

# Abstract Brochure

Abstracts for the LCT Annual Meeting Poster Session

June 19-20, 2022

*This brochure contains abstracts for posters being presented at the LCT annual meeting of 2023, hosted by the University of Malta. The section numbers in this brochure correspond to the respective poster slots in the poster session.*



With the support of the  
Erasmus+ Programme  
of the European Union



## Contents

1	COMET for Low-Resource Machine Translation Evaluation	4
2	BART meets Macbeth: Summarization of Shakespeare’s plays with BART based models	5
3	Exploiting the Structure of HTML to Learn Document Representations	6
4	Multi-Modal fine-tuning to improve the readability of Automatic Speech Recognition outputs	7
5	RDF-DST: A case study of rich representations for Dialogue-State Tracking	8
6	Captioning artworks: What can AI say about paintings?	9
7	Generating Solutions to Morphology-Based Analogical Equations	10
8	Multilingual Coreference Resolution Using PromptEngineering for Mention Detection	11
9	Early-exit Transformers for resource-aware ASR	12
10	Cross-Lingual Transfer Learning with Task and Language Adapters for Multilingual Question Generation	13
11	Multi-Teacher Distillation for Pretraining a Low-Resource Language	14
12	Discourse Entity State Tracking with Explicit Memory	15
13	Temporal Reasoning in Vision and Language Models	16

<b>14 Symbolic Numerical Reasoning on Structured Tables</b>	<b>17</b>
<b>15 Compositional Discourse Representation Structure Parsing</b>	<b>18</b>
<b>16 Can GPT-4 Outperform Fine-tuned BERT Models in Relation Extraction Tasks within the Clinical Domain?</b>	<b>19</b>
<b>17 Role of intonation in intercomprehension scenario</b>	<b>20</b>
<b>18 Tackling misinformation in Serbo-Croatian: corpora and experiments</b>	<b>21</b>
<b>19 Analyzing Modular Cross-Lingual Transfer Learning</b>	<b>22</b>
<b>20 Discovering Emerging Entities Descriptions via Search Engine</b>	<b>23</b>
<b>21 Evaluating fairness in job recommendation system using fairness indicator parity</b>	<b>24</b>
<b>22 Learning sentence similarity with pre-trained LM</b>	<b>25</b>
<b>23 Utilization of Adapters in Pre-trained Speech Models for Parameter Efficient Automatic Speech Recognition</b>	<b>26</b>
<b>24 Quantifying and Mitigating Gender Bias in Chinese Word Embedding</b>	<b>27</b>
<b>25 OPEN INFORMATION EXTRACTION USING LARGE LANGUAGE MODELS</b>	<b>28</b>
<b>26 GPL at SemEval-2023 Task 1: WordNet and CLIP to Disambiguate Images</b>	<b>29</b>

# 1 COMET for Low-Resource Machine Translation Evaluation

**Author:** Júlia Falcão

**Presented by:** Júlia Falcão

## Abstract

Human judgement is widely regarded as the gold standard for measuring translation quality in Machine Translation, but with the rapid development of neural systems, simple and costeffective metrics such as BLEU have been almost exclusively used for evaluation instead. Researchers have been raising concerns about the shortcomings of these string-based metrics for years, and meta-evaluations have proven again and again that they fail to correlate with human judgements. As a response, neural metrics such as COMET have been proposed, aiming explicitly to learn and predict translation quality scores based on human judgements. These metrics have been performing impressively in the latest meta-evaluations, but these tasks mostly feature high-resource languages. We are interested in exploring how well COMET works with pairs including two under-resourced language pairs: Maltese-English and Basque-Spanish. We plan on experimenting with COMET out-of-the-box to see how well it generalizes to unseen languages, and furthermore, we aim to extend it by attaching Maltese and Basque multilingual Language Models and training on parallel data. We will evaluate the results by obtaining our own human evaluations on a small set of translations, and calculating how well COMET correlates with them in comparison to other metrics.

