Conference Report: The 17th Conference of the European Chapter of the Association for Computational Linguistics (EACL 2023)*

Roman Klinger

2023-05-12

Last week, May 2-6, 2023, the 17th Conference of the European Chapter of the Association for Computational Linguistics (EACL) took place in Dubrovnik. I think this has been my first EACL conference, and the program was very interesting, with a good mixture of recent state-of-the-art research on fashionable topics while maintaining a good diversity across various research fields. It's been one of my favorite conferences that I attended so far.

Figure 1: Plenary Session

The conference hat 281 papers accepted, out of which there were 229 long and 41 short papers. The acceptance rate was 24,1%. Such a comparably low rate leads to many good papers not finding a space in the conference. EMNLP invented the model of the Findings of ACL/EMNLP/EACL/NAACL, in which papers are included that might not fit in the main conference but are still worth publishing. In some conferences these papers are presented as posters, sometimes invited to be presented in workshops, and sometimes presented only as videos online. For EACL, all 201 Findings papers (149 long) were presented in a video on the Underline platform, and some papers were additionally presented in workshops.

---

*Published as blog post at https://www.romanklinger.de/blog/2023-05-12-eacl/

I am wondering if the invitations to present in workshops shaped the perception that Findings papers are somehow between the main conference and the workshops - I heard this opinion multiple times at the conference. Personally, my perception is more that Findings papers are in the same category as main conference papers and not more similar to workshop papers. Workshop papers are not worse – they are more focused on specific topics. Findings papers often lack this special focus and therefore they are not submitted to a dedicated workshop but to the main conference.
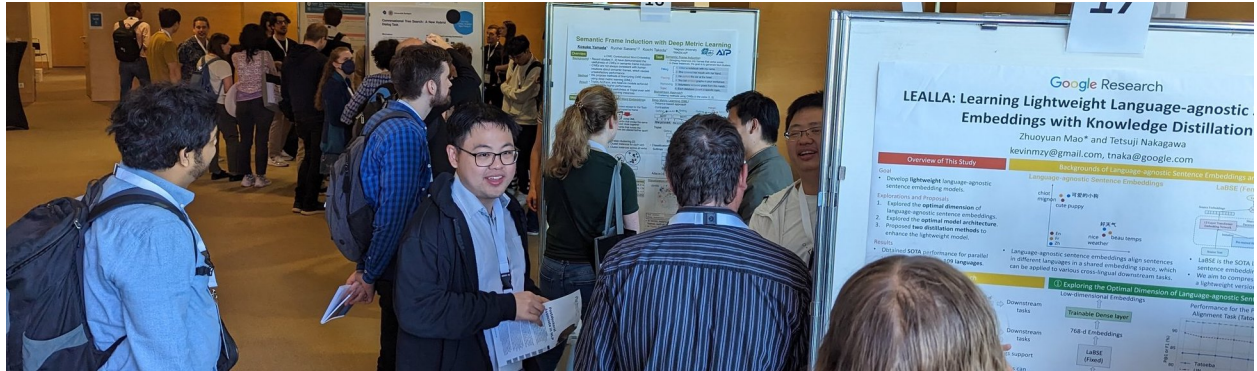

Figure 2: Poster Session

Next to many oral presentations, there were very nice poster sessions. At *ACL/EMNLP conferences, poster papers are considered to be of the same quality as those presented as talks: there is no difference in the proceedings.

The program of the conference, with talks and posters, has been complemented by tutorials and so-called Birds-of-a-feather (BoF) sessions, in which people interested in a specific topic introduced each other, as a networking event. This format has been introduced during the COVID-online conference times (I think) and still exists. It's actually quite nice to get in touch with a subcommunity that one did not know yet. I did participate in such sessions for the first time and can only recommend it.

Contributions from University of Stuttgart

Tutorials

The IMS at the University of Stuttgart had many contributions, and it felt very nice to be at the conference with so many nice colleagues. I did not have this experience as a PhD student (where I was essentially the only person targeting ACL conferences for publications in the group), and I really appreciated it. It is so much easier to meet new people if you already know many.

