

Seminar Natural Language Processing (NLP) — Part 2

Overview of Seminar Topics

Henning Wachsmuth

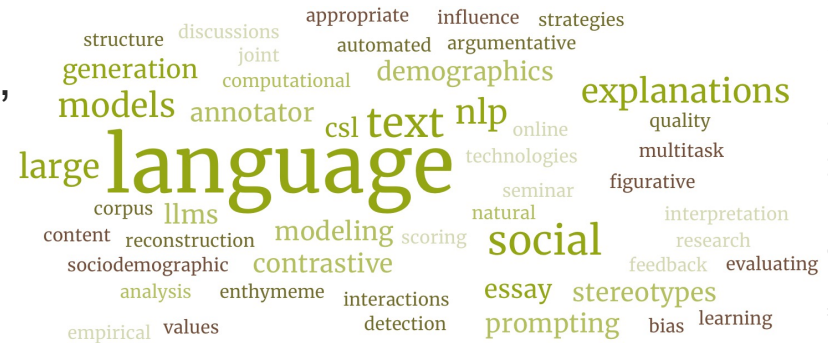
<https://ai.uni-hannover.de>



Assignment of seminar topics

■ This talk

- Overview of 15 possible seminar topics, grouped by the respective advisors
- For each topic, one paper is given that provides the basis of the topic



■ Concept behind

- Each seminar participant will be assigned one topic
- The paper should be the focus of the talk and of the article
- Further literature on basics and related work should be added where useful
- You have to choose topic preferences, we then assign topics

■ Your task

- Inform yourself about the topics and papers in this presentation
- Choose 3 topics with preferences
- Until Monday, October 28, 23:59 UTC+2. Send e-mail with preferences

Both direct e-mail and Stud.IP message are fine.



Choosing preferences: eMail and subsequent process

■ Your e-mail

- **Recipient.** h.wachsmuth@ai.uni-hannover.de
- **Subject.** "[nlp] Topic preferences"
- **Content.** Your name, matriculation number, and 3 topic preferences
Notice: Less than 3 does not increase chances to get one of them
- **Example.** On the right, you see how the content of your e-mail could look like

Name:

Gabriella Skitalinska

Matriculation number:

12345678

Topic preferences:

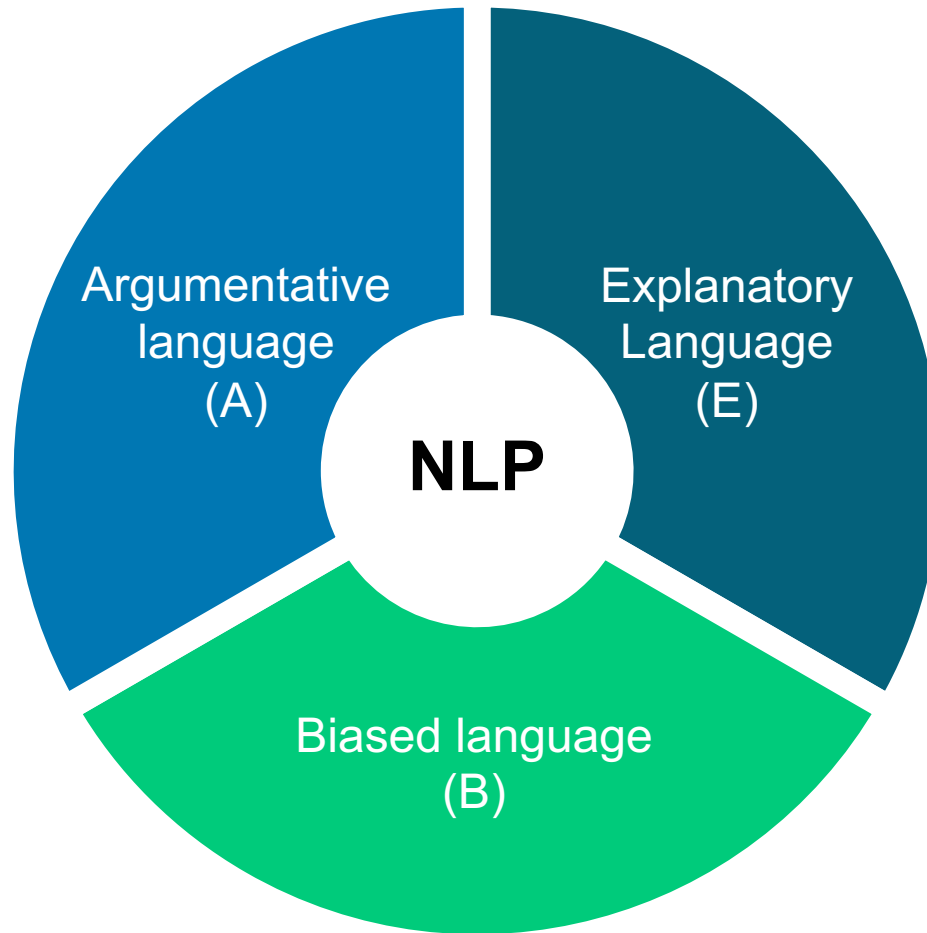
- 1) B4. Evaluating Stereotypes in LLMs
- 2) A1. Modeling Appropriate Language
- 3) B5. Values and opinions in LLMs