## 2 BART meets Macbeth: Summarization of Shakespeare’s plays with BART based models

**Authors:** Rishu Kumar, Katja Konermann, Jingyan Chen, Thierry Declerck

**Presented by:** Rishu Kumar

### Abstract

With the rise in multilingual models and their increasing capacity in multiple downstream tasks, this study explores a recent state-of-the-art model for the downstream task of summarization. We explore different summarization techniques, namely abstractive and dialogue summarization. Dialogues pose a special challenge to the task of summarization as their format differs quite a bit from other types of text. Especially in plays, important actions and crucial pieces of information are often only stated indirectly or not in the dialogue itself but in short stage directions. Additionally, summaries of dialogues require a large amount of rewording as well as rephrasing. For example, instances of the pronoun in 1st person usually have to be replaced by a 3rd person pronoun. For this study, we worked with Shakespeare plays in their original English version and German translations of them. This process included creating our own dataset by manually mapping chunks of the plays to short summaries. The antiquated language in these plays presents a further challenge, mainly because it has to be translated into present-time English for the summaries. Our approach tests the language-agnostic vocabulary space of BART-based multilingual models by providing input in Old English, New English and German. We further fine-tune our models for dialogue summarization with the SAMSUM dataset in English and its machine-translated equivalent in German as well as our own hand-crafted datasets.

### **3 Exploiting the Structure of HTML to Learn Document Representations**

**Authors:** Guenter Neumann (UdS), Christophe Cerisara (UL), Lara Perinetti (Qwant), Adelia Khasanova

**Presented by:** Adelia Khasanova

#### **Abstract**

Information Extraction (IE) is a crucial sub-field of Natural Language Processing. It comprises various business applications obtainable in natural language and this encoded information plays a significant role in business decisions. IE in distinction to these documents provides extracted features to carry out groundwork for further research: in making efficient and feasible business solutions. For years, various multimodal pretraining techniques have been proposed, and this work attempts to implement transfer learning using those models to smaller neural network layers. Furthermore, this thesis incorporates the impact of linguistic features when trained explicitly in IE.

## 4 Multi-Modal fine-tuning to improve the readability of Automatic Speech Recognition outputs

**Authors:** Md Zobaer Hossain

**Presented by:** Md Zobaer Hossain

Due to the recent introduction of deep neural network models, the performance of automatic speech recognition has significantly improved. Although Automatic Speech Recognition (ASR) models can correctly recognize words, their final output typically lacks capitalization and punctuation. Post-editing techniques are found to be highly effective in addressing such issues and can significantly improve readability performance. However, the majority of post-editing models are not yet robust to ASR errors, and these models are primarily designed for only restoring punctuation and do not consider restoring capitalization. In this study, we proposed a multi-modal fine-tuning method for improving the readability of ASR outputs through post-editing. The key idea of the method is to use noisy training outputs from ASR models to fine-tune encoder-decoder-based large pre-trained transformer models such as T5 in order to predict the sentences with accurate capitalization and punctuation. During the evaluation, we found the effectiveness of our proposed method to improve the readability of ASR outputs, particularly in the case of low-resource language.

## 5 RDF-DST: A case study of rich representations for Dialogue-State Tracking

**Authors:** Andres Gonzalez

**Presented by:** Andres Gonzalez

### Abstract

The flow of a task-oriented dialogue between a virtual assistant and a human user is dependent on intentionality, i.e. the user’s goals. Unlike open dialogue systems, a user may have a precise goal such as booking a restaurant or calling 911, and as such, require precise information to address the user’s issues. Dialogue State Tracking (DST) is a task that attempts to track this information with slot-values, e.g. (Slot=Restaurant, Value=Chinese). Although current solutions can accurately predict this information, these annotations are usually not enough for a system to learn what action to take in a multi-turn, multi-task, and multi-domain dialogue context. In this research project, we have introduced a richer annotation scheme that includes additional knowledge ranging from what information should the system be looking for, what has been found, and what actions have been taken so far. All of this has been annotated in the form of RDFs and our goal is to leverage large language models such as T5 to correctly generate these sequences. We have used a T5 slot-value baseline to compare our results with our T5 RDF annotations. Such analysis is crucial to gauge whether costlier annotations and representations, which equate to costlier training, yield better actions in task-oriented dialogue systems with virtual assistants.



