemma Terms of Use (Gemma 3), Apache 2.0 (Apertus), MRL (Minstral 3). Use of the OpenAI API for ChatGPT 4o and 5.2 complies with the OpenAI Terms of Use.

Download



The augmented dataset with clause-level genericity and RST annotations is available at: https://github.com/SorenKF/general_machines

Researchers



- Søren Fomsgaard
- Martial Pastor
- Gael Dias
- Nelleke Oostdijk

These researchers are affiliated with Universite de Caen Normandie and Radboud University.

Summary

The Augmented OUTFOX Dataset is a large-scale collection of 215,600 argumentative essays comprising human-written texts and LLM-generated texts from 12 models, annotated with clause-level genericity labels and RST discourse relations. The dataset extends the original OUTFOX corpus (Koike et al., 2024) with 123,200 new synthetic texts generated from recent open and proprietary models, providing a comprehensive resource for studying the linguistic characteristics of machine-generated text.

What We Offer

- Augmented texts: 123,200 LLM-generated essays using OUTFOX prompts and recent models.
- Dual-layer annotations: Clause-level genericity + document-level RST discourse trees.
- Cross-population validation: Validation on Aeon-essays (2,235 expository texts by 1,655 adult expert writers), confirming generalization across writers and genres.

Key Features

- Multi-model coverage: Texts from 12 LLMs (old and recent) plus 15,400 human essays (PERSUADE 2.0).
- Clause-level annotations: 17 labels on genericity, eventivity, and boundedness using a trained RoBERTa + bi-GRU model.
- RST discourse annotations: Full discourse trees with 18 coherence relations generated via a DMRST parser.
- Scale: 215,600 texts, enabling robust cross-model analysis.

Collaboration Objectives

- Provide a benchmark to study differences between human and LLM-generated argumentative text.
- Enable misinformation research through discourse and genericity annotations.
- Support cross-model analysis of discourse structure and rhetorical strategies.
- Encourage extensions to multilingual settings and new genres.