The IMS contributed two tutorials. I was part of the Emotion Analysis Tutorial (Štajner and Klinger 2023), offered together with Sanja Štajner. It's been my first tutorial, as usual, I did not plan enough time for the material that I wanted to cover. Thanks to Sanja, who was flexible enough with her timing, we did not overrun too much.

Another tutorial with substantial involvement by IMS people was given by Gabriella Lapesa, Eva Maria Vecchi, Serena Villata, and Henning Wachsmuth (Lapesa et al. 2023). The topic was argument mining, and it unfortunately took place in parallel,
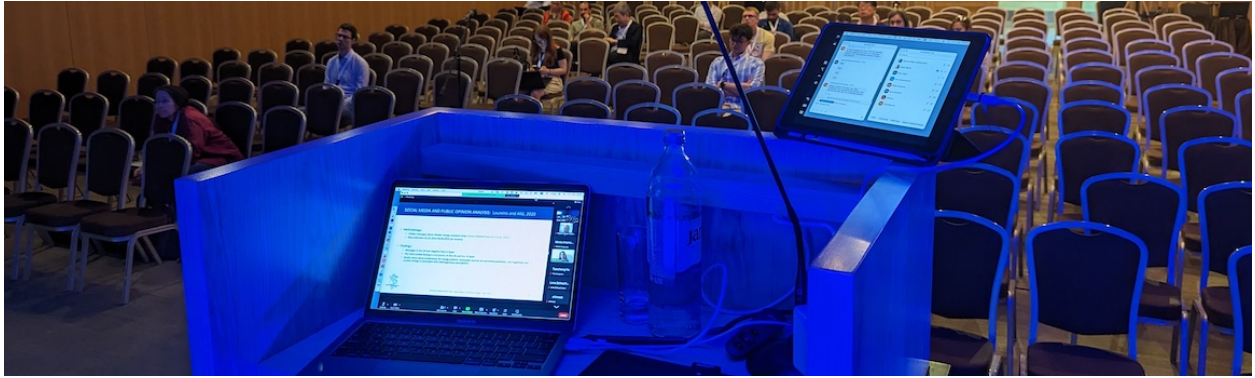
Figure 3: EA Tutorial

which was a pity because the same people might have been interested in both. Luckily the tutorials were recorded and will be online on the underline platform.

Papers

We further had a set of papers in the main conference, the workshops, and Findings.

Wuehrl, Grimminger, and Klinger (2023) proposed a real-world pipeline for biomedical fact-checking, based on the idea that reformulated claims can be better checked against scientific text than the original formulation of a claim as it occurs in social media. Miletic and Schulte im Walde (2023) showed how compositionality information can be extracted from BERT. Eichel, Schlipf, and Schulte im Walde (2023) investigate how LLM can be prompted for plausability with applications in the material sciences. Falk and Lapesa (2023) show how adapters can be used to efficiently predict argument quality, based on a large set of datasets and quality dimensions. Nikolaev and Padó (2023) study representation biases in sentence transformers, Gaser et al. (2023) explore segmentation approaches for neural machine translation with code-switching, and Väth, Vanderlyn, and Vu (2023) show how dialog systems allow for more complex tree-like conversations with intelligent agents.

My Favorite Contributions

In addition to the IMS contributions to the conference, I found a set of talks and papers very interesting. This only reflects my personal opinion, and that I do not mention a particular paper probably only means that I did not have the time to go to its presentation. There were many interesting papers in the program, I did not go through all of them yet.

Invited Talks

Before I say something about the papers I liked a lot, I would like to point out two of the three invited talks. Joyce Chai talked about embodied AI. As a student of computer science, I often heard the phrase "intelligence needs a body", and I must say, I never

really understood. Now, with this talk, and the nice demonstration videos that Joyce showed, I finally got a grasp of what's behind this phrase. Full understanding of the whole context is only possible in multimodal interactions. That does not mean that every researcher needs to work on multimodal interaction analysis, but there needs to be such integration efforts to not miss important aspects. I found that very intuitive.

