Opinion diversity through hybrid intelligence

Meer, M.T. van der

Bibliography

References

[1] Gavin Abercrombie, Amanda Cercas Curry, Tanvi Dinkar, Verena Rieser, and Zeerak Talat. Mirages. on anthropomorphism in dialogue systems. In Houda Bouamor, Juan Pino, and Kalika Bali, editors, *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 4776–4790, Singapore, December 2023. Association for Computational Linguistics. doi: 10.18653/v1/2023.emnlp-main.290.

[2] Francisca Adoma Acheampong, Chen Wenyu, and Henry Nunoo-Mensah. Text-based emotion detection: Advances, challenges, and opportunities. *Engineering Reports*, 2(7):e12189, 2020.

[3] G. Adomavicius and A. Tuzhilin. Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering*, 17(6):734–749, 2005. doi: 10.1109/TKDE.2005.99.

[4] Mikel Aickin and Helen Gensler. Adjusting for multiple testing when reporting research results: the bonferroni vs holm methods. *American journal of public health*, 86 (5):726–728, 1996.

[5] Zeynep Akata, Dan Balliet, Maarten De Rijke, Frank Dignum, Virginia Dignum, Guszti Eiben, Antske Fokkens, Davide Grossi, Koen Hindriks, Holger Hoos, et al. A research agenda for hybrid intelligence: augmenting human intellect with collaborative, adaptive, responsible, and explainable artificial intelligence. *Computer*, 53(8): 18–28, 2020. doi: 10.1109/MC.2020.2996587.

[6] Khalid Al Khatib, Michael Völske, Shahbaz Syed, Nikolay Kolyada, and Benno Stein. Exploiting personal characteristics of debaters for predicting persuasiveness. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 7067–7072, 2020. doi: 10.18653/v1/2020.acl-main.632.

[7] Abeer ALDayel and Walid Magdy. Stance detection on social media: State of the art and trends. *Information Processing & Management*, 58(4):102597, 2021. ISSN 0306-4573. doi: https://doi.org/10.1016/j.ipm.2021.102597.

[8] Dimitrios Alivanistos, Selene Báez Santamaría, Michael Cochez, Jan-Christoph Kalo, Emile van Krieken, and Thiviyan Thanapalasingam. Prompting as probing: Using language models for knowledge base construction, 2022.

[9] Emily Allaway and Kathleen McKeown. Zero-Shot Stance Detection: A Dataset and Model using Generalized Topic Representations. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 8913–8931, Online, November 2020. Association for Computational Linguistics. doi: 10.18653/v1/2020.emnlp-main.717.

[10] Milad Alshomary, Timon Gurcke, Shahbaz Syed, Philipp Heinisch, Maximilian Spliethöver, Philipp Cimiano, Martin Potthast, and Henning Wachsmuth. Key point analysis via contrastive learning and extractive argument summarization. In *Proceedings of the 8th Workshop on Argument Mining*, pages 184–189, 2021.

[11] Milad Alshomary, Roxanne El Baff, Timon Gurcke, and Henning Wachsmuth. The moral debater: A study on the computational generation of morally framed arguments. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 8782–8797, Dublin, Ireland, May 2022. Association for Computational Linguistics. doi: 10.18653/v1/2022.acl-long.601.

[12] Milad Alshomary, Jonas Rieskamp, and Henning Wachsmuth. Generating contrastive snippets for argument search. In *Computational Models of Argument*, pages 21–31. IOS Press, 2022.

[13] Reinald Kim Amplayo, Stefanos Angelidis, and Mirella Lapata. Aspect-controllable opinion summarization. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 6578–6593, 2021.

[14] Anitha Anandhan, Liyana Shuib, Maizatul Akmar Ismail, and Ghulam Mujtaba. Social media recommender systems: review and open research issues. *IEEE Access*, 6: 15608–15628, 2018.

[15] Stefanos Angelidis and Mirella Lapata. Summarizing opinions: Aspect extraction meets sentiment prediction and they are both weakly supervised. In Ellen Riloff, David Chiang, Julia Hockenmaier, and Jun'ichi Tsujii, editors, *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 3675–3686, Brussels, Belgium, October-November 2018. Association for Computational Linguistics. doi: 10.18653/v1/D18-1403.

[16] Stefanos Angelidis, Reinald Kim Amplayo, Yoshihiko Suhara, Xiaolan Wang, and Mirella Lapata. Extractive opinion summarization in quantized transformer spaces. *Transactions of the Association for Computational Linguistics*, 9:277–293, 2021.

[17] Miguel Arana-Catania, Felix-Anselm Van Lier, Rob Procter, Nataliya Tkachenko, Yulan He, Arkaitz Zubiaga, and Maria Liakata. Citizen participation and machine learning for a better democracy. *Digital Government: Research and Practice*, 2(3):1–22, 2021.

[18] Nikolay Archak, Anindya Ghose, and Panagiotis G Ipeirotis. Show me the money! deriving the pricing power of product features by mining consumer reviews. In *Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 56–65, 2007.

[19] Lisa P. Argyle, Christopher A. Bail, Ethan C. Busby, Joshua R. Gubler, Thomas Howe, Christopher Rytting, Taylor Sorensen, and David Wingate. Leveraging ai for democratic discourse: Chat interventions can improve online political conversations at scale. *Proceedings of the National Academy of Sciences*, 120(41):e2311627120, 2023. doi: 10.1073/pnas.2311627120.

[20] Lisa P Argyle, Ethan C Busby, Nancy Fulda, Joshua R Gubler, Christopher Rytting, and David Wingate. Out of one, many: Using language models to simulate human samples. *Political Analysis*, 31(3):337–351, 2023. doi: 10.1017/pan.2023.2.

[21] Lora Aroyo and Chris Welty. Truth is a lie: Crowd truth and the seven myths of human annotation. *AI Magazine*, 36(1):15–24, Mar. 2015. doi: 10.1609/aimag.v36i1.2564.

[22] Lora Aroyo, Alex S Taylor, Mark Diaz, Christopher M Homan, Alicia Parrish, Greg Serapio-Garcia, Vinodkumar Prabhakaran, and Ding Wang. Dices dataset: Diversity in conversational ai evaluation for safety. *arXiv preprint arXiv:2306.11247*, 2023.

[23] Erfan Sadeqi Azer, Daniel Khashabi, Ashish Sabharwal, and Dan Roth. Not all claims are created equal: Choosing the right statistical approach to assess hypotheses. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 5715–5725, 2020.

[24] Joris Baan, Wilker Aziz, Barbara Plank, and Raquel Fernández. Stop Measuring Calibration When Humans Disagree. *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, pages 1892–1915, 2022. doi: 10.18653/v1/2022.emnlp-main.124.

[25] André Bächtiger, Simon Niemeyer, Michael Neblo, Marco R Steenbergen, and Jürg Steiner. Disentangling diversity in deliberative democracy: Competing theories, their blind spots and complementarities. *Journal of political philosophy*, 18(1), 2010.

[26] André Bächtiger, John S Dryzek, Jane Mansbridge, and Mark Warren. Deliberative democracy. *The Oxford handbook of deliberative democracy*, pages 1–34, 2018.

[27] Christopher A. Bail, Lisa P. Argyle, Taylor W. Brown, John P. Bumpus, Haohan Chen, M. B. Fallin Hunzaker, Jaemin Lee, Marcus Mann, Friedolin Merhout, and Alexander Volfovsky. Exposure to opposing views on social media can increase political polarization. *Proceedings of the National Academy of Sciences*, 115(37):9216–9221, 2018. doi: 10.1073/pnas.1804840115.

[28] Michiel Bakker, Martin Chadwick, Hannah Sheahan, Michael Tessler, Lucy Campbell-Gillingham, Jan Balaguer, Nat McAleese, Amelia Glaese, John Aslanides, Matt Botvinick, et al. Fine-tuning language models to find agreement among humans with diverse preferences. *Advances in Neural Information Processing Systems*, 35:38176–38189, 2022.

[29] Eytan Bakshy, Solomon Messing, and Lada A Adamic. Exposure to ideologically diverse news and opinion on facebook. *Science*, 348(6239):1130–1132, 2015. doi: 10.1126/science.aaa1160.

[30] Dan Bang and Chris D Frith. Making better decisions in groups. *Royal Society open science*, 4(8):170193, 2017.

[31] Roy Bar-Haim, Indrajit Bhattacharya, Francesco Dinuzzo, Amrita Saha, and Noam Slonim. Stance classification of context-dependent claims. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 1, Long Papers*, pages 251–261, 2017.

[32] Roy Bar-Haim, Lilach Eden, Roni Friedman, Yoav Kantor, Dan Lahav, and Noam Slonim. From arguments to key points: Towards automatic argument summarization. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 4029–4039, Online, July 2020. Association for Computational Linguistics, Association for Computational Linguistics. doi: 10.18653/v1/2020.acl-main.371.

[33] Roy Bar-Haim, Yoav Kantor, Lilach Eden, Roni Friedman, Dan Lahav, and Noam Slonim. Quantitative argument summarization and beyond: Cross-domain key point analysis. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 39–49, Online, November 2020. Association for Computational Linguistics. doi: 10.18653/v1/2020.emnlp-main.3.

[34] Roy Bar-Haim, Lilach Eden, Yoav Kantor, Roni Friedman, and Noam Slonim. Every bite is an experience: Key point analysis of business reviews. In Chengqing Zong, Fei Xia, Wenjie Li, and Roberto Navigli, editors, *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 3376–3386, Online, August 2021. Association for Computational Linguistics. doi: 10.18653/v1/2021.acl-long.262.

[35] Roy Bar-Haim, Yoav Kantor, Elad Venezian, Yoav Katz, and Noam Slonim. Project Debater APIs: Decomposing the AI grand challenge. In Heike Adel and Shuming Shi, editors, *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, pages 267–274, Online and Punta Cana, Dominican Republic, November 2021. Association for Computational Linguistics. doi: 10.18653/v1/2021.emnlp-demo.31.