## 6 Captioning artworks: What can AI say about paintings?

**Author:** Joseph KEENAN, Karolin Izabel BOCZON, and Luis Antonio VASQUEZ REINA

**Presented by:** Luis Antonio VASQUEZ REINA

### **Abstract**

Image captioning on artworks is a relatively unexplored area. Our project researches this area in order to extend the work of Lamiroy and Potier (2022), LaMuse, an artwork style transfer model, into the textual domain. We design a pipeline which separates image processing and text generation. The former consists of object recognition, color detection, and emotion classification. The obtained information is then preprocessed and fed into our own fine-tuned language generation models. We also implement a rule-based approach for comparison with the fine-tuned models. Although we do not obtain the best results from the evaluation, our models are not far from the state-of-the-art. Through several advancements in the future work we aim to improve our results and we plan on integrating our pipeline with LaMuse.

## 7 Generating Solutions to Morphology-Based Analogical Equations

**Authors:** Shane Kaszefski

**Presented by:** Shane Kaszefski

### Abstract

”Analogical proportions are statements of the form “ $A$  is to  $B$  as  $C$  is to  $D$ ” hereafter denoted as  $A : B :: C : D$ . A morphological analogy, therefore, is an analogical proportion at the morphological level. For example: *easy : easier :: sunny : sunnier*. To “solve” an analogy, given  $A$ ,  $B$ , and  $C$ , one needs to find a solution  $X$  such that  $A : B :: C : X$  is a valid analogy. The work presented is an attempt at solving morphological analogies through generation, rather than relying on retrieving the solution from a dictionary of possible solutions. Preliminary experiments with a novel autoencoder architecture show promising results for some languages and reveal the feasibility of the approach in generating solutions of analogical equations in the morphology setting.”

## 8 Multilingual Coreference Resolution Using PromptEngineering for Mention Detection

**Author:** Anna Mokhova

**Presented by:** Anna Mokhova

### Abstract

Generative NLP models have recently become the center of attention for an enormous variety of tasks worldwide. Prompt learning, or prompt engineering, that originates from generative architectures, hence, is a promising concept of adapting pre-trained language models for solving downstream NLP by modifying input text with textual templates. On the other hand, multilingual coreference resolution, defined as a task for clustering mentions referred to the same entity in multiple languages, has received a new wave of popularity after the coreference community proposed a universal annotation scheme for harmonizing existing data. This work focuses on solving multilingual coreference resolution in the end-to-end setting as a part of CRAC-2023 Shared Task and suggests using prompt engineering for mention detection. The motivation behind this approach is straightforward: we propose that additional mention labeling will increase the performance for coreference resolution. The intermediate results include mention detection performance using prompt-learning for 14 languages as well as a discussion of possible directions for future research. The presentation, apart from the above-mentioned, highlights the overview of the research review, and relevant concepts, and specifies linguistic explanations for the stated problem.

## 9 Early-exit Transformers for resource-aware ASR

**Authors: George August Wright**

**Presented by: George August Wright**

### **Abstract**

In an edge-cloud continuum infrastructure, the workload of model training and inference for automatic speech recognition (ASR) is distributed across a variety of devices with variable computational resources. An ideal ASR model should dynamically adapt its computational load to fit the capabilities of the device it runs on. We present the early-exit architecture as a framework for training a dynamic Transformer-based ASR model, in which input only traverses a subset of the total number of layers in the model, saving time and computational resources. Conformer and Wav2Vec 2.0, two Transformer-based ASR models, are respectively trained and retrained using an early-exit loss, which is the joint loss computed on the decoded output of every other layer in the model. Knowledge distillation is also incorporated into the early-exit training through a teacher-student architecture. We find the early-exit training significantly improves ASR word error rate at intermediate layers of the models as compared to typical single-exit training, indicating its effectiveness at training better-performing small-size models. Furthermore, the Conformer model, trained from scratch with early exits, outperforms and converges faster than a single-exit-trained Conformer, suggesting that the early-exit loss regularizes training.