Figure 4: Keynote Joyce Chai

The other keynote, given by Edward Grefenstette included discussions on the efficiency of LLM and their future use. He mentioned work by Lyle et al. (2020) who studied why LLM can actually generalize. Apparently, it is crucial to only have one epoch during pre-training. On a more entertaining side, he pointed out that LLM currently mostly fail with pragmatics ("Have you seen my phone" – "Yes, I have seen your phone.")

Figure 5: Keynote Grefenstette

Papers

My Favorite papers:

- Eisenschlos et al. (2023) study how LLM can learn new words in-context at inference time and develop a method to measure such word acquisition (by prompt-based coreference resolution with new words). This paper also won a best paper award.
- Ishibashi et al. (2023) analyze the robustness of prompts by prompt pertubation. One interesting finding is that manual prompts are more robust than automatically learned prompts in few-shot settings. Very interesting study to get a better understanding what "good prompts" are.
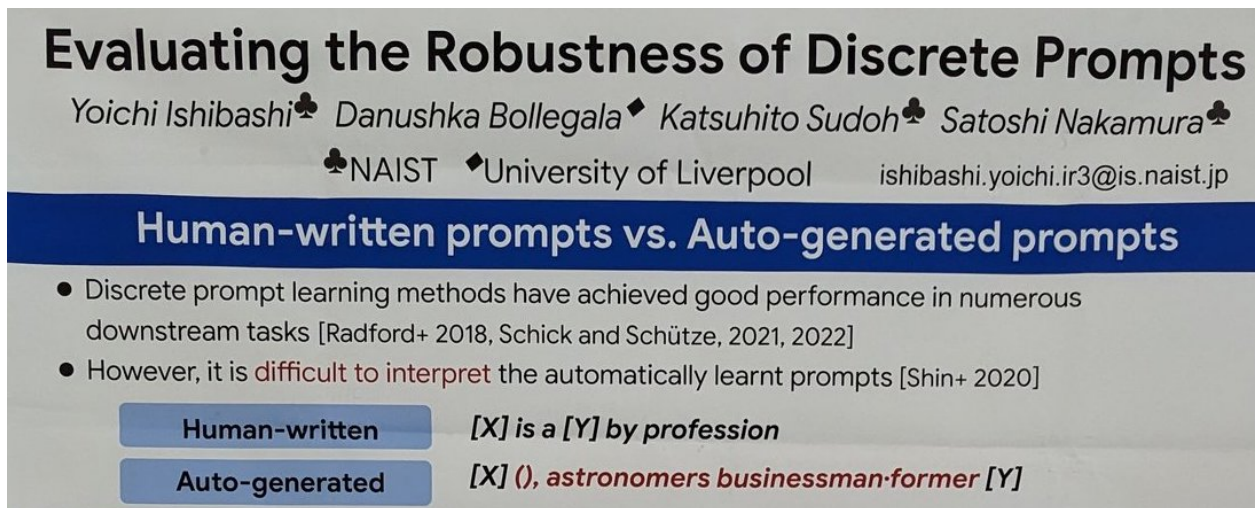
Figure 6: Ishibashi Poster

- We know that LLM tend to hallucinate content. This can also happen during machine translation (I did not know that, and it sounds pretty scary!). Understanding such hallucinations is the topic of the work by Guerreiro, Voita, and Martins (2023). They also propose a method to mitigate the issue by regularization during inference.
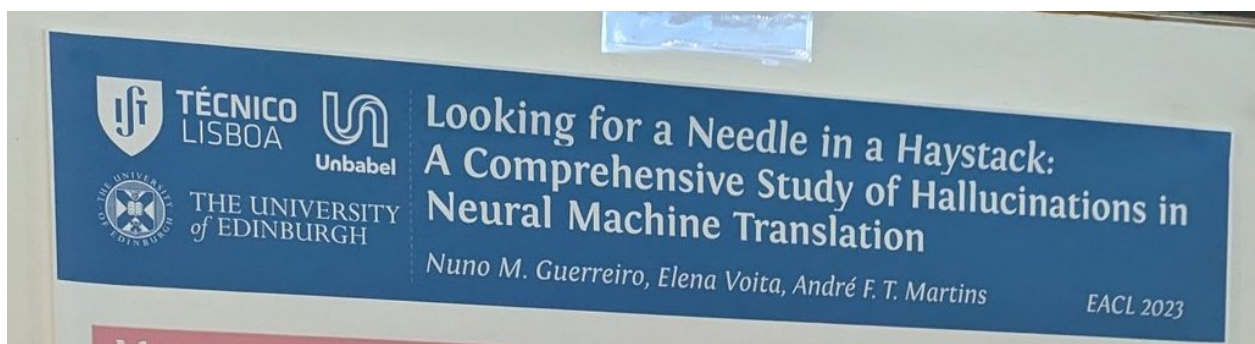