[36] Jason Barabas. How deliberation affects policy opinions. *American political science review*, 98(4):687–701, 2004.

[37] Sugato Basu, Arindam Banerjee, and Raymond J. Mooney. *Active Semi-Supervision for Pairwise Constrained Clustering*, chapter 32, pages 333–344. Society for Industrial and Applied Mathematics, 2004. doi: 10.1137/1.9781611972740.31.

[38] Eric PS Baumer, Mahmood Jasim, Ali Sarvghad, and Narges Mahyar. Of course it's political! a critical inquiry into underemphasized dimensions in civic text visualization. In *Computer Graphics Forum*, volume 41, pages 1–14. Wiley Online Library, 2022.

[39] Connor Baumler, Anna Sotnikova, and Hal Daumé III. Which Examples Should be Multiply Annotated? Active Learning When Annotators May Disagree. In *Findings of the Association for Computational Linguistics: ACL 2023*, pages 10352–10371. ACL, 2023.

[40] Connor Baumler, Anna Sotnikova, and Hal Daumé III. Which examples should be multiply annotated? active learning when annotators may disagree. In Anna Rogers, Jordan Boyd-Graber, and Naoaki Okazaki, editors, *Findings of the Association for Computational Linguistics: ACL 2023*, pages 10352–10371, Toronto, Canada, July 2023. Association for Computational Linguistics. doi: 10.18653/v1/2023.findings-acl.658.

[41] Jordan Beck, Bikalpa Neupane, and John M. Carroll. Managing conflict in online debate communities. *First Monday*, 24(7), Jun. 2019. doi: 10.5210/fm.v24i7.9585.

[42] Tilman Beck, Ji-Ung Lee, Christina Viehmann, Marcus Maurer, Oliver Quiring, and Iryna Gurevych. Investigating label suggestions for opinion mining in German covid-19 social media. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 1–13, Online, August 2021. Association for Computational Linguistics. doi: 10.18653/v1/2021.acl-long.1.

[43] Tilman Beck, Hendrik Schuff, Anne Lauscher, and Iryna Gurevych. How (not) to use sociodemographic information for subjective NLP tasks. *arXiv preprint arXiv:2309.07034*, 2023.

[44] Tilman Beck, Hendrik Schuff, Anne Lauscher, and Iryna Gurevych. Sensitivity, performance, robustness: Deconstructing the effect of sociodemographic prompting. In Yvette Graham and Matthew Purver, editors, *Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 2589–2615, St. Julian's, Malta, March 2024. Association for Computational Linguistics.

[45] Jacob Beel, Tong Xiang, Sandeep Soni, and Diyi Yang. Linguistic characterization of divisive topics online: Case studies on contentiousness in abortion, climate change, and gun control. *Proceedings of the International AAAI Conference on Web and Social Media*, 16(1):32–42, May 2022. doi: 10.1609/icwsm.v16i1.19270.

[46] Emily M Bender, Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell. On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, pages 610–623, 2021.

[47] Jamal Bentahar, Bernard Moulin, and Micheline Bélanger. A taxonomy of argumentation models used for knowledge representation. *The Artificial Intelligence Review*, 33(3):211, 2010.

[48] Surbhi Bhatia. A comparative study of opinion summarization techniques. *IEEE Transactions on Computational Social Systems*, 8(1):110–117, 2020.

[49] Haji Binali, Vidyasagar Potdar, and Chen Wu. A state of the art opinion mining and its application domains. In *2009 IEEE International Conference on Industrial Technology*, pages 1–6. IEEE, 2009.

[50] Or Biran and Owen Rambow. Identifying justifications in written dialogs. In *2011 IEEE Fifth International Conference on Semantic Computing*, pages 162–168, 2011. doi: 10.1109/ICSC.2011.41.

[51] Su Lin Blodgett, Solon Barocas, Hal Daumé III, and Hanna Wallach. Language (technology) is power: A critical survey of "bias"in nlp. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 5454–5476, 2020.

[52] Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10):P10008, oct 2008. doi: 10.1088/1742-5468/2008/10/p10008.

[53] James Bohman. Deliberative democracy and the epistemic benefits of diversity. *Episteme*, 3(3):175–191, 2006.

[54] Filip Boltužić and Jan Šnajder. Identifying prominent arguments in online debates using semantic textual similarity. In *Proceedings of the 2nd Workshop on Argumentation Mining*, pages 110–115, 2015.

[55] Thijs Bouman, Linda Steg, and Henk A. L. Kiers. Measuring values in environmental research: A test of an environmental portrait value questionnaire. *Frontiers in Psychology*, 9, 2018. ISSN 1664-1078. doi: 10.3389/fpsyg.2018.00564.

[56] Samuel R. Bowman and George Dahl. What will it take to fix benchmarking in natural language understanding? In Kristina Toutanova, Anna Rumshisky, Luke Zettlemoyer, Dilek Hakkani-Tur, Iz Beltagy, Steven Bethard, Ryan Cotterell, Tanmoy Chakraborty, and Yichao Zhou, editors, *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 4843–4855, Online, June 2021. Association for Computational Linguistics. doi: 10.18653/v1/2021.naacl-main.385.

[57] Ryan Boyd, Steven Wilson, James Pennebaker, Michal Kosinski, David Stillwell, and Rada Mihalcea. Values in words: Using language to evaluate and understand personal values. *Proceedings of the International AAAI Conference on Web and Social Media*, 9 (1):31–40, Aug. 2021. doi: 10.1609/icwsm.v9i1.14589.

[58] Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. Language models are few-shot learners. *Advances in neural information processing systems*, 33:1877–1901, 2020.

[59] Moritz Büchi, Eduard Fosch-Villaronga, Christoph Lutz, Aurelia Tamò-Larrieux, Shruthi Velidi, and Salome Viljoen. The chilling effects of algorithmic profiling: Mapping the issues. *Computer law & security review*, 36:105367, 2020.

[60] Federico Cabitza, Andrea Campagner, and Valerio Basile. Toward a perspectivist turn in ground truthing for predictive computing. *Proceedings of the AAAI Conference on Artificial Intelligence*, 37(6):6860–6868, Jun. 2023. doi: 10.1609/aaai.v37i6.25840.

[61] Federico Cabitza, Andrea Campagner, and Valerio Basile. Toward a perspectivist turn in ground truthing for predictive computing. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, pages 6860–6868, Jun. 2023. doi: 10.1609/aaai.v37i6.25840.

[62] Elena Cabrio and Serena Villata. Five years of argument mining: A data-driven analysis. In *Proceedings of the 27th International Joint Conference on Artificial Intelligence*, IJCAI'18, page 5427–5433. AAAI Press, 7 2018. ISBN 9780999241127. doi: 10.24963/ijcai.2018/766.

[63] Philippe Caillou, Jonas Renault, Jean-Daniel Fekete, Anne-Catherine Letournel, and Michèle Sebag. Cartolabe: A web-based scalable visualization of large document collections. *IEEE Computer Graphics and Applications*, 41(2):76–88, 2020.

[64] Paul Cairney. *The politics of evidence-based policy making*. Springer, 2016.

[65] Joel Castaño, Silverio Martínez-Fernández, Xavier Franch, and Justus Bogner. Analyzing the evolution and maintenance of ml models on hugging face. *arXiv preprint arXiv:2311.13380*, 2023.

[66] Arie Cattan, Lilach Eden, Yoav Kantor, and Roy Bar-Haim. From key points to key point hierarchy: Structured and expressive opinion summarization. In Anna Rogers, Jordan Boyd-Graber, and Naoaki Okazaki, editors, *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 912–928, Toronto, Canada, July 2023. Association for Computational Linguistics. doi: 10.18653/v1/2023.acl-long.52.

[67] Chengliang Chai, Guoliang Li, Jian Li, Dong Deng, and Jianhua Feng. Cost-effective crowdsourced entity resolution: A partial-order approach. In *Proceedings of the 2016 International Conference on Management of Data*, SIGMOD '16, page 969–984, New York, NY, USA, 2016. Association for Computing Machinery. ISBN 9781450335317. doi: 10.1145/2882903.2915252.

[68] Tuhin Chakrabarty, Christopher Hidey, Smaranda Muresan, Kathy Mckeown, and Alyssa Hwang. Ampersand: Argument mining for persuasive online discussions. *arXiv preprint arXiv:2004.14677*, pages 2933–2943, November 2020. doi: 10.18653/v1/d19-1291.

[69] Yupeng Chang, Xu Wang, Jindong Wang, Yuan Wu, Linyi Yang, Kaijie Zhu, Hao Chen, Xiaoyuan Yi, Cunxiang Wang, Yidong Wang, Wei Ye, Yue Zhang, Yi Chang, Philip S. Yu, Qiang Yang, and Xing Xie. A survey on evaluation of large language models. *ACM Trans. Intell. Syst. Technol.*, 15(3), mar 2024. ISSN 2157-6904. doi: 10.1145/3641289.

[70] Jilin Chen, Gary Hsieh, Jalal U. Mahmud, and Jeffrey Nichols. Understanding individuals' personal values from social media word use. In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing*, CSCW '14, page 405–414, New York, NY, USA, 2014. Association for Computing Machinery. ISBN 9781450325400. doi: 10.1145/2531602.2531608.

[71] Sihao Chen, Daniel Khashabi, Wenpeng Yin, Chris Callison-Burch, and Dan Roth. Seeing things from a different angle: Discovering diverse perspectives about claims. In Jill Burstein, Christy Doran, and Thamar Solorio, editors, *Proceedings of the 2019*

Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), pages 542–557, Minneapolis, Minnesota, June 2019. Association for Computational Linguistics. doi: 10.18653/v1/N19-1053.

[72] Sihao Chen, Siyi Liu, Xander Uyttendaele, Yi Zhang, William Bruno, and Dan Roth. Design challenges for a multi-perspective search engine. In Marine Carpuat, Marie-Catherine de Marneffe, and Ivan Vladimir Meza Ruiz, editors, *Findings of the Association for Computational Linguistics: NAACL 2022*, pages 293–303, Seattle, United States, July 2022. Association for Computational Linguistics. doi: 10.18653/v1/2022.findings-naacl.22.