■ Subsequent process

- We will assign topics based on preferences, special reasons, and randomly
- **If you don't send your e-mail in time, you will *not* be assigned any topic**
- The final schedule will be decided based on the topic assignment
You can get a rough idea of the schedule from the ordering on the next slides.
- Topic assignment and schedule will be announced by end of next week

Overview of topics

Seminar topics organized into three areas



Topics for the seminar talks and articles (1 out of 2)

▪ **Argumentative language**

- A1. Modeling Appropriate Language
- A2. LLMs as Content Moderators
- A3. Interactions of Argument Structure and Quality
- A4. Joint Essay Scoring and Feedback Generation
- A5. Enthymeme Detection and Reconstruction

Timon
Carlotta
Maja
Maja
Maja



▪ **Biased language**

- B1. Influence of Annotator Demographics
- B2. Sociodemographic Prompting
- B3. Addressing Stereotypes in Online Discussions
- B4. Evaluating Stereotypes in LLMs
- B5. Values and Opinions in LLMs
- B6. Identifying Social Bias in Texts

Leandra
Leandra
Max
Max
Yamen
Max



Explanatory language on next slide...

Topics for the seminar talks and articles (2 out of 2)

- **Explanatory language**

E1. Multimodal Figurative Language

E2. Metaphor Interpretation

E3. Explainability in LLMs

E4. Contrastive Explanations

Meghdut

Meghdut

Meghdut

Leandra



Notice: E4 is going to be schedule between A5 and B1

Topics and literature

Topics supervised by Timon, Carlotta, and Yamen

A1. Modeling Appropriate Language

- Timon Ziegenbein, Shahbaz Syed, Felix Lange, Martin Potthast, and Henning Wachsmuth. 2023. Modeling Appropriate Language in Argumentation.
<https://aclanthology.org/2023.acl-long.238/>



A2. LLMs as Content Moderators

- Mahi Kolla, Siddharth Salunkhe, Eshwar Chandrasekharan, and Koustuv Saha. 2024. LLM-Mod: Can Large Language Models Assist Content Moderation?
<https://doi.org/10.1145/3613905.3650828>



B5. Values and Opinions in LLMs

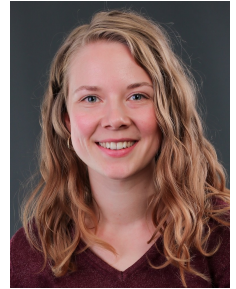
- Paul Röttger, Valentin Hofmann, Valentina Pyatkin, Musashi Hinck, Hannah Rose Kirk, Hinrich Schütze, Dirk Hovy. 2024. Political Compass or Spinning Arrow? Towards More Meaningful Evaluations for Values and Opinions in Large Language Models.
<https://aclanthology.org/2024.acl-long.816>



Topics supervised by Maja

A3. Interactions of Argument Structure and Quality

- Maja Stahl, Nadine Michel, Sebastian Kilsbach, Julian Schmidtke, Sara Rezat, Henning Wachsmuth. 2024. A School Student Essay Corpus for Analyzing Interactions of Argumentative Structure and Quality.
<https://aclanthology.org/2024.naacl-long.145>