Figure 7: Guerreiro Poster

- Govindarajan et al. (2023) point out that there is no such thing as unbiased language! They look at interpersonal bias and emotion.
- Zhong, Dhuliawala, and Stoehr (2023) study a task that sounds like it should be super-straight-forward to solve: extract mentions of counts from social media (here: victim counts). Apparently, the task is really difficult, because models need to understand enumerations, implicit references, next to actual mentions of numbers. This is a very interesting paper, because it shows another case where general models fail and that specifically developed models for particular tasks are important.
- Mohammad (2023) wrote a paper about how to use emotion lexicons and how to build them. The paper style is worth pointing out: it's written in a question-answer style, and I think that this is very accessible.
- Most of our language models are huge these days, and luckily, there are methods to compress them, such that they can not only be used on GPU clusters. Du et

al. (2023) show that such compression comes with disadvantages: it reinforces biases.

- Recently, several methods have been proposed to automatically find well-performing prompts (for instance Shin et al. (2020), Ding et al. (2022)). In their paper, Prasad et al. (2023) focus on instruction tuning without a need to calculate gradients.
- Narayanan Venkit et al. (2023) look into nationality bias (instead of a bias towards particular languages).
- Parmar et al. (2023) (also outstanding paper) study if instructions in crowd-sourcing tasks create biases (unfortunately, the answer is yes).
- Cortal et al. (2023) build on top of previous work that we published, on the emotion component process model and appraisal theories for emotion analysis, particularly Casel, Heindl, and Klinger (2021). They build on top of another appraisal theory that focuses on the cognitive component of emotions and create a corpus in French, with a focus on emotion regulation. This is the first work that I have seen that puts emotion regulation into focus in NLP work.

Awards

I won't go into detail regarding the paper awards, these papers have already been evaluated by other people to be interesting, I'll just list them here:

Outstanding papers:

- Nguyen et al. (2023)
- Madusanka, Batista-navarro, and Pratt-hartmann (2023)
- Milich and Akbik (2023)
- Chen et al. (2023)
- Lu and Lam (2023)
- Epure and Hennequin (2023)
- Parmar et al. (2023)
- Belouadi and Eger (2023)
- Krishna et al. (2023)
- Winata et al. (2023)

Best papers:

- Eisenschlos et al. (2023)
- Misra, Rayz, and Ettinger (2023)

Venue and Place

At the end of this blog post, I'd like to comment on the location. Personally, I like to categorize conference locations into three types:

1. Nice conference centers in some city with hotels around, without a hotel directly associated to the venue. An example was COLING 2018 in a very nice place downtown Santa Fe. The advantage is: such conference locations are nice! The disadvantage is: participants do not tend to hang out at the venue.

2. Conference hotels downtown of some big city. An example was NAACL-HLT 2019 in Minneapolis or ACL 2014 in Baltimore. If one can effort the conference hotel, that's nice because people hang around also outside of the conference schedule. Unfortunately, these hotels are sometimes prohibitively expensive, and then people are just elsewhere - and in contrast to (1), the conference center is typically not even nice but arbitrary.
3. Conference hotels somewhere where nothing else is that might motivate one to be elsewhere. This has been the setup of RANLP 2009 and RANLP2011 (not sure about later). Of course both places had a lot to offer outside of the conference hotel, but given that these places were pretty empty outside of the respective season, one ran into conference people everywhere.