[73] Liying Cheng, Lidong Bing, Qian Yu, Wei Lu, and Luo Si. Ape: argument pair extraction from peer review and rebuttal via multi-task learning. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 7000–7011, 2020.

[74] Robert Chew, Caroline Kery, Laura Baum, Thomas Bukowski, Annice Kim, and Mario Navarro. Predicting age groups of reddit users based on posting behavior and metadata: Classification model development and validation. *JMIR Public Health Surveill*, 7(3):e25807, Mar 2021. ISSN 2369-2960. doi: 10.2196/25807.

[75] Eric Chu and Peter Liu. Meansum: A neural model for unsupervised multi-document abstractive summarization. In *International Conference on Machine Learning*, pages 1223–1232. PMLR, 2019.

[76] Hyung Won Chung, Thibault Fevry, Henry Tsai, Melvin Johnson, and Sebastian Ruder. Rethinking embedding coupling in pre-trained language models. In *International Conference on Learning Representations*, 2020.

[77] Matteo Cinelli, Gianmarco De Francisci Morales, Alessandro Galeazzi, Walter Quattrociocchi, and Michele Starnini. The echo chamber effect on social media. *Proceedings of the National Academy of Sciences*, 118(9):e2023301118, 2021. doi: 10.1073/pnas.2023301118.

[78] Paul Clough, Robert Gaizauskas, Scott SL Piao, and Yorick Wilks. Measuring text reuse. In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, pages 152–159, 2002.

[79] Katherine M Collins, Umang Bhatt, and Adrian Weller. Eliciting and learning with soft labels from every annotator. In *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing*, volume 10, pages 40–52, 2022.

[80] Luke Collins and Brigitte Nerlich. Examining user comments for deliberative democracy: A corpus-driven analysis of the climate change debate online. In *Climate Change Communication and the Internet*, pages 41–59. Routledge, 2019.

[81] Alexis Conneau, Douwe Kiela, Holger Schwenk, Loïc Barrault, and Antoine Bordes. Supervised learning of universal sentence representations from natural language inference data. In *Proceedings of the 2017 Conference on Empirical Methods in Natural*

Language Processing, pages 670–680, Copenhagen, Denmark, September 2017. Association for Computational Linguistics. doi: 10.18653/v1/D17-1070.

[82] Marta R Costa-jussà, James Cross, Onur Çelebi, Maha Elbayad, Kenneth Heafield, Kevin Heffernan, Elahe Kalbassi, Janice Lam, Daniel Licht, Jean Maillard, et al. No language left behind: Scaling human-centered machine translation. *arXiv preprint arXiv:2207.04672*, 2022.

[83] Michael Crawshaw. Multi-task learning with deep neural networks: A survey. *arXiv preprint arXiv:2009.09796*, 2020.

[84] Scott Crossley, Luc Paquette, Mihai Dascalu, Danielle S McNamara, and Ryan S Baker. Combining click-stream data with nlp tools to better understand mooc completion. In *Proceedings of the sixth international conference on learning analytics & knowledge*, LAK '16, pages 6–14, New York, NY, USA, 2016. Association for Computing Machinery. ISBN 9781450341905. doi: 10.1145/2883851.2883931.

[85] Christophe Croux and Catherine Dehon. Influence functions of the spearman and kendall correlation measures. *Statistical Methods & Applications*, 19(4):497–515, Nov 2010. ISSN 1613-981X. doi: 10.1007/s10260-010-0142-z.

[86] Rowena Cullen. Addressing the digital divide. *Online information review*, 25(5):311–320, 2001. doi: 10.1108/14684520110410517.

[87] Nicole Curato, John S Dryzek, Selen A Ercan, Carolyn M Hendriks, and Simon Niemeyer. Twelve key findings in deliberative democracy research. *Daedalus*, 146 (3):28–38, 2017.

[88] Lincoln Dahlberg. The internet and democratic discourse: Exploring the prospects of online deliberative forums extending the public sphere. *Information, communication & society*, 4(4):615–633, 2001. doi: 10.1080/13691180110097030.

[89] Aida Mostafazadeh Davani, Mark Díaz, and Vinodkumar Prabhakaran. Dealing with disagreements: Looking beyond the majority vote in subjective annotations. *Transactions of the Association for Computational Linguistics*, 10:92–110, 2022.

[90] Jo Davies and Martin Graff. Performance in e-learning: online participation and student grades. *British Journal of Educational Technology*, 36(4):657–663, 2005. doi: https://doi.org/10.1111/j.1467-8535.2005.00542.x.

[91] Todd Davies and Reid Chandler. *Democracy in motion: evaluating the practice and impact of deliberative civic engagement*, chapter Online Deilberation Design: Choices, Criteria, and Evidence, pages 103–131. Oxford University Press, 2012.

[92] Johannes Daxenberger, Steffen Eger, Ivan Habernal, Christian Stab, and Iryna Gurevych. What is the essence of a claim? cross-domain claim identification. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 2055–2066, Copenhagen, Denmark, September 2017. Association for Computational Linguistics. doi: 10.18653/v1/D17-1218.

[93] Johannes Daxenberger, Benjamin Schiller, Chris Stahlhut, Erik Kaiser, and Iryna Gurevych. Argumentext: argument classification and clustering in a generalized search scenario. *Datenbank-Spektrum*, 20:115–121, 2020.

[94] Angel Daza and Anette Frank. X-SRL: A parallel cross-lingual semantic role labeling dataset. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 3904–3914, Online, November 2020. Association for Computational Linguistics. doi: 10.18653/v1/2020.emnlp-main.321.

[95] Maaike H de Boer, Jasper van der Waa, Sophie van Gent, Quirine T.S. Smit, Wouter Korteling, Robin M. van Stokkum, and Mark Neerincx. A contextual hybrid intelligent system design for diabetes lifestyle management. In *International Workshop Modelling and Representing Context, ECAI*, volume 23, 2023.

[96] Christine De Kock and Andreas Vlachos. I beg to differ: A study of constructive disagreement in online conversations. In *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume*, pages 2017–2027, 2021.

[97] Davide Dell'Anna, Pradeep K. Murukannaiah, Bernd Dudzik, Davide Grossi, Catholijn M. Jonker, Catharine Oertel, and Pınar Yolum. Toward a quality model for hybrid intelligence teams. In *Proceedings of the 23rd International Conference on Autonomous Agents and Multiagent Systems*, pages 1–10, Auckland, 2024. To appear.

[98] Dominik Dellermann, Philipp Ebel, Matthias Söllner, and Jan Marco Leimeister. Hybrid intelligence. *Business & Information Systems Engineering*, 61:637–643, 2019. doi: 10.1007/s12599-019-00595-2.

[99] Emily Denton, Alex Hanna, Razvan Amironesei, Andrew Smart, Hilary Nicole, and Morgan Klaus Scheuerman. Bringing the people back in: Contesting benchmark machine learning datasets, 2020.

[100] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 4171–4186, Minneapolis, Minnesota, June 2019. Association for Computational Linguistics. doi: 10.18653/v1/N19-1423.

[101] Thomas G Dietterich. Approximate statistical tests for comparing supervised classification learning algorithms. *Neural computation*, 10(7):1895–1923, 1998.

[102] Zijian Ding, Alison Smith-Renner, Wenjuan Zhang, Joel R Tetreault, and Alejandro Jaimes. Harnessing the power of llms: Evaluating human-ai text co-creation through the lens of news headline generation. In Houda Bouamor, Juan Pino, and Kalika Bali, editors, *The 2023 Conference on Empirical Methods in Natural Language Processing*, pages 3321–3339, Singapore, December 2023. Association for Computational Linguistics. doi: 10.18653/v1/2023.findings-emnlp.217.

[103] Mark Dingemanse and Andreas Liesenfeld. From text to talk: Harnessing conversational corpora for humane and diversity-aware language technology. In Smaranda Muresan, Preslav Nakov, and Aline Villavicencio, editors, *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 5614–5633, Dublin, Ireland, May 2022. Association for Computational Linguistics. doi: 10.18653/v1/2022.acl-long.385.

[104] Tim Draws, Oana Inel, Nava Tintarev, Christian Baden, and Benjamin Timmermans. Comprehensive viewpoint representations for a deeper understanding of user interactions with debated topics. In *Proceedings of the 2022 Conference on Human Information Interaction and Retrieval*, CHIIR '22, pages 135–145, New York, NY, USA, 2022. Association for Computing Machinery. ISBN 9781450391863. doi: 10.1145/3498366.3505812.

[105] Tim Draws, Nirmal Roy, Oana Inel, Alisa Rieger, Rishav Hada, Mehmet Orcun Yalcin, Benjamin Timmermans, and Nava Tintarev. Viewpoint diversity in search results. In *European Conference on Information Retrieval*, pages 279–297. Springer, 2023.

[106] John S Dryzek, André Bächtiger, Simone Chambers, Joshua Cohen, James N Druckman, Andrea Felicetti, James S Fishkin, David M Farrell, Archon Fung, Amy Gutmann, et al. The crisis of democracy and the science of deliberation. *Science*, 363 (6432):1144–1146, 2019. doi: 10.1126/science.aaw2694.

[107] Olive Jean Dunn. Multiple comparisons using rank sums. *Technometrics*, 6(3):241–252, 1964.

[108] Esin Durmus, Karina Nyugen, Thomas I Liao, Nicholas Schiefer, Amanda Askell, Anton Bakhtin, Carol Chen, Zac Hatfield-Dodds, Danny Hernandez, Nicholas Joseph, et al. Towards measuring the representation of subjective global opinions in language models. *arXiv preprint arXiv:2306.16388*, 2023.

[109] Charlie Egan, Advaith Siddharthan, and Adam Wyner. Summarising the points made in online political debates. In *Proceedings of the 3rd Workshop on Argument Mining, The 54th Annual Meeting of the Association for Computational Linguistics*, pages 134–143. Association for Computational Linguistics (ACL), 2016.

