



Natural Language Processing

Natural Language Processing

- Natural Language Processing is a **cross-disciplinary** research field that draws heavily from **artificial intelligence (AI)**, **machine learning (ML)**, mathematics, and linguistics.
- Personal assistants, recommender systems, fake news identification, financial stock analysis, chatbots, autocorrection, auto-completion, intelligent search engines, and automatic translation or captioning are just a few examples of how NLP and AI are helping us manage the flood of data. However, systems to process natural language are far from perfect, which leaves much space for research.
- Some of the areas we work are:
 - Natural language understanding
 - Paraphrase detection
 - Text summarization
 - Media bias/Fake news detection
 - Semantic analysis/extraction
 - Sentiment analysis

For a complete list of our research topics visit our [website!](#)



Paraphrase Types: Data and Task Generation

Background

Current paraphrase generation and detection systems are yet unaware of the lexical variables they manipulate. Generative models cannot be asked to perform certain types of perturbations, and detection models are unable to understand which paraphrase types they detect or learn limited language aspects (e.g., primarily syntax). The shallow notion of what composes paraphrases used by these systems limit their understanding of the task and makes it challenging to interpret detection decisions in practice. Thus, we need to leverage existing datasets and tasks used in Paraphrasing with more granular information so we can assess the problem better and develop more robust techniques.

Goal

- Extend current datasets used in paraphrase related tasks to include paraphrase types

Tasks

- Literature review on paraphrase types (atomic paraphrase types)
- Probe existing LLM to generate/classify pair sentences including selected paraphrase types (e.g., prompting, few-, or zero-shot) using the ETPC dataset as a reference
- Correlate (e.g., BLEU, similarity, ROUGE, BERTScore) generated paraphrase with existing data and select the best paraphrase types
- Extend the best paraphrase types to generate/classify new data from other paraphrase datasets
- Propose new tasks for the BIG-bench and/or GEM benchmarks based on Paraphrase Types



Jan Philip Wahle
wahle@gipplab.org



Terry L. Ruas
ruas@gipplab.org



The Paraphrase Type Taxonomy

Background

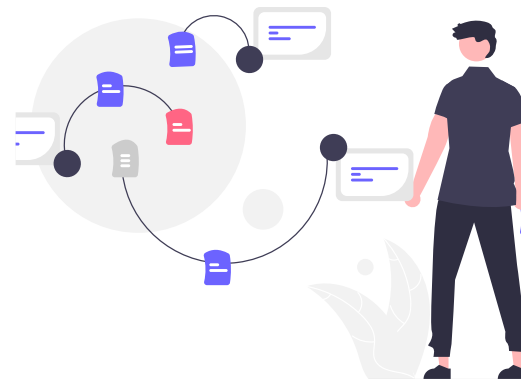
This project proposes an extensive literature review to identify and critically evaluate various paraphrase types that have been proposed in linguistic, computational, and educational domains. By synthesizing these diverse perspectives, the project aims to develop a cohesive framework that categorizes paraphrase types based on linguistic features, context, and communicative intent. Through rigorous analysis and categorization, the project aims to establish a comprehensive taxonomy of paraphrase types. Furthermore, the research team plans to develop an open-access online repository, where the findings and the framework will be made available to the public, promoting collaboration and further research in this domain.

Goal

- Investigate and (re)organize available taxonomies and language models used in paraphrase types

Tasks

- Investigate available taxonomies used in paraphrase (types)
- Critical evaluation of existing ones (agreement and disagreements between them)
- Investigate available models used in paraphrase generation and detection
- Propose a new taxonomy (with definitions, examples, and instructions) for paraphrase types (generation and detection)



Jan Philip Wahle
wahle@gipplab.org



Terry L. Ruas
ruas@gipplab.org



CS-Insights – State of the art in Computer Science Publications

Background

DBLP is the largest open-access repository of scientific articles on computer science and provides metadata associated with publications, authors, and venues. We retrieved more than 6 million publications from DBLP and extracted pertinent metadata (e.g., abstracts, author affiliations, citations) from the publication texts to create the DBLP Discovery Dataset (D3). Now, on [CS-Insights](#) we devised a system (back- and front-end) to explore our dataset and uncover all the trends regarding computer science publications. As [CS-Insights](#) is an ongoing project we need to fix its open issues and extend its functionalities.