The conference hotel Valamar Lacroma Dubrovnik Hotel of EACL 2023 was something in-between. It was slightly too expensive such that all particants would decide to stay there, but downtown was sufficiently far away such that people did not go elsewhere during breaks, even if they had the accomodation elsewhere. I must say that I found this to be a very good setup. The city of Dubrovnik was beautiful, and in the evening we ran into hundreds of EACL people. But the conference venue/hotel had enough to offer that one could also stay there and talk.

Figure 8: Conference Dinner

Figure 9: Conference Room

Bibliography

Belouadi, Jonas, and Steffen Eger. 2023. "UScore: An Effective Approach to Fully Unsupervised Evaluation Metrics for Machine Translation." In *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics*, 358–74. Dubrovnik, Croatia: Association for Computational Linguistics. https://aclanthology.org/2023.eacl-main.27.

Casel, Felix, Amelie Heindl, and Roman Klinger. 2021. "Emotion Recognition Under Consideration of the Emotion Component Process Model." In *Proceedings of the 17th Conference on Natural Language Processing (KONVENS 2021)*, 49–61. Düsseldorf, Germany: KONVENS 2021 Organizers. https://aclanthology.org/2021.konvens-1.5.

Chen, Yunmo, William Gantt, Weiwei Gu, Tongfei Chen, Aaron White, and Benjamin Van Durme. 2023. "Iterative Document-Level Information Extraction via Imitation

Learning." In *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics*, 1858–74. Dubrovnik, Croatia: Association for Computational Linguistics. https://aclanthology.org/2023.eacl-main.136.

Cortal, Gustave, Alain Finkel, Patrick Paroubek, and Lina Ye. 2023. "Emotion Recognition Based on Psychological Components in Guided Narratives for Emotion Regulation." In *Proceedings of the 7th Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature*, 72–81. Dubrovnik, Croatia: Association for Computational Linguistics. https://aclanthology.org/2023.latechclfl-1.8.

Ding, Ning, Shengding Hu, Weilin Zhao, Yulin Chen, Zhiyuan Liu, Haitao Zheng, and Maosong Sun. 2022. "OpenPrompt: An Open-Source Framework for Prompt-Learning." In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, 105–13. Dublin, Ireland: Association for Computational Linguistics. https://doi.org/10.18653/v1/2022.acl-demo.10.

Du, Mengnan, Subhabrata Mukherjee, Yu Cheng, Milad Shokouhi, Xia Hu, and Ahmed Hassan Awadallah. 2023. "Robustness Challenges in Model Distillation and Pruning for Natural Language Understanding." In *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics*, 1766–78. Dubrovnik, Croatia: Association for Computational Linguistics. https://aclanthology.org/2023.eacl-main.129.

Eichel, Annerose, Helena Schlipf, and Sabine Schulte im Walde. 2023. "Made of Steel? Learning Plausible Materials for Components in the Vehicle Repair Domain." In *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics*, 1420–35. Dubrovnik, Croatia: Association for Computational Linguistics. https://aclanthology.org/2023.eacl-main.104.

Eisenschlos, Julian Martin, Jeremy R. Cole, Fangyu Liu, and William W. Cohen. 2023. "WinoDict: Probing Language Models for in-Context Word Acquisition." In *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics*, 94–102. Dubrovnik, Croatia: Association for Computational Linguistics. https://aclanthology.org/2023.eacl-main.7.

Epure, Elena, and Romain Hennequin. 2023. "A Human Subject Study of Named Entity Recognition in Conversational Music Recommendation Queries." In *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics*, 1281–96. Dubrovnik, Croatia: Association for Computational Linguistics. https://aclanthology.org/2023.eacl-main.92.

Falk, Neele, and Gabriella Lapesa. 2023. "Bridging Argument Quality and Deliberative Quality Annotations with Adapters." In *Findings of the Association for Computational Linguistics: EACL 2023*, 2469–88. Dubrovnik, Croatia: Association for Computational Linguistics. https://aclanthology.org/2023.findings-eacl.187.

Gaser, Marwa, Manuel Mager, Injy Hamed, Nizar Habash, Slim Abdennadher, and Ngoc Thang Vu. 2023. "Exploring Segmentation Approaches for Neural Machine Translation of Code-Switched Egyptian Arabic-English Text." In *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics*, 3523–38. Dubrovnik, Croatia: Association for Computational Linguistics. https://aclanthology.org/2023.eacl-main.256.