[110] Steffen Eger, Johannes Daxenberger, and Iryna Gurevych. Neural end-to-end learning for computational argumentation mining. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 11–22, Vancouver, Canada, July 2017. Association for Computational Linguistics. doi: 10.18653/v1/P17-1002.

[111] Steffen Eger, Johannes Daxenberger, Christian Stab, and Iryna Gurevych. Cross-lingual argumentation mining: Machine translation (and a bit of projection) is all you need! In *Proceedings of the 27th International Conference on Computational Linguistics*, pages 831–844, Santa Fe, New Mexico, USA, August 2018. Association for Computational Linguistics.

[112] Liat Ein-Dor, Eyal Shnarch, Lena Dankin, Alon Halfon, Benjamin Sznajder, Ariel Gera, Carlos Alzate, Martin Gleize, Leshem Choshen, Yufang Hou, Yonatan Bilu, Ranit Aharonov, and Noam Slonim. Corpus wide argument mining—a working solution. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(05):7683–7691, Apr. 2020. doi: 10.1609/aaai.v34i05.6270.

[113] Wafaa S El-Kassas, Cherif R Salama, Ahmed A Rafea, and Hoda K Mohamed. Automatic text summarization: A comprehensive survey. *Expert systems with applications*, 165:113679, 2021.

[114] Tina Eliassi-Rad, Henry Farrell, David Garcia, Stephan Lewandowsky, Patricia Palacios, Don Ross, Didier Sornette, Karim Thébault, and Karoline Wiesner. What science can do for democracy: a complexity science approach. *Humanities and Social Sciences Communications*, 7(1):1–4, 2020.

[115] Stephen Elstub and Peter McLaverty. *Deliberative Democracy: Issues and Cases*. Edinburgh University Press, Edinburgh, 2014. ISBN 9780748643509. doi: doi: 10.1515/9780748643509.

[116] Alexander Erdmann, David Joseph Wrisley, Benjamin Allen, Christopher Brown, Sophie Cohen-Bodénès, Micha Elsner, Yukun Feng, Brian Joseph, Béatrice Joyeux-Prunel, and Marie Catherine de Marneffe. Practical, Efficient, and Customizable Active Learning for Named Entity Recognition in the Digital Humanities. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics*, NAACL '19, pages 2223–2234, Minneapolis, Minnesota, USA, 2019. ACL. doi: 10.18653/v1/n19-1231.

[117] K Esau and Dennis Friess. What creates listening online? exploring reciprocity in online political discussions with relational content analysis. *Journal of Deliberative Democracy*, 18(1):1–16, June 2022. doi: 10.16997/jdd.1021.

[118] Katharina Esau, Dennis Friess, and Christiane Eilders. Design matters! an empirical analysis of online deliberation on different news platforms. *Policy & Internet*, 9(3): 321–342, 2017. doi: 10.1002/poi3.154.

[119] Kevin M Esterling, Archon Fung, and Taeku Lee. How much disagreement is good for democratic deliberation? *Political Communication*, 32(4):529–551, 2015.

[120] Alexander R Fabbri, Wojciech Kryściński, Bryan McCann, Caiming Xiong, Richard Socher, and Dragomir Radev. Summeval: Re-evaluating summarization evaluation. *Transactions of the Association for Computational Linguistics*, 9:391–409, 2021.

[121] Neele Falk, Iman Jundi, Eva Maria Vecchi, and Gabriella Lapesa. Predicting moderation of deliberative arguments: Is argument quality the key? In Khalid Al-Khatib, Yufang Hou, and Manfred Stede, editors, *Proceedings of the 8th Workshop on Argument Mining*, pages 133–141, Punta Cana, Dominican Republic, November 2021. Association for Computational Linguistics. doi: 10.18653/v1/2021.argmining-1.13.

[122] Wei Fang, Moin Nadeem, Mitra Mohtarami, and James Glass. Neural multi-task learning for stance prediction. In *Proceedings of the Second Workshop on Fact Extraction and VERification (FEVER)*, pages 13–19, Hong Kong, China, November 2019. Association for Computational Linguistics. doi: 10.18653/v1/D19-6603.

[123] Andrea Felicetti, Simon Niemeyer, and Nicole Curato. Improving deliberative participation: Connecting mini-publics to deliberative systems. *European Political Science Review*, 8(3):427–448, 2016.

[124] Shangbin Feng, Chan Young Park, Yuhan Liu, and Yulia Tsvetkov. From pretraining data to language models to downstream tasks: Tracking the trails of political biases leading to unfair nlp models. *arXiv preprint arXiv:2305.08283*, pages 11737–11762, July 2023. doi: 10.18653/v1/2023.acl-long.656.

[125] Oliver Ferschke, Iryna Gurevych, and Yevgen Chebotar. Behind the article: Recognizing dialog acts in wikipedia talk pages. In *Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics*, pages 777–786, 2012.

[126] Ken Fischer, Justin Reedy, Cameron Piercy, and Rashmi Thapaliya. A typology of reasoning in deliberative processes: A study of the 2010 oregon citizens'initiative review. *Journal of Deliberative Democracy*, 18(2), 2022.

[127] Lukas Fluri, Daniel Paleka, and Florian Tramèr. Evaluating superhuman models with consistency checks. *arXiv preprint arXiv:2306.09983*, 2023.

[128] Chase Foster and Jeffry Frieden. Crisis of trust: Socio-economic determinants of europeans'confidence in government. *European Union Politics*, 18(4):511–535, 2017.

[129] James B Freeman. *Argument structure:*. Argumentation Library. Springer, Dordrecht, Netherlands, 2011 edition, March 2011. ISBN 978-94-007-3553-8. doi: 10.1007/978-94-007-0357-5.

[130] Batya Friedman, Peter H. Kahn, Alan Borning, and Alina Huldtgren. *Value Sensitive Design and Information Systems*, pages 55–95. Springer Netherlands, Dordrecht, 2013. ISBN 978-94-007-7844-3. doi: 10.1007/978-94-007-7844-3_4.

[131] Roni Friedman-Melamed, Lena Dankin, Yufang Hou, Ranit Aharonov, Yoav Katz, and Noam Slonim. Overview of the 2021 key point analysis shared task. In *Conference on Empirical Methods in Natural Language Processing*, 2021.

[132] Dennis Friess and Christiane Eilders. A systematic review of online deliberation research. *Policy & Internet*, 7(3):319–339, 2015. doi: 10.1002/poi3.95.

[133] Adam D Galinsky, William W Maddux, Debra Gilin, and Judith B White. Why it pays to get inside the head of your opponent: The differential effects of perspective taking and empathy in negotiations. *Psychological science*, 19(4):378–384, 2008.

[134] Tianyu Gao, Xingcheng Yao, and Danqi Chen. SimCSE: Simple contrastive learning of sentence embeddings. In *Empirical Methods in Natural Language Processing (EMNLP)*, 2021.

[135] Albert Gatt and Emiel Krahmer. Survey of the state of the art in natural language generation: Core tasks, applications and evaluation. *Journal of Artificial Intelligence Research*, 61:65–170, 2018.

[136] Hunter Gehlbach, Maureen E Brinkworth, and Ming-Te Wang. The social perspective taking process: What motivates individuals to take another's perspective? *Teachers College Record*, 114(1):1–29, 2012.

[137] Samuel Gehman, Suchin Gururangan, Maarten Sap, Yejin Choi, and Noah A Smith. Realtoxicityprompts: Evaluating neural toxic degeneration in language models. In *Findings of the Association for Computational Linguistics: EMNLP 2020*, pages 3356–3369, 2020.

[138] Robert Geirhos, Jörn-Henrik Jacobsen, Claudio Michaelis, Richard Zemel, Wieland Brendel, Matthias Bethge, and Felix A Wichmann. Shortcut learning in deep neural networks. *Nature Machine Intelligence*, 2(11):665–673, 2020. doi: 10.1038/s42256-020-00257-z.

[139] Debanjan Ghosh, Ritvik Shrivastava, and Smaranda Muresan. "laughing at you or with you": The role of sarcasm in shaping the disagreement space. In *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume*, pages 1998–2010, Online, April 2021. Association for Computational Linguistics. doi: 10.18653/v1/2021.eacl-main.171.

[140] Robert E Goodin and John S Dryzek. Deliberative impacts: The macro-political uptake of mini-publics. *Politics & society*, 34(2):219–244, 2006.

[141] Jesse Graham, Jonathan Haidt, and Brian A Nosek. Liberals and conservatives rely on different sets of moral foundations. *Journal of personality and social psychology*, 96(5):1029–1046, 2009. doi: https://doi.org/10.1037/a0015141.

[142] Jesse Graham, Jonathan Haidt, Sena Koleva, Matt Motyl, Ravi Iyer, Sean P. Wojcik, and Peter H. Ditto. Moral Foundations Theory: The Pragmatic Validity of Moral Pluralism. In *Advances in Experimental Social Psychology*, volume 47, pages 55–130. Elsevier, Amsterdam, the Netherlands, 2013. doi: 10.1016/B978-0-12-407236-7.00002-4.

[143] Jesse Graham, Jonathan Haidt, Sena Koleva, Matt Motyl, Ravi Iyer, Sean P Wojcik, and Peter H Ditto. Moral foundations theory: The pragmatic validity of moral pluralism. In Patricia Devine and Ashby Plant, editors, *Advances in experimental social psychology*, volume 47 of *Advances in Experimental Social Psychology*, pages 55–130. Elsevier, 2013. doi: https://doi.org/10.1016/B978-0-12-407236-7.00002-4.

[144] Shai Gretz, Roni Friedman, Edo Cohen-Karlik, Assaf Toledo, Dan Lahav, Ranit Aharonov, and Noam Slonim. A large-scale dataset for argument quality ranking: Construction and analysis. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(05):7805–7813, Apr. 2020. doi: 10.1609/aaai.v34i05.6285.