Goal

- Solve existing issues in [CS-Insights-Roadmap](#)

Tasks

- Work on project roadmap for CS-Insights
 - Backlog and additional features
- Propose extension for CS-Insights
 - Authors features (e.g., h-index)



Jan Philip Wahle
wahle@gipplab.org



Terry L. Ruas
ruas@gipplab.org



Multi-Agent Reasoning with Large Language Models

Background

Advances in human collaboration, shown by our ability to communicate through natural language, have led to significant achievements. In contrast, current large language models (LLMs) like ChatGPT and LLaMA operate in isolation, without the benefit of interaction that mirrors human collaboration. Preliminary research indicates potential advantages when multiple LLM agents engage in dialogue, yet systematic exploration into how LLMs could communicate to make decisions and tackle reasoning tasks remains sparse. This gap in research highlights the need for an in-depth investigation into the dynamics of LLM interactions, including the roles of participants, discussion formats, and decision-making protocols.

Goal

- Explore the potential of discourse between LLMs in improving decision-making and task performance.

Tasks

- Conduct a thorough literature review focusing on the interaction between LLM agents, their roles in discussions, and the impact on task performance.
- Examine different discussion formats and their suitability for specific tasks, assessing how they influence the representation of viewpoints and conclusion of discussions.
- Develop and evaluate decision-making protocols, including consensus-building and voting mechanisms, to determine the most effective strategies for collective LLM decision-making.



Jan Philip Wahle
wahle@gipplab.org



Terry L. Ruas
ruas@gipplab.org



Making Large Language Models Safer

Background

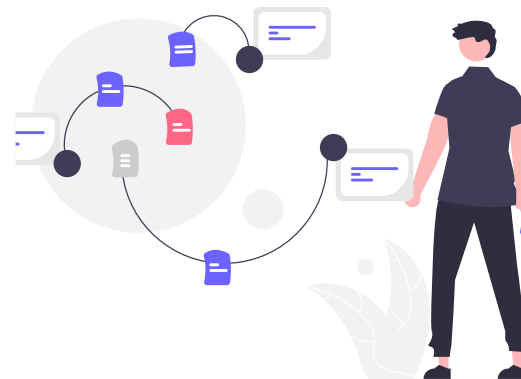
The rapid advancement and integration of Large Language Models (LLMs) into various sectors underscore the necessity for safety and ethical considerations in their development and deployment. Current challenges include ensuring bias control, privacy, non-discrimination, and robustness across legal, ethical, and technical dimensions. Furthermore, the security of AI models, their training and test data against manipulation, safeguarding against the extraction of sensitive information, and the need for secure operating environments for AI systems are crucial. The authentication of AI-generated outputs and the adaptation of existing AI technologies to meet the needs of public interest organizations also present significant areas for development and innovation.

Goal

- To enhance the safety, security, and ethical alignment of LLMs.

Tasks

- Develop and implement bias control mechanisms in LLMs, ensuring they are effective with smaller language corpora and computationally efficient.
- Incorporate privacy, non-discrimination, legal and ethical robustness, reliability, trustworthiness, interpretability, and explainability into LLM design and operation.
- Implement protective measures against manipulation of AI models, their training, and test data, as well as strategies to prevent sensitive data extraction from AI models.
- Create secure operational environments for AI systems that may not be fully safe, and develop methods for authenticating AI-generated results.



Jan Philip Wahle
wahle@gipplab.org



Terry L. Ruas
ruas@gipplab.org



Meeting Summarization System Testbench

Background

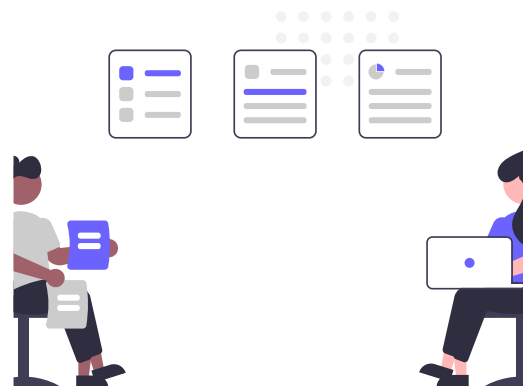
The field of natural language processing has seen a significant amount of research in recent years on the task of meeting summarization. With the increasing availability of meeting transcripts, there is a growing need for efficient methods to automatically summarize the content of these meetings. As of now, due to the different formats of meetings and the dynamic, idiosyncratic nature, many domain- and problem-specific techniques have been introduced. However, the area lacks a standardized benchmark for evaluating these methods. Thus, it is difficult to compare and identify the strengths and weaknesses of the individual techniques.

Goal

- Design and develop a unified framework to test meeting summarization techniques (evaluation harness).

Tasks

- Design a solution to transform any kind of input format for models and datasets into one common form
- Develop a functionality to automatically add noise to the input text to assess models' robustness
- Implement a general applicable evaluation environment to test different models, datasets and metrics simultaneously
- Evaluate the most prominent techniques



Frederic Kirstein
kirstein@gipplab.org



Terry L. Ruas
ruas@gipplab.org