A4. Joint Essay Scoring and Feedback Generation

- Maja Stahl, Leon Biermann, Andreas Nehring, Henning Wachsmuth. 2024. Exploring LLM Prompting Strategies for Joint Essay Scoring and Feedback Generation.
<https://aclanthology.org/2024.bea-1.23>

A5. Enthymeme Detection and Reconstruction

- Maja Stahl, Nick Düsterhus, Mei-Hua Chen, Henning Wachsmuth. 2023. Mind the Gap: Automated Corpus Creation for Enthymeme Detection and Reconstruction in Learner Arguments.
<https://aclanthology.org/2023.findings-emnlp.312>

Topics supervised by Leandra

E4. Contrastive Explanations

- Oliver Eberle, Ilias Chalkidis, Laura Cabello, Stephanie Brandl. 2023. Rather a Nurse than a Physician - Contrastive Explanations under Investigation.
<https://aclanthology.org/2023.emnlp-main.427>



B1. Influence of Annotator Demographics

- Jiaxin Pei, David Jurgens. 2023. When Do Annotator Demographics Matter? Measuring the Influence of Annotator Demographics with the POPQUORN Dataset.
<https://aclanthology.org/2023.law-1.25>

B2. Sociodemographic Prompting

- Tilman Beck, Hendrik Schuff, Anne Lauscher, Iryna Gurevych. 2024. Sensitivity, Performance, Robustness: Deconstructing the Effect of Sociodemographic Prompting.
<https://aclanthology.org/2024.eacl-long.159>

Topics supervised by Max

B3. Addressing Stereotypes in Online Discussions

- Kathleen Fraser, Svetlana Kiritchenko, Isar Nejadgholi, and Anna Kerkhof. 2023. What Makes a Good Counter-Stereotype? Evaluating Strategies for Automated Responses to Stereotypical Text. <https://doi.org/10.18653/v1/2023.sicon-1.4>



B4. Evaluating Stereotypes in LLMs

- Flor Plaza-del-Arco, Amanda Curry, Alba Cercas Curry, Gavin Abercrombie, and Dirk Hovy. 2024. Angry Men, Sad Women: Large Language Models Reflect Gendered Stereotypes in Emotion Attribution. <https://aclanthology.org/2024.acl-long.415/>

B6. Identifying Social Bias in Texts

- Maximilian Spliethöver, Sai Nikhil Menon, and Henning Wachsmuth. 2024. Disentangling Dialect from Social Bias via Multitask Learning to Improve Fairness. <https://aclanthology.org/2024.findings-acl.553>

Topics supervised by Meghdut

E1. Multimodal Figurative Language

- Arkadiy Saakyan, Shreyas Kulkarni, Tuhin Chakrabarty, and Smaranda Muresan. 2024. Understanding Figurative Meaning through Explainable Visual Entailment.
<https://arxiv.org/abs/2405.01474>



E2. Metaphor Interpretation

- Meghdut Sengupta, Milad Alshomary, Ingrid Scharlau, and Henning Wachsmuth. 2023. Modeling Highlighting of Metaphors in Multitask Contrastive Learning Paradigms.
<https://aclanthology.org/2023.findings-emnlp.308>

E3. Explainability in LLMs

- James Enouen, Hootan Nakhost, Sayna Ebrahimi, Sercan Arik, Yan Liu, and Tomas Pfister. 2024. TextGenSHAP: Scalable Post-hoc Explanations in Text Generation with Long Documents.
<https://aclanthology.org/2024.findings-acl.832>

Sum-up

Conclusion

▪ Seminar topics

- 15 candidate topics related to three research areas
- Each of you will be assigned one of these topics
- Given (+ further) literature forms the basis of talk and article



▪ Topic assignment

- You choose topic preferences, we assign topics
- Inform yourself about the topics of the given literature
- [Send me your topic preferences by Monday next week!](#)



▪ Next up

- Topic assignment will be announced by end of next week
- Talk preparation starts then
- Basics of scientific presentation on November 7