[145] Shai Gretz, Assaf Toledo, Roni Friedman, Dan Lahav, Rose Weeks, Naor Bar-zeev, João Sedoc, Pooja Sangha, Yoav Katz, and Noam Slonim. Benchmark data and evaluation framework for intent discovery around covid-19 vaccine hesitancy. In *Findings of the Association for Computational Linguistics: EACL 2023*, pages 1328–1340, 2023.

[146] Leo Groarke. Informal Logic. In Edward N. Zalta and Uri Nodelman, editors, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Spring 2024 edition, 2024.

[147] Maarten Grootendorst. Bertopic: Neural topic modeling with a class-based tf-idf procedure, 2022.

[148] Davide Grossi, Ulrike Hahn, Michael Mäs, Andreas Nitsche, Jan Behrens, Niclas Boehmer, Markus Brill, Ulle Endriss, Umberto Grandi, Adrian Haret, et al. Enabling the digital democratic revival: A research program for digital democracy. *arXiv preprint arXiv:2401.16863*, 2024.

[149] Cornelia Gruber, Katharina Hechinger, Matthias Assenmacher, Göran Kauermann, and Barbara Plank. More labels or cases? assessing label variation in natural language inference. In *Proceedings of the Third Workshop on Understanding Implicit and Underspecified Language*, pages 22–32, Malta, March 2024. Association for Computational Linguistics.

[150] Max Grusky. Rogue scores. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1914–1934, 2023.

[151] Ken Gu and Akshay Budhkar. A package for learning on tabular and text data with transformers. In *Proceedings of the Third Workshop on Multimodal Artificial Intelligence*, pages 69–73, Mexico City, Mexico, June 2021. Association for Computational Linguistics. doi: 10.18653/v1/2021.maiworkshop-1.10.

[152] Zhen Guo, Zhe Zhang, and Munindar Singh. In opinion holders'shoes: Modeling cumulative influence for view change in online argumentation. In *Proceedings of The Web Conference 2020*, pages 2388–2399, 2020. doi: 10.1145/3366423.3380302.

[153] Ali Gürkan, Luca Iandoli, Mark Klein, and Giuseppe Zollo. Mediating debate through on-line large-scale argumentation: Evidence from the field. *Information Sciences*, 180 (19):3686–3702, 2010.

[154] Suchin Gururangan, Swabha Swayamdipta, Omer Levy, Roy Schwartz, Samuel Bowman, and Noah A. Smith. Annotation artifacts in natural language inference data. In Marilyn Walker, Heng Ji, and Amanda Stent, editors, *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 2 (Short Papers)*, pages 107–112, New Orleans, Louisiana, June 2018. Association for Computational Linguistics. doi: 10.18653/v1/N18-2017.

[155] J. Habermas and T. Burger. *The Structural Transformation of the Public Sphere: An Inquiry into a Category of Bourgeois Society*. Studies in Contemporary German Social Thought. MIT Press, 1991. ISBN 9780262581080.

[156] Ivan Habernal and Iryna Gurevych. Argumentation mining in user-generated web discourse. *Computational Linguistics*, 43(1):125–179, April 2017. doi: 10.1162/COLI_a_00276.

[157] Ivan Habernal, Henning Wachsmuth, Iryna Gurevych, and Benno Stein. The argument reasoning comprehension task: Identification and reconstruction of implicit warrants. In Marilyn Walker, Heng Ji, and Amanda Stent, editors, *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*, pages 1930–1940, New Orleans, Louisiana, June 2018. Association for Computational Linguistics. doi: 10.18653/v1/N18-1175.

[158] Rafik Hadfi and Takayuki Ito. Augmented democratic deliberation: Can conversational agents boost deliberation in social media? In *Proceedings of the 21st International Conference on Autonomous Agents and Multiagent Systems*, pages 1794–1798, 2022.

[159] Daniel Halpern and Jennifer Gibbs. Social media as a catalyst for online deliberation? exploring the affordances of facebook and youtube for political expression. *Computers in human behavior*, 29(3):1159–1168, 2013.

[160] Xuehua Han, Juanle Wang, Min Zhang, and Xiaojie Wang. Using social media to mine and analyze public opinion related to covid-19 in china. *International journal of environmental research and public health*, 17(8):2788, 2020.

[161] Momchil Hardalov, Arnav Arora, Preslav Nakov, and Isabelle Augenstein. Cross-domain label-adaptive stance detection. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 9011–9028, Online and Punta Cana, Dominican Republic, November 2021. Association for Computational Linguistics. doi: 10.18653/v1/2021.emnlp-main.710.

[162] Jochen Hartmann, Jasper Schwenzow, and Maximilian Witte. The political ideology of conversational ai: Converging evidence on chatgpt's pro-environmental, left-libertarian orientation. *arXiv preprint arXiv:2301.01768*, 2023.

[163] Brian W Head and John Alford. Wicked problems: Implications for public policy and management. *Administration & society*, 47(6):711–739, 2015.

[164] Jeffrey Heer. Agency plus automation: Designing artificial intelligence into interactive systems. *Proceedings of the National Academy of Sciences*, 116(6):1844–1850, 2019. doi: 10.1073/pnas.1807184115.

[165] Philipp Heinisch, Anette Frank, Juri Opitz, Moritz Plenz, and Philipp Cimiano. Overview of the 2022 validity and novelty prediction shared task. In Gabriella Lapesa, Jodi Schneider, Yohan Jo, and Sougata Saha, editors, *Proceedings of the 9th Workshop on Argument Mining*, pages 84–94, Online and in Gyeongju, Republic of Korea, October 2022. International Conference on Computational Linguistics.

[166] Dan Hendrycks, Collin Burns, Steven Basart, Andrew Critch, Jerry Li, Dawn Song, and Jacob Steinhardt. Aligning ai with shared human values. In *International Conference on Learning Representations*, 2020.

[167] Kristin Page Hocevar, Andrew J Flanagin, and Miriam J Metzger. Social media self-efficacy and information evaluation online. *Computers in Human Behavior*, 39:254–262, 2014. ISSN 0747-5632. doi: https://doi.org/10.1016/j.chb.2014.07.020.

[168] Lu Hong and Scott E Page. Groups of diverse problem solvers can outperform groups of high-ability problem solvers. *Proceedings of the National Academy of Sciences*, 101 (46):16385–16389, 2004.

[169] Joe Hoover, Gwenyth Portillo-Wightman, Leigh Yeh, Shreya Havaldar, Aida Mostafazadeh Davani, Ying Lin, Brendan Kennedy, Mohammad Atari, Zahra Kamel, Madelyn Mendlen, Gabriela Moreno, Christina Park, Tingyee E. Chang, Jenna Chin, Christian Leong, Jun Yen Leung, Arineh Mirinjian, and Morteza Dehghani. Moral foundations twitter corpus: A collection of 35k tweets annotated for moral sentiment. *Social Psychological and Personality Science*, 11(8):1057–1071, 11 2020. ISSN 19485514. doi: 10.1177/1948550619876629.

[170] Dirk Hovy and Shrimai Prabhumoye. Five sources of bias in natural language processing. *Language and Linguistics Compass*, 15(8):e12432, 2021.

[171] Dirk Hovy, Taylor Berg-Kirkpatrick, Ashish Vaswani, and Eduard Hovy. Learning whom to trust with mace. In *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 1120–1130, 2013.

[172] Jeremy Howard and Sebastian Ruder. Universal language model fine-tuning for text classification. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 328–339, Melbourne, Australia, July 2018. Association for Computational Linguistics. doi: 10.18653/v1/P18-1031.

[173] Edward J Hu, yelong shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. LoRA: Low-rank adaptation of large language models. In *International Conference on Learning Representations*, 2022.

[174] J. Edward Hu, Abhinav Singh, Nils Holzenberger, Matt Post, and Benjamin Van Durme. Large-scale, diverse, paraphrastic bitexts via sampling and clustering. In *Proceedings of the 23rd Conference on Computational Natural Language Learning (CoNLL)*, pages 44–54, Hong Kong, China, November 2019. Association for Computational Linguistics. doi: 10.18653/v1/K19-1005.

[175] Lei Huang, Weijiang Yu, Weitao Ma, Weihong Zhong, Zhangyin Feng, Haotian Wang, Qianglong Chen, Weihua Peng, Xiaocheng Feng, Bing Qin, et al. A survey on hallucination in large language models: Principles, taxonomy, challenges, and open questions. *arXiv preprint arXiv:2311.05232*, 2023.

[176] Nannan Huang, Lin Tian, Haytham Fayek, and Xiuzhen Zhang. Examining bias in opinion summarisation through the perspective of opinion diversity. In *Proceedings of the 13th Workshop on Computational Approaches to Subjectivity, Sentiment, & Social Media Analysis*, pages 149–161, Toronto, Canada, July 2023. Association for Computational Linguistics. doi: 10.18653/v1/2023.wassa-1.14.

[177] Dieuwke Hupkes, Mario Giulianelli, Verna Dankers, Mikel Artetxe, Yanai Elazar, Tiago Pimentel, Christos Christodoulopoulos, Karim Lasri, Naomi Saphra, Arabella Sinclair, et al. A taxonomy and review of generalization research in nlp. *Nature Machine Intelligence*, 5(10):1161–1174, 2023.

[178] Luca Iandoli, Ivana Quinto, Anna De Liddo, and Simon Buckingham Shum. Socially augmented argumentation tools: Rationale, design and evaluation of a debate dashboard. *International Journal of Human-Computer Studies*, 72(3):298–319, 2014. ISSN 1071-5819. doi: https://doi.org/10.1016/j.ijhcs.2013.08.006.

[179] IBM Research. The Project Debater Service API. `https://developer.ibm.com/apis/catalog/debater--project-debater-service-api/Introduction`, 2023. Accessed: October 2023.

[180] David Inouye and Jugal K Kalita. Comparing twitter summarization algorithms for multiple post summaries. In *2011 IEEE Third international conference on privacy, security, risk and trust and 2011 IEEE third international conference on social computing*, pages 298–306. IEEE, 2011.