## 10 Cross-Lingual Transfer Learning with Task and Language Adapters for Multilingual Question Generation

**Authors:** Silviya Silwal

**Presented by:** Silviya Silwal

### Abstract

As chatbots become central to online services, the need for their improved understanding and addressing of user queries intensifies. This research focuses on the creation and evaluation of a question generation system that enhances the performance of chatbots. Leveraging the principles of Question Answering (QA) in Natural Language Understanding, the study aims to invert the conventional QA process by generating questions from answers and relevant text. The objective is to assess a machine's reading comprehension by generating questions that mirror the comprehension of a given corpus.

The research examines diverse strategies for creating questions from a documentary corpus for Synapse's chatbot system. For the DearBot solution, generating queries from the

corpus holds potential for considerable enhancement of the chatbot's knowledge base and overall performance. This technique will capacitate the chatbot to handle novel user queries that resemble pre-existing ones, leading to heightened efficiency and customer satisfaction. Moreover, the question generation system will identify gaps in the knowledge base, fostering further content creation or research.

The study employs adapter modules inspired by the MAD-X framework, training interchangeable language and task-specific adapters atop the mT5 base model for cross-lingual

question generation. This research offers significant insights into improving chatbot performance and broadening its functional repertoire, thereby enhancing user experience.

## 11 Multi-Teacher Distillation for Pretraining a Low-Resource Language

**Author:** Amit Kumar Chaudhary

**Presented by:** Amit Kumar Chaudhary

### **Abstract**

We aim to study the effect of pre-training masked language models for low-resource languages through the use of knowledge distillation from related high-resources languages. Specifically, we will evaluate the effectiveness of this pre-training approach in the Maltese language, with Arabic, English, and Italian serving as the teacher language models. The results will help us understand how the performance of more specific multilingual models compares to the performance of massively multilingual models on low-resource languages.

## 12 Discourse Entity State Tracking with Explicit Memory

**Author:** Qiankun Zheng

**Presented by:** Qiankun Zheng

### **Abstract**

Entities play an important role in human narratives. The changing state of entities influences the flow and establishes the logic of narratives, thus making them more coherent. Tracking entity state changes in discourse is essential for long-context understanding. Recent work has shown that the entity state tracking capabilities of Transformer-based models are far from satisfactory, suggesting that there is much room for improvement. To fill this gap, we develop an external memory on top of the T5 model to store and update entity representations. Our approach is expected to help the model capture long-range dependencies in action sequences, thus facilitating the model's ability to track entity state changes.

## 13 Temporal Reasoning in Vision and Language Models

**Author:** Andrew McIsaac

**Presented by:** Andrew McIsaac

### **Abstract**

Vision and Language Models show reasonable performance on the task of video question answering, but it is unclear whether the ability to answer questions that seem to rely on temporal reasoning comes from a model’s learned ability to reason across time, or whether multiple choice questions lead to other factors such as object detection providing a stronger signal to the model. We explore one model, Merlot Reserve, and show that it is unable to determine the correct answer in before/after questions above chance. We hypothesise that this comes in part from the contrastive pretraining objective, as has been shown in previous work on verb understanding and compositional relationships. To explore, we finetune this model with additional hard negatives explicitly related to temporal understanding and evaluate performance on temporal-aware subsets of video question answering datasets.



## 14 Symbolic Numerical Reasoning on Structured Tables

**Author:** Ekaterina Garanina

**Supervisors:** Ondřej Dušek, Ph.D. (Charles University in Prague)  
Gertjan van Noord, Prof Dr. (University of Groningen)