Govindarajan, Venkata Subrahmanyan, Katherine Atwell, Barea Sinno, Malihe Alikhani, David I. Beaver, and Junyi Jessy Li. 2023. "How People Talk about Each Other: Modeling Generalized Intergroup Bias and Emotion." In *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics*, 2496–506. Dubrovnik, Croatia: Association for Computational Linguistics. https:

//aclanthology.org/2023.eacl-main.183.

Guerreiro, Nuno M., Elena Voita, and André Martins. 2023. "Looking for a Needle in a Haystack: A Comprehensive Study of Hallucinations in Neural Machine Translation." In *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics*, 1059–75. Dubrovnik, Croatia: Association for Computational Linguistics. https://aclanthology.org/2023.eacl-main.75.

Ishibashi, Yoichi, Danushka Bollegala, Katsuhito Sudoh, and Satoshi Nakamura. 2023. "Evaluating the Robustness of Discrete Prompts." In *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics*, 2373–84. Dubrovnik, Croatia: Association for Computational Linguistics. https://aclanthology.org/2023.eacl-main.174.

Krishna, Kalpesh, Erin Bransom, Bailey Kuehl, Mohit Iyyer, Pradeep Dasigi, Arman Cohan, and Kyle Lo. 2023. "LongEval: Guidelines for Human Evaluation of Faithfulness in Long-Form Summarization." In *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics*, 1650–69. Dubrovnik, Croatia: Association for Computational Linguistics. https://aclanthology.org/2023.eacl-main.121.

Lapesa, Gabriella, Eva Maria Vecchi, Serena Villata, and Henning Wachsmuth. 2023. "Mining, Assessing, and Improving Arguments in NLP and the Social Sciences." In *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics: Tutorial Abstracts*, 1–6. Dubrovnik, Croatia: Association for Computational Linguistics. https://aclanthology.org/2023.eacl-tutorials.1.

Lu, Hongyuan, and Wai Lam. 2023. "PCC: Paraphrasing with Bottom-k Sampling and Cyclic Learning for Curriculum Data Augmentation." In *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics*, 68–82. Dubrovnik, Croatia: Association for Computational Linguistics. https://aclanthology.org/2023.eacl-main.5.

Lyle, Clare, Lisa Schut, Binxin Ru, Yarin Gal, and Mark van der Wilk. 2020. "A Bayesian Perspective on Training Speed and Model Selection." In *34th Conference on Neural Information Processing Systems (NeurIPS 2020)*. Vancouver, Canada. https://proceedings.neurips.cc/paper/2020/file/75a7c30fc0063c4952d7eb044a3c0897-Paper.pdf.

Madusanka, Tharindu, Riza Batista-navarro, and Ian Pratt-hartmann. 2023. "Identifying the Limits of Transformers When Performing Model-Checking with Natural Language." In *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics*, 3539–50. Dubrovnik, Croatia: Association for Computational Linguistics. https://aclanthology.org/2023.eacl-main.257.

Miletic, Filip, and Sabine Schulte im Walde. 2023. "A Systematic Search for Compound Semantics in Pretrained BERT Architectures." In *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics*, 1499–1512. Dubrovnik, Croatia: Association for Computational Linguistics. https://aclanthology.org/2023.eacl-main.110.

Milich, Marcel, and Alan Akbik. 2023. "ZELDA: A Comprehensive Benchmark for Supervised Entity Disambiguation." In *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics*, 2061–72. Dubrovnik, Croatia: Association for Computational Linguistics. https://aclanthology.org/2023.eacl-main.151.

Misra, Kanishka, Julia Rayz, and Allyson Ettinger. 2023. "COMPS: Conceptual Minimal Pair Sentences for Testing Robust Property Knowledge and Its Inheritance in Pre-Trained Language Models." In *Proceedings of the 17th Conference of the European*

Chapter of the Association for Computational Linguistics, 2928–49. Dubrovnik, Croatia: Association for Computational Linguistics. https://aclanthology.org/2023.eacl-main.213.