[181] IPCC. Climate change 2022: Impacts, adaptation, and vulnerability. https://www.ipcc.ch/report/ar6/syr/, 2022.

[182] Karen A. Jehn. Enhancing effectiveness: An investigation of advantages and disadvantages of value□based intragroup conflict. *International Journal of Conflict Management*, 5(3):223–238, Jan 1994. ISSN 1044-4068. doi: 10.1108/eb022744.

[183] Hamed Jelodar, Yongli Wang, Rita Orji, and Shucheng Huang. Deep sentiment classification and topic discovery on novel coronavirus or covid-19 online discussions: Nlp using lstm recurrent neural network approach. *IEEE Journal of Biomedical and Health Informatics*, 24(10):2733–2742, 2020. doi: 10.1109/JBHI.2020.3001216.

[184] Shagun Jhaver, Amy Bruckman, and Eric Gilbert. Does transparency in moderation really matter? user behavior after content removal explanations on reddit. *Proc. ACM Hum.-Comput. Interact.*, 3(CSCW), nov 2019. doi: 10.1145/3359252.

[185] Ziwei Ji, Nayeon Lee, Rita Frieske, Tiezheng Yu, Dan Su, Yan Xu, Etsuko Ishii, Ye Jin Bang, Andrea Madotto, and Pascale Fung. Survey of hallucination in natural language generation. *ACM Comput. Surv.*, 55(12), mar 2023. ISSN 0360-0300. doi: 10.1145/3571730.

[186] Albert Q Jiang, Alexandre Sablayrolles, Arthur Mensch, Chris Bamford, Devendra Singh Chaplot, Diego de las Casas, Florian Bressand, Gianna Lengyel, Guillaume Lample, Lucile Saulnier, et al. Mistral 7b. *arXiv preprint arXiv:2310.06825*, 2023.

[187] Xiaoqi Jiao, Yichun Yin, Lifeng Shang, Xin Jiang, Xiao Chen, Linlin Li, Fang Wang, and Qun Liu. Tinybert: Distilling bert for natural language understanding. *arXiv preprint arXiv:1909.10351*, 2019.

[188] Yohan Jo and Alice H Oh. Aspect and sentiment unification model for online review analysis. In *Proceedings of the fourth ACM international conference on Web search and data mining*, WSDM '11, pages 815–824, New York, NY, USA, 2011. Association for Computing Machinery. ISBN 9781450304931. doi: 10.1145/1935826.1935932.

[189] Carolina Centeio Jorge, Emma M van Zoelen, Ruben Verhagen, Siddharth Mehrotra, Catholijn M Jonker, and Myrthe L Tielman. Appropriate context-dependent artificial trust in human-machine teamwork. In *Putting AI in the Critical Loop*, pages 41–60. Elsevier, 2024.

[190] Pratik Joshi, Sebastin Santy, Amar Budhiraja, Kalika Bali, and Monojit Choudhury. The state and fate of linguistic diversity and inclusion in the NLP world. In Dan Jurafsky, Joyce Chai, Natalie Schluter, and Joel Tetreault, editors, *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 6282–6293, Online, July 2020. Association for Computational Linguistics. doi: 10.18653/v1/2020.acl-main.560.

[191] Armand Joulin, Edouard Grave, Piotr Bojanowski, and Tomas Mikolov. Bag of tricks for efficient text classification. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers*, pages 427–431, Valencia, Spain, April 2017. Association for Computational Linguistics.

[192] Kamil Kanclerz, Konrad Karanowski, Julita Bielaniewicz, Marcin Gruza, Piotr Miłkowski, Jan Kocon, and Przemyslaw Kazienko. Pals: Personalized active learning for subjective tasks in nlp. In *The 2023 Conference on Empirical Methods in Natural Language Processing*, pages 13326–13341, 2023.

[193] Nikhil Kandpal, Haikang Deng, Adam Roberts, Eric Wallace, and Colin Raffel. Large language models struggle to learn long-tail knowledge. In *International Conference on Machine Learning*, pages 15696–15707. PMLR, 2023.

[194] Georgi Karadzhov, Andreas Vlachos, and Tom Stafford. The effect of diversity on group decision-making. *arXiv preprint arXiv:2402.01427*, 2024.

[195] Robert E. Kass and Adrian E. Raftery. Bayes factors. *Journal of the American Statistical Association*, 90(430):773–795, 1995. doi: 10.1080/01621459.1995.10476572.

[196] Brendan Kennedy, Mohammad Atari, Aida Mostafazadeh Davani, Joe Hoover, Ali Omrani, Jesse Graham, and Morteza Dehghani. Moral concerns are differentially observable in language. *Cognition*, 212:104696, 2021.

[197] Dana Khartabil, Christopher Collins, Simon Wells, Benjamin Bach, and Jessie Kennedy. Design and evaluation of visualization techniques to facilitate argument exploration. *Computer Graphics Forum*, 40(6):447–465, 2021.

[198] Urja Khurana, Eric Nalisnick, Antske Fokkens, and Swabha Swayamdipta. Crowdcalibrator: Can annotator disagreement inform calibration in subjective tasks? In *First Conference on Language Modeling*, Philadelphia, PA, 2024.

[199] Urja Khurana, Eric Nalisnick, Antske Fokkens, and Swabha Swayamdipta. Crowdcalibrator: Can annotator disagreement inform calibration in subjective tasks?, 2024.

[200] Douwe Kiela, Max Bartolo, Yixin Nie, Divyansh Kaushik, Atticus Geiger, Zhengxuan Wu, Bertie Vidgen, Grusha Prasad, Amanpreet Singh, Pratik Ringshia, Zhiyi Ma, Tristan Thrush, Sebastian Riedel, Zeerak Waseem, Pontus Stenetorp, Robin Jia, Mohit Bansal, Christopher Potts, and Adina Williams. Dynabench: Rethinking benchmarking in NLP. In Kristina Toutanova, Anna Rumshisky, Luke Zettlemoyer, Dilek Hakkani-Tur, Iz Beltagy, Steven Bethard, Ryan Cotterell, Tanmoy Chakraborty, and Yichao Zhou, editors, *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 4110–4124, Online, June 2021. Association for Computational Linguistics. doi: 10.18653/v1/2021.naacl-main.324.

[201] Johannes Kiesel, Milad Alshomary, Nicolas Handke, Xiaoni Cai, Henning Wachsmuth, and Benno Stein. Identifying the human values behind arguments. In Smaranda Muresan, Preslav Nakov, and Aline Villavicencio, editors, *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 4459–4471, Dublin, Ireland, May 2022. Association for Computational Linguistics. doi: 10.18653/v1/2022.acl-long.306.

[202] Johannes Kiesel, Milad Alshomary, Nailia Mirzakhmedova, Maximilian Heinrich, Nicolas Handke, Henning Wachsmuth, and Benno Stein. Semeval-2023 task 4: Valueeval: Identification of human values behind arguments. In Atul Kr. Ojha, A. Seza Doğruöz, Giovanni Da San Martino, Harish Tayyar Madabushi, Ritesh Kumar, and Elisa Sartori, editors, *Proceedings of the 17th International Workshop on Semantic Evaluation (SemEval-2023)*, pages 2287–2303, Toronto, Canada, July 2023. Association for Computational Linguistics. doi: 10.18653/v1/2023.semeval-1.313.

[203] Johannes Kiesel, Milad Alshomary, Nailia Mirzakhmedova, Maximilian Heinrich, Nicolas Handke, Henning Wachsmuth, and Benno Stein. SemEval-2023 Task 4: ValueEval: Identification of Human Values behind Arguments. In *17th International Workshop on Semantic Evaluation*, SemEval '23, pages 2290–2306, Toronto, Canada, July 2023. Association for Computational Linguistics.

[204] Mark Klein. Enabling large-scale deliberation using attention-mediation metrics. *Computer Supported Cooperative Work (CSCW)*, 21(4-5):449–473, 2012. doi: 10.1007/s10606-012-9156-4.

[205] Mark Klein. Crowd-scale deliberation for group decision-making. *Handbook of group decision and negotiation*, pages 355–369, 2021.

[206] Michele Knobel and Colin Lankshear. Digital literacy and participation in online social networking spaces. *Digital literacies: Concepts, policies and practices*, 11:249–278, 2008.

[207] Jonathan Kobbe, Ines Rehbein, Ioana Hulpuș, and Heiner Stuckenschmidt. Exploring morality in argumentation. In *Proceedings of the 7th Workshop on Argument Mining*, pages 30–40, Online, December 2020. Association for Computational Linguistics.

[208] Quyu Kong, Emily Booth, Francesco Bailo, Amelia Johns, and Marian-Andrei Rizoiu. Slipping to the extreme: A mixed method to explain how extreme opinions infiltrate online discussions. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 16, pages 524–535, May 2022. doi: 10.1609/icwsm.v16i1.19312.

[209] Erik Körner, Ahmad Dawar Hakimi, Gerhard Heyer, and Martin Potthast. Casting the same sentiment classification problem. In Marie-Francine Moens, Xuanjing Huang, Lucia Specia, and Scott Wen-tau Yih, editors, *Findings of the Association for Computational Linguistics: EMNLP 2021*, pages 584–590, Punta Cana, Dominican Republic, November 2021. Association for Computational Linguistics. doi: 10.18653/v1/2021.findings-emnlp.53.

[210] Erik Körner, Gregor Wiedemann, Ahmad Dawar Hakimi, Gerhard Heyer, and Martin Potthast. On classifying whether two texts are on the same side of an argument. In Marie-Francine Moens, Xuanjing Huang, Lucia Specia, and Scott Wen-tau Yih, editors, *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 10130–10138, Online and Punta Cana, Dominican Republic, November 2021. Association for Computational Linguistics. doi: 10.18653/v1/2021.emnlp-main.795.

[211] John K. Kruschke. Rejecting or accepting parameter values in bayesian estimation. *Advances in Methods and Practices in Psychological Science*, 1(2):270–280, 2018. doi: 10.1177/2515245918771304.