**Presented by:** Ekaterina Garanina

### Abstract

Table-to-text generation is a sub-task of Natural Language Generation (NLG) that aims at generating text in natural language from structured tables, sum-marizing the main or most interesting pieces of information contained therein (Bao et al., 2018; Liu et al., 2018). While multiple works have been addressing this task in the past few years, mostly using pre-trained language models (Chen et al., 2020a; Liu et al., 2022), it still poses significant problems, mainly (1) pre-selecting the most important content (table cells) to be verbalized, and (2) making language models work consistently with numeric operations over the table, which are often needed for summarizing statements. My work addresses both problems by generating sentences with logical inference (Saha et al., 2022), using the raw table as input. I explore whether a pipeline-based approach using separate planning/reasoning and surface realization steps (Gatt and Krahmer, 2018) is superior to end-to-end language-model-based generation (Kale and Rastogi, 2020) in terms of accuracy and consistency. My approach predominantly focuses on the planning and reasoning steps using a formal semantic representation, while the surface realization step makes maximum use of existing approaches. The project is based on Logic2Text (Chen et al., 2020b) and LogicNLG (Chen et al., 2020a) datasets. For evaluation, I employ automatic metrics and conduct a small-scale human evaluation. This work is currently in progress, and I anticipate that the pipeline-based approach with symbolic reasoning will substantially outperform an end-to-end model in terms of fidelity to the input, while maintaining the same high level of language fluency.

## 15 Compositional Discourse Representation Structure Parsing

**Author:** Xiulin Yang

**Supervisor:** Professor Alexander Koller (Saarland University) and Professor Johan Bos (University of Groningen)

**Presented by:** Xiulin Yang

### Abstract

Sequence-to-sequence models have shown impressive performance in semantic parsing tasks. However, they often struggle with compositional generalization, prompting the need for the exploration of compositional parsers. In line with Montague’s principle of compositionality, this project aims to investigate approaches to compositional Discourse Representation Theory (DRT) parsing, to take advantage of the strong expressive power of DRT in comparison to other meaning representation frameworks. More specifically, our focus is on parsing a new representation format called Simplified Box Notation (SBN), known for its graph format and variable-free features. To maintain the compositional nature of the parser and enhance the node-token alignment, we propose to simplify SBN by assuming the implicit box membership. We also experiment with different node-token aligners and compositional parsers. This project endeavors to offer fresh insights into scope parsing and the compositionality of meaning.

## 16 Can GPT-4 Outperform Fine-tuned BERT Models in Relation Extraction Tasks within the Clinical Domain?

**Author:** Goutham Karunakaran

**Presented by:** Goutham Karunakaran

### **Abstract**

Clinical statements, an essential component of healthcare systems, encapsulate crucial information about patients' health statuses, diagnostic tests, and medical procedures. Extracting structured information from these records, specifically linking laboratory test mentions to their corresponding results (a task termed 'relation extraction'), has profound implications for healthcare delivery, enabling more effective information retrieval, disease surveillance, and clinical decision-making.

State-of-the-art transformer models like BERT, Multilingual BERT, XLM-RoBERTa, and BioBERT have been successfully fine-tuned for this purpose. However, the advent of GPT-4, a large language model trained by OpenAI, brings new possibilities to the table.

This study investigates whether GPT-4, even without any fine-tuning, can outperform these aforementioned fine-tuned models in relation extraction tasks. Leveraging a dataset of annotated clinical statements from the E3C Corpus, this research compares the performance of GPT-4 to that of other models, striving to ascertain the feasibility and effectiveness of using GPT-4 for such intricate tasks. The results have the potential to influence the application of NLP in healthcare, opening doors to efficient and robust information extraction without the need for labour-intensive fine-tuning processes.

## 17 Role of intonation in intercomprehension scenario

**Author:** Sofia Shkapa

**Presented by:** Sofia Shkapa

### **Abstract**

Prosody, and intonation in particular, plays a key role in language intercomprehension, which could be both facilitating and confounding. This study focuses on exploring intonational features in intercomprehension, with a specific emphasis on four Slavic languages: Bulgarian, Czech, Polish, and Russian. The investigation centers on the intonation patterns of polar questions and statements. Slavic languages have been shown to exhibit notable similarities at the prosodic level. Additionally, the results of our computational analysis suggest that intonation contours specifically for statement-type sentences of target languages are more resembling than those for question-type sentences. The findings provide a basis for further exploring the impact of intonation on language intercomprehension. Specifically, we aim to extend the research to study the native perception of intonation on question-statements stimuli. To mitigate the influence of lexical barriers, we propose using cognates, thereby isolating the impact of intonation. We will employ sentence type identification (statement, question) as the primary task, which would allow us to elucidate the role of intonation in facilitating or hindering intercomprehension.

