News is More than a Collection of Facts

Moral Frame Preserving News Summarization

Find me on LinkedIn!

Enrico Liscio

Michela Lorandi

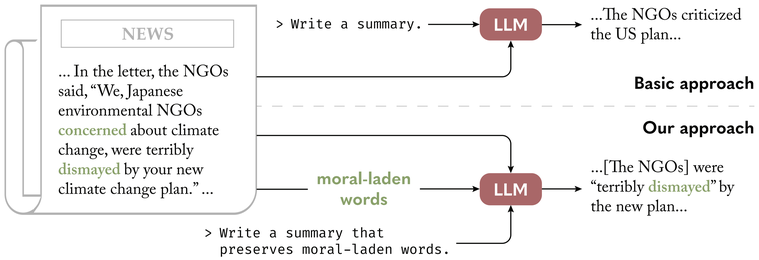
Pradeep K. Murukannaiah

Moral framing of news

Framing refers to **the way information is presented** to shape audience perceptions of an issue, e.g., by emphasizing some aspects of an event while downplaying others.

In this work, we focus on the **moral dimension of framing**, which evaluates actions, behaviors, or situations as right or wrong based on the underlying moral principles.

LLM-based summarization may overlook the article framing



Automated evaluation

Prompting the model to preserve moral-laden words **does not affect the overall summary quality**.

Llama-3-70B-Instruct is the best-performing tested model.

Supervised classification beats Chain-of-Thought in detecting moral-laden words (47.2 vs. 22.3 F_1).

Crowd evaluation

Classifier and **Oracle** are the best-performing methods.

Preserving moral-laden words leads to **higher evaluation scores**.

There is **no correlation** between crowd and automated evaluations (i.e., human evaluation is needed).

Expert evaluation

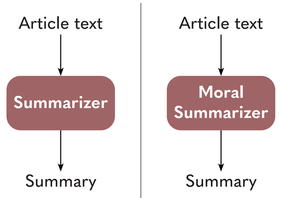
All moral-sensitive approaches **outperform the basic approach**.

Preserving a **spokesperson's words verbatim** is important.

LLMs can **add or modify** the moral framing of the article.

The five compared approaches to news summarization

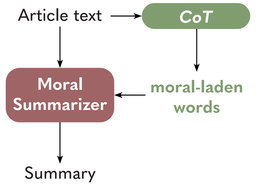
Basic prompts



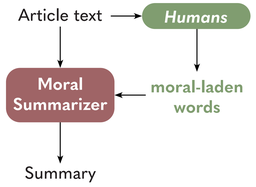
***Plain* approach:** The summarizer (a zero-shot LLM) is prompted to write a summary of the article.

***Direct* approach:** The summarizer is prompted to write a summary that "preserves the moral framing of the original article".

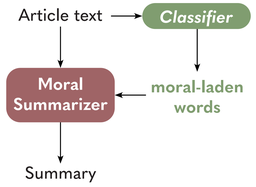
Word-preserving prompts



***CoT* approach:** The summarizer is first prompted to identify the list of moral-laden words and then to generate a summary that preserves them, in a Chain-of-Thought fashion.



***Oracle* approach:** The summarizer is provided with the list of words annotated by humans as moral-laden in the article and asked to preserve them in the summary.



***Class* approach:** A classifier is trained on human annotations to identify moral-laden words in a news article. The summarizer is then prompted to generate a summary that preserves the words identified by the classifier.

Takeaways

We can **improve moral framing preservation** while maintaing overall summary quality.

Better moral-ladenness prediction **does not correlate** with better moral framing preservation.

We don't know how the model decides **which moral-laden words** to preserve in the summary.

Human judgment is required to evaluate moral framing preservation.

Experts' judgment varies across approaches

Experts made pairwise comparisons of summaries generated by different approaches. This is the **distribution of labels of the experts' motivations** of summaries' differences.

Category	Label	*Plain*	*Direct*	*CoT*	*Oracle*	*Class*
Positive	Moral Framing Alignment	23.7%	20.4%	29.8%	13.9%	68.3%
	Quote Preservation	2.6%	8.2%	2.1%	11.1%	**17.1%**
	Examples Inclusion	0.0%	2.0%	0.0%	8.3%	0.0%
Negative	Moral Framing Loss	57.9%	34.7%	21.3%	41.7%	2.4%
	Quote Omission	2.6%	8.2%	12.8%	5.6%	0.0%
	Examples Omission	2.6%	0.0%	4.3%	0.0%	4.9%
	Moral Framing Modification	5.3%	**6.1%**	2.1%	0.0%	0.0%
	Moral Framing Addition	2.6%	**14.3%**	**23.4%**	16.7%	2.4%
Neutral	Similarity	2.6%	4.1%	2.1%	2.8%	4.9%

TU Delft AlgoSoc / Hybrid Intelligence DCU