[212] William H Kruskal and W Allen Wallis. Use of ranks in one-criterion variance analysis. *Journal of the American statistical Association*, 47(260):583–621, 1952.

[213] Gillian Ku, Cynthia S Wang, and Adam D Galinsky. The promise and perversity of perspective-taking in organizations. *Research in Organizational Behavior*, 35:79–102, 2015.

[214] Dilek Küçük and Fazli Can. Stance detection: A survey. *ACM Computing Surveys (CSUR)*, 53(1):1–37, feb 2020. ISSN 0360-0300. doi: 10.1145/3369026.

[215] Andrew P. Kythreotis, Chrystal Mantyka-Pringle, Theresa G. Mercer, Lorraine E. Whitmarsh, Adam Corner, Jouni Paavola, Chris Chambers, Byron A. Miller, and Noel Castree. Citizen social science for more integrative and effective climate action: A science-policy perspective. *Frontiers in Environmental Science*, 7, 2019. ISSN 2296-665X. doi: 10.3389/fenvs.2019.00010.

[216] Philippe Laban, Chien-Sheng Wu, Lidiya Murakhovs'ka, Xiang Chen, and Caiming Xiong. Discord questions: A computational approach to diversity analysis in news coverage. In Yoav Goldberg, Zornitsa Kozareva, and Yue Zhang, editors, *Findings of the Association for Computational Linguistics: EMNLP 2022*, pages 5180–5194, Abu Dhabi, United Arab Emirates, December 2022. Association for Computational Linguistics. doi: 10.18653/v1/2022.findings-emnlp.380.

[217] Preethi Lahoti, Nicholas Blumm, Xiao Ma, Raghavendra Kotikalapudi, Sahitya Potluri, Qijun Tan, Hansa Srinivasan, Ben Packer, Ahmad Beirami, Alex Beutel, and Jilin Chen. Improving diversity of demographic representation in large language models via collective-critiques and self-voting. In Houda Bouamor, Juan Pino, and Kalika Bali, editors, *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 10383–10405, Singapore, December 2023. Association for Computational Linguistics. doi: 10.18653/v1/2023.emnlp-main.643.

[218] Balaji Lakshminarayanan, Alexander Pritzel, and Charles Blundell. Simple and scalable predictive uncertainty estimation using deep ensembles. In I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017.

[219] Hélène Landemore. *Democratic reason: Politics, collective intelligence, and the rule of the many*. Princeton University Press, 2012.

[220] Gabriella Lapesa, Eva Maria Vecchi, Serena Villata, and Henning Wachsmuth. Mining, assessing, and improving arguments in NLP and the social sciences. In Fabio Massimo Zanzotto and Sameer Pradhan, editors, *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics: Tutorial Abstracts*, pages 1–6, Dubrovnik, Croatia, May 2023. Association for Computational Linguistics. doi: 10.18653/v1/2023.eacl-tutorials.1.

[221] Dora C Lau and J Keith Murnighan. Demographic diversity and faultlines: The compositional dynamics of organizational groups. *Academy of management review*, 23(2): 325–340, 1998.

[222] Anne Lauscher, Lily Ng, Courtney Napoles, and Joel Tetreault. Rhetoric, logic, and dialectic: Advancing theory-based argument quality assessment in natural language processing. In Donia Scott, Nuria Bel, and Chengqing Zong, editors, *Proceedings of the 28th International Conference on Computational Linguistics*, pages 4563–4574, Barcelona, Spain (Online), December 2020. International Committee on Computational Linguistics. doi: 10.18653/v1/2020.coling-main.402.

[223] Anne Lauscher, Henning Wachsmuth, Iryna Gurevych, and Goran Glavaš. Scientia potentia est—on the role of knowledge in computational argumentation. *Transactions of the Association for Computational Linguistics*, 10:1392–1422, 2022.

[224] John Lawrence and Chris Reed. Argument Mining: A Survey. *Computational Linguistics*, 45(4):765–818, 01 2020. ISSN 0891-2017. doi: 10.1162/coli_a_00364.

[225] David Lazer, Alex Pentland, Lada Adamic, Sinan Aral, Albert-László Barabási, Devon Brewer, Nicholas Christakis, Noshir Contractor, James Fowler, Myron Gutmann, et al. Computational social science. *Science*, 323(5915):721–723, 2009. doi: 10.1126/science.1167742.

[226] Sabinne Lee, Changho Hwang, and M. Jae Moon. Policy learning and crisis policymaking: quadruple-loop learning and covid-19 responses in south korea. *Policy*

and Society, 39(3):363–381, 2020. doi: 10.1080/14494035.2020.1785195. PMID: 35039726.

[227] Jan Marco Leimeister. Collective intelligence. *Business & Information Systems Engineering*, 2:245–248, 2010.

[228] Elisa Leonardelli, Stefano Menini, Alessio Palmero Aprosio, Marco Guerini, and Sara Tonelli. Agreeing to disagree: Annotating offensive language datasets with annotators' disagreement. In Marie-Francine Moens, Xuanjing Huang, Lucia Specia, and Scott Wen-tau Yih, editors, *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 10528–10539, Online and Punta Cana, Dominican Republic, November 2021. Association for Computational Linguistics. doi: 10.18653/v1/2021.emnlp-main.822.

[229] Roger X. Lera-Leri, Enrico Liscio, Filippo Bistaffa, Catholijn M. Jonker, Maite Lopez-Sanchez, Pradeep K. Murukannaiah, Juan A. Rodriguez-Aguilar, and Francisco Salas-Molina. Aggregating value systems for decision support. *Knowledge-Based Systems*, 287:111453, 2024.

[230] Piyawat Lertvittayakumjorn and Francesca Toni. Explanation-based human debugging of nlp models: A survey. *Transactions of the Association for Computational Linguistics*, 9:1508–1528, 12 2021. ISSN 2307-387X. doi: 10.1162/tacl_a_00440.

[231] Hao Li, Viktor Schlegel, Riza Batista-Navarro, and Goran Nenadic. Do you hear the people sing? key point analysis via iterative clustering and abstractive summarisation. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 14064–14080, Toronto, Canada, July 2023. Association for Computational Linguistics.

[232] Paul Pu Liang, Chiyu Wu, Louis-Philippe Morency, and Ruslan Salakhutdinov. Towards understanding and mitigating social biases in language models. In Marina Meila and Tong Zhang, editors, *International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pages 6565–6576. PMLR, PMLR, 18–24 Jul 2021.

[233] Fu-Ren Lin, Lu-Shih Hsieh, and Fu-Tai Chuang. Discovering genres of online discussion threads via text mining. *Computers & Education*, 52(2):481–495, 2009. ISSN 0360-1315. doi: https://doi.org/10.1016/j.compedu.2008.10.005.

[234] Han Lin and Yonghwan Kim. Learning from disagreement on social media: The mediating role of like-minded and cross-cutting discussion and the moderating role of fact-checking. *Computers in Human Behavior*, 139:107558, 2023. ISSN 0747-5632. doi: https://doi.org/10.1016/j.chb.2022.107558.

[235] Enrico Liscio, Michiel van der Meer, Luciano C. Siebert, Catholijn M. Jonker, Niek Mouter, and Pradeep K. Murukannaiah. Axies: Identifying and evaluating context-specific values. In *Proceedings of the 20th Conference on Autonomous Agents and MultiAgent Systems*, AAMAS '21, pages 799–808, London, 2021.

[236] Enrico Liscio, Michiel van der Meer, Luciano C. Siebert, Catholijn M. Jonker, Niek Mouter, and Pradeep K. Murukannaiah. Axies: Identifying and evaluating context-specific values. In *Proceedings of the 20th International Conference on Autonomous Agents and MultiAgent Systems*, AAMAS '21, page 799–808, Richland, SC, 2021. International Foundation for Autonomous Agents and Multiagent Systems. ISBN 9781450383073.

[237] Enrico Liscio, Alin E Dondera, Andrei Geadau, Catholijn M Jonker, and Pradeep K Murukannaiah. Cross-domain classification of moral values. In *Findings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics. NAACL*, volume 22, pages 1–13, 2022.

[238] Enrico Liscio, Michiel van der Meer, Luciano C. Siebert, Catholijn M. Jonker, and Pradeep K. Murukannaiah. What values should an agent align with? an empirical comparison of general and context-specific values. *Autonomous Agents and Multi-Agent Systems*, 36(1):23, apr 2022. ISSN 1387-2532. doi: 10.1007/s10458-022-09550-0.

[239] Enrico Liscio, Oscar Araque, Lorenzo Gatti, Ionut Constantinescu, Catholijn Jonker, Kyriaki Kalimeri, and Pradeep Kumar Murukannaiah. What does a Text Classifier Learn about Morality? An Explainable Method for Cross-Domain Comparison of Moral Rhetoric. In Anna Rogers, Jordan Boyd-Graber, and Naoaki Okazaki, editors, *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, ACL '23, pages 14113–14132, Toronto, Canada, July 2023. Association for Computational Linguistics. doi: 10.18653/v1/2023.acl-long.789.

[240] Enrico Liscio, Roger Lera-Leri, Filippo Bistaffa, Roel IJ Dobbe, Catholijn M Jonker, Maite Lopez-Sanchez, Juan A Rodriguez-Aguilar, and Pradeep K Murukannaiah. Inferring values via hybrid intelligence. In *HHAI 2023: Augmenting Human Intellect*, pages 373–378. IOS Press, 2023.

[241] Enrico Liscio, Roger Lera-Leri, Filippo Bistaffa, Roel IJ Dobbe, Catholijn M Jonker, Maite Lopez-Sanchez, Juan A Rodriguez-Aguilar, and Pradeep K Murukannaiah. Value inference in sociotechnical systems. In *Proceedings of the 2023 International Conference on Autonomous Agents and Multiagent Systems*, volume 23 of *AAMAS '23*, pages 1774–1780, Richland, SC, 2023. International Foundation for Autonomous Agents and Multiagent Systems. ISBN 9781450394321.