## 18 Tackling misinformation in Serbo-Croatian: corpora and experiments

**Author:** Antonije Petrović

**Supervisor:** Rodrigo Agerri, IXA group, University of the Basque Country Jiri Hana, Institute of Formal and Applied Linguistics, Charles University

**Presented by:** Antonije Petrović

### Abstract

Societies around the world are faced with the issue of handling misinformation, including fake news, pseudoscience, spin, propaganda, etc. In an attempt to tackle this issue, we build the first dataset of fake news and misinformation in Serbo-Croatian, a low-resourced South Slavic language. We show that well-structured fact-checking websites help to build such datasets. We describe the process of gathering the data and organizing it in a suitable way so that experiments on different tasks such as news article classification, claim sentence detection, and claim span detection, are possible. Furthermore, we show some preliminary results on these tasks.

## 19 Analyzing Modular Cross-Lingual Transfer Learning

**Author:** Jacobo Del Valle

**Presented by:** Jacobo Del Valle

### **Abstract**

Cross-lingual abilities have been evident in large multilingual language models over the past few years. However, understanding why and under what circumstances they work is not entirely clear. In this work, we work towards a better understanding of these aspects in a specific subset of multilingual models, namely modular ones. The core idea is to evaluate how parameters change and how they remain compatible with data from other languages, i.e. how and under what circumstances shared representations arise. Furthermore, we carry out an exploration of robustness of current approaches in different adaptation scenarios.

## 20 Discovering Emerging Entities Descriptions via Search Engine

**Author:** Averie So

**Presented by:** Averie So

### Abstract

Entity Linking (EL) is the task of connecting a 'mention' of a named entity in natural language to an entity in a knowledge base (KB), such as a Wikipedia page. This task is popular due to its potential downstream application to other NLP tasks such as question-answering and information retrieval. As such, recent state-of-the-art approaches have achieved impressive performance in conventional benchmarks in terms of accuracy, speed and zero-shot capabilities. However, in real-world applications, there are always entities missing from existing Knowledge Bases, especially in constantly updating domains such as news and web. This project aims to utilise a search engine (Bing API) to discover these Emerging Entities from online web pages. The main challenges will involve 1. generating the right query to the search engine to avoid ambiguous entities (eg. Kevin Lynch (Apple software engineer) vs. Kevin Lynch (Irish republican)), and 2. selecting credible and appropriate descriptions of these entities from online web pages.

## **21 Evaluating fairness in job recommendation system using fairness indicator parity**

**Author:** Sahil Chopra

**Presented by:** Sahil Chopra

### **Abstract**

Fairness is a fundamental principle of AI ethics that seeks to ensure that AI systems do not discriminate against certain demographic groups based on their protected attributes such as gender, ethnicity, or religion. Recommendation System (RS) is said to be unfair when it does not perform equally well for different groups of users. The AI fairness literature proposes a wide variety of fairness notions and debiasing strategies that compare the model's performance across groups based on the outcome. While these methods can reduce unfairness, it may lead to significant drop in model performance. We evaluate a variety of RS algorithms in job recommendation in terms of accuracy and beyond-accuracy metrics to explore the fairness in the RS results toward a specific gender group.



## **22 Learning sentence similarity with pre-trained LM**

**Author:** NATALIA FLECHAS MANRIQUE

**Presented by:** NATALIA FLECHAS MANRIQUE

### **Abstract**

This work investigates how three types of large language models pre-trained on different types of corpora (monolingual, multilingual and multimodal), perform on a Semantic Similarity Task after being fine-tuned on Spanish or English dataset. It finds: 1) A comparable performance between monolingual and multilingual models; 2) A sharp advantage of monolingual models over multimodal models and 3) A similar performance to the state-of-the-art presented in the SemEval-2017 STS shared-task (Cer et al., 2017).