Mohammad, Saif. 2023. "Best Practices in the Creation and Use of Emotion Lexicons." In *Findings of the Association for Computational Linguistics: EACL 2023*, 1825–36. Dubrovnik, Croatia: Association for Computational Linguistics. https://aclanthology.org/2023.findings-eacl.136.

Narayanan Venkit, Pranav, Sanjana Gautam, Ruchi Panchanadikar, Ting-Hao Huang, and Shomir Wilson. 2023. "Nationality Bias in Text Generation." In *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics*, 116–22. Dubrovnik, Croatia: Association for Computational Linguistics. https://aclanthology.org/2023.eacl-main.9.

Nguyen, Laura, Thomas Scialom, Benjamin Piwowarski, and Jacopo Staiano. 2023. "LoRaLay: A Multilingual and Multimodal Dataset for Long Range and Layout-Aware Summarization." In *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics*, 636–51. Dubrovnik, Croatia: Association for Computational Linguistics. https://aclanthology.org/2023.eacl-main.46.

Nikolaev, Dmitry, and Sebastian Padó. 2023. "Representation Biases in Sentence Transformers." In *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics*, 3701–16. Dubrovnik, Croatia: Association for Computational Linguistics. https://aclanthology.org/2023.eacl-main.268.

Parmar, Mihir, Swaroop Mishra, Mor Geva, and Chitta Baral. 2023. "Don't Blame the Annotator: Bias Already Starts in the Annotation Instructions." In *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics*, 1779–89. Dubrovnik, Croatia: Association for Computational Linguistics. https://aclanthology.org/2023.eacl-main.130.

Prasad, Archiki, Peter Hase, Xiang Zhou, and Mohit Bansal. 2023. "GrIPS: Gradient-Free, Edit-Based Instruction Search for Prompting Large Language Models." In *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics*, 3845–64. Dubrovnik, Croatia: Association for Computational Linguistics. https://aclanthology.org/2023.eacl-main.277.

Shin, Taylor, Yasaman Razeghi, Robert L. Logan IV, Eric Wallace, and Sameer Singh. 2020. "AutoPrompt: Eliciting Knowledge from Language Models with Automatically Generated Prompts." In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 4222–35. Online: Association for Computational Linguistics. https://doi.org/10.18653/v1/2020.emnlp-main.346.

Štajner, Sanja, and Roman Klinger. 2023. "Emotion Analysis from Texts." In *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics: Tutorial Abstracts*, 7–12. Dubrovnik, Croatia: Association for Computational Linguistics. https://aclanthology.org/2023.eacl-tutorials.2.

Väth, Dirk, Lindsey Vanderlyn, and Ngoc Thang Vu. 2023. "Conversational Tree Search: A New Hybrid Dialog Task." In *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics*, 1264–80. Dubrovnik, Croatia: Association for Computational Linguistics. https://aclanthology.org/2023.eacl-main.91.

Winata, Genta Indra, Alham Fikri Aji, Samuel Cahyawijaya, Rahmad Mahendra, Fajri Koto, Ade Romadhony, Kemal Kurniawan, et al. 2023. "NusaX: Multilingual Parallel Sentiment Dataset for 10 Indonesian Local Languages." In *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Lin-*

guistics, 815–34. Dubrovnik, Croatia: Association for Computational Linguistics. https://aclanthology.org/2023.eacl-main.57.

Wuehrl, Amelie, Lara Grimminger, and Roman Klinger. 2023. "An Entity-Based Claim Extraction Pipeline for Real-World Biomedical Fact-Checking." In *Proceedings of the Sixth Fact Extraction and VERification Workshop (FEVER)*, 29–37. Dubrovnik, Croatia: Association for Computational Linguistics. https://aclanthology.org/2023.fever-1.3.

Zhong, Mian, Shehzaad Dhuliawala, and Niklas Stoehr. 2023. "Extracting Victim Counts from Text." In *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics*, 1925–40. Dubrovnik, Croatia: Association for Computational Linguistics. https://aclanthology.org/2023.eacl-main.141.

[Download this post as PDF]