## 23 Utilization of Adapters in Pre-trained Speech Models for Parameter Efficient Automatic Speech Recognition

**Author:** Ahnaf Mozib Samin

**Presented by:** Ahnaf Mozib Samin

### **Abstract**

Although pre-training speech models and then fine-tuning has now become the state-of-the-art approach, the models contain millions (or billions) of parameters, making them difficult to deploy in offline mobile devices. Moreover, these models suffer from the catastrophic forgetting issue in multi-task learning scenarios. Thus, we propose utilizing parameter-efficient adapters for automatic speech recognition. In this study, we evaluate the data requirement for the adapter-tuning approach to perform well. Moreover, we provide solutions to improve the performance of adapter-tuning in scenarios where we have a large amount of data, however, the parameter-efficient adapters cannot produce satisfactory results. Our experiments are conducted in three diverse languages including English, Maltese, and Bengali.

## 24 Quantifying and Mitigating Gender Bias in Chinese Word Embedding

**Author:** Chen Jingyan

**Presented by:** Chen Jingyan

### **Abstract**

Gender bias is the preference or prejudice toward one gender over another. Men, for example, are frequently over-represented than women due to asymmetrical valuations. Gender bias is explicitly expressed in language and thus has been analyzed both by linguistics and Natural Language Processing (NLP) communities. Word embeddings are an important model in NLP. The study consists of three stages: collecting data from Chinese news sources to create a corpus, assessing bias through word analogy tests, and addressing bias using two approaches. The first approach involves neutralizing the gender projection of words using predefined gender directions, while the second approach ensures equal proximity of neutral words to gendered word pairs. The objective is to evaluate and mitigate gender bias in Chinese word embeddings, making valuable contributions to the NLP field.

## 25 OPEN INFORMATION EXTRACTION USING LARGE LANGUAGE MODELS

**Author:** Amrita H. Nair

**Presented by:** Amrita H. Nair

### **Abstract**

Open Information Extraction(OIE) is a task in NLP which involves generating a structured n-ary representation of text to facilitate various functions like knowledge graph construction. OIE facilitates the retrieval of relevant entities without the need for pre-defined relations. Initial research in this area involved a rule-based system and also included assessing the linguistic composition of text to extract relevant data. Different datasets and evaluation benchmarks have emerged over the years. In the last few years, LLMs (Large Language Models) have revolutionised the field of NLP and proven to be capable and as good as, if not better than state-of-the-art models for tasks like sentiment analysis, name entity recognition, machine translation and so on. Recent studies have proven that LLM's store relational as well as linguistic knowledge, proving that OIE for LLM's is a significant area of research. Through leveraging the attention mechanism, relevant relations between two noun-phrase entities can be extracted which results in an F1 score that is competitive to the state-of-the-art on well-known OIE datasets. This poster focuses on the broader idea and research done on OIE for LLM's and attempts to provide a literature survey for the same.

## 26 GPL at SemEval-2023 Task 1: WordNet and CLIP to Disambiguate Images

Shibingfeng Zhang<sup>1,2</sup> Shantanu Nath<sup>1</sup> Davide Mazzaccara<sup>1</sup>

<sup>1</sup>CIMeC, University of Trento <sup>2</sup>Saarland University

{shibingfeng.zhang,shantanu.nath,davide.mazzaccara}@unitn.it

**Presented by:** Shantanu Nath

**Abstract** Given a word in context, the task of Visual Word Sense Disambiguation consists of selecting the correct image among a set of candidates. To select the correct image, we propose a solution blending text augmentation and multimodal models. Text augmentation leverages the fine-grained semantic annotation from WordNet to get a better representation of the textual component. We then compare this sense-augmented text to the set of image using pre-trained multimodal models CLIP and ViLT. Our system has been ranked 16<sup>th</sup> for the English language, achieving 68.5 points for hit rate and 79.2 for mean reciprocal rank. The code to this project is available on Github.