[242] Enrico Liscio, Roger Lera-Leri, Filippo Bistaffa, Roel I.J. Dobbe, Catholijn M. Jonker, Maite Lopez-Sanchez, Juan A. Rodriguez-Aguilar, and Pradeep K. Murukannaiah. Value inference in sociotechnical systems. In *Proceedings of the 2023 International Conference on Autonomous Agents and Multiagent Systems*, AAMAS '23, page 1774–1780, Richland, SC, 2023. International Foundation for Autonomous Agents and Multiagent Systems. ISBN 9781450394321.

[243] Enrico Liscio, Luciano C. Siebert, Catholijn M. Jonker, and Pradeep K. Murukannaiah. Value preferences estimation and disambiguation in hybrid participatory systems, 2024.

[244] Bing Liu. *Sentiment Analysis and Opinion Mining*. Springer International Publishing, 2012. ISBN 9783031021459. doi: 10.1007/978-3-031-02145-9.

[245] Pengfei Liu, Weizhe Yuan, Jinlan Fu, Zhengbao Jiang, Hiroaki Hayashi, and Graham Neubig. Pre-train, prompt, and predict: A systematic survey of prompting methods in natural language processing. *ACM Computing Surveys*, 55(9):1–35, 2023.

[246] Rui Liu, Zheng Lin, Yutong Tan, and Weiping Wang. Enhancing zero-shot and few-shot stance detection with commonsense knowledge graph. In *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, pages 3152–3157, Online, August 2021. Association for Computational Linguistics. doi: 10.18653/v1/2021.findings-acl.278.

[247] Yang Liu and Mirella Lapata. Text summarization with pretrained encoders. *arXiv preprint arXiv:1908.08345*, 2019. doi: 10.18653/v1/d19-1387.

[248] Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. Roberta: A robustly optimized bert pretraining approach. *arXiv preprint arXiv:1907.11692*, 1(1), 2019.

[249] Ines Lörcher and Monika Taddicken. Discussing climate change online. topics and perceptions in online climate change communication in different online public arenas. *Journal of Science Communication*, 16(2):A03, 2017.

[250] Jan Lorenz, Heiko Rauhut, Frank Schweitzer, and Dirk Helbing. How social influence can undermine the wisdom of crowd effect. *Proceedings of the national academy of sciences*, 108(22):9020–9025, 2011. doi: 10.1073/pnas.1008636108.

[251] Philipp Lorenz-Spreen, Lisa Oswald, Stephan Lewandowsky, and Ralph Hertwig. A systematic review of worldwide causal and correlational evidence on digital media and democracy. *Nature human behaviour*, 7(1):74–101, 2023.

[252] Ilya Loshchilov and Frank Hutter. Decoupled weight decay regularization. In *International Conference on Learning Representations*, 2018.

[253] Nicholas Lourie, Ronan Le Bras, and Yejin Choi. Scruples: A corpus of community ethical judgments on 32,000 real-life anecdotes. *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(15):13470–13479, May 2021. doi: 10.1609/aaai.v35i15.17589.

[254] Yao Lu, Max Bartolo, Alastair Moore, Sebastian Riedel, and Pontus Stenetorp. Fantastically ordered prompts and where to find them: Overcoming few-shot prompt order sensitivity. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 8086–8098, 2022.

[255] Yun Luo, Zihan Liu, Stan Z. Li, and Yue Zhang. Improving (dis)agreement detection with inductive social relation information from comment-reply interactions. In *Proceedings of the ACM Web Conference 2023*, WWW '23, page 1584–1593, New York, NY, USA, 2023. Association for Computing Machinery. ISBN 9781450394161. doi: 10.1145/3543507.3583314.

[256] Veronica Lynn, Youngseo Son, Vivek Kulkarni, Niranjan Balasubramanian, and H Andrew Schwartz. Human centered nlp with user-factor adaptation. In Martha Palmer, Rebecca Hwa, and Sebastian Riedel, editors, *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 1146–1155, Copenhagen, Denmark, September 2017. Association for Computational Linguistics. doi: 10.18653/v1/D17-1119.

[257] Qingsong Ma, Johnny Wei, Ondřej Bojar, and Yvette Graham. Results of the WMT19 metrics shared task: Segment-level and strong MT systems pose big challenges. In *Proceedings of the Fourth Conference on Machine Translation (Volume 2: Shared Task Papers, Day 1)*, pages 62–90, Florence, Italy, August 2019. Association for Computational Linguistics. doi: 10.18653/v1/W19-5302.

[258] Noni E. MacDonald. Vaccine hesitancy: Definition, scope and determinants. *Vaccine*, 33(34):4161–4164, 2015. ISSN 0264-410X. doi: https://doi.org/10.1016/j.vaccine.2015.04.036. WHO Recommendations Regarding Vaccine Hesitancy.

[259] David Madras, Toni Pitassi, and Richard Zemel. Predict responsibly: improving fairness and accuracy by learning to defer. *Advances in Neural Information Processing Systems*, 31, 2018.

[260] Jane Mansbridge, James Bohman, Simone Chambers, Thomas Christiano, Archon Fung, John Parkinson, Dennis F Thompson, and Mark E Warren. A systemic approach to deliberative democracy. *Deliberative systems: Deliberative democracy at the large scale*, pages 1–26, 2012. doi: 10.1017/cbo9781139178914.002.

[261] Katerina Margatina and Nikolaos Aletras. On the limitations of simulating active learning. In Anna Rogers, Jordan Boyd-Graber, and Naoaki Okazaki, editors, *Findings of the Association for Computational Linguistics: ACL 2023*, pages 4402–4419, Toronto, Canada, July 2023. Association for Computational Linguistics. doi: 10.18653/v1/2023.findings-acl.269.

[262] Tom McCoy, Ellie Pavlick, and Tal Linzen. Right for the wrong reasons: Diagnosing syntactic heuristics in natural language inference. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 3428–3448, Florence, Italy, July 2019. Association for Computational Linguistics. doi: 10.18653/v1/P19-1334.

[263] Spencer McKay and Chris Tenove. Disinformation as a threat to deliberative democracy. *Political research quarterly*, 74(3):703–717, 2021.

[264] Duncan McLaren, Karen A Parkhill, Adam Corner, Naomi E Vaughan, and Nicholas F Pidgeon. Public conceptions of justice in climate engineering: Evidence from secondary analysis of public deliberation. *Global Environmental Change*, 41:64–73, 2016.

[265] Ninareh Mehrabi, Fred Morstatter, Nripsuta Saxena, Kristina Lerman, and Aram Galstyan. A Survey on Bias and Fairness in Machine Learning. *ACM Computing Surveys*, 54(6), jul 2021. ISSN 0360-0300. doi: 10.1145/3457607.

[266] Bonan Min, Hayley Ross, Elior Sulem, Amir Pouran Ben Veyseh, Thien Huu Nguyen, Oscar Sainz, Eneko Agirre, Ilana Heintz, and Dan Roth. Recent advances in natural language processing via large pre-trained language models: A survey. *ACM Computing Surveys*, 56(2):1–40, sep 2023. ISSN 0360-0300. doi: 10.1145/3605943.

[267] Amita Misra, Pranav Anand, Jean E Fox Tree, and Marilyn Walker. Using summarization to discover argument facets in online idealogical dialog. In *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 430–440, 2015.

[268] Mirjam Moerbeek. Bayesian updating: increasing sample size during the course of a study. *BMC Medical Research Methodology*, 21(1):137, Jul 2021. ISSN 1471-2288. doi: 10.1186/s12874-021-01334-6.

[269] Negar Mokhberian, Myrl Marmarelis, Frederic Hopp, Valerio Basile, Fred Morstatter, and Kristina Lerman. Capturing perspectives of crowdsourced annotators in subjective learning tasks. In Kevin Duh, Helena Gomez, and Steven Bethard, editors, *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers)*, pages 7337–7349, Mexico City, Mexico, June 2024. Association for Computational Linguistics. doi: 10.18653/v1/2024.naacl-long.407.

[270] Ines Montani and Matthew Honnibal. Prodigy: A modern and scriptable annotation tool for creating training data for machine learning models, 2022.

[271] Marlon Mooijman, Joe Hoover, Ying Lin, Heng Ji, and Morteza Dehghani. Moralization in social networks and the emergence of violence during protests. *Nature human behaviour*, 2(6):389–396, 2018. doi: 10.1038/s41562-018-0353-0.

[272] Roser Morante, Chantal Van Son, Isa Maks, and Piek Vossen. Annotating perspectives on vaccination. In Nicoletta Calzolari, Frédéric Béchet, Philippe Blache, Khalid Choukri, Christopher Cieri, Thierry Declerck, Sara Goggi, Hitoshi Isahara, Bente Maegaard, Joseph Mariani, Hélène Mazo, Asuncion Moreno, Jan Odijk, and Stelios Piperidis, editors, *Proceedings of the Twelfth Language Resources and Evaluation Conference*, pages 4964–4973, Marseille, France, May 2020. European Language Resources Association. ISBN 979-10-95546-34-4.

[273] Eduardo Mosqueira-Rey, Elena Hernández-Pereira, David Alonso-Ríos, José Bobes-Bascarán, and Ángel Fernández-Leal. Human-in-the-loop machine learning: a state of the art. *Artificial Intelligence Review*, 56(4):3005–3054, 2023.

[274] Niek Mouter, Jose Ignacio Hernandez, and Anatol Valerian Itten. Public participation in crisis policymaking. how 30,000 dutch citizens advised their government on relaxing covid-19 lockdown measures. *PLOS ONE*, 16(5):1–42, 05 2021. doi: 10.1371/journal.pone.0250614.

[275] Niek Mouter, Paul Koster, and Thijs Dekker. Contrasting the recommendations of participatory value evaluation and cost-benefit analysis in the context of urban mobility investments. *Transportation research part A: policy and practice*, 144:54–73, 2021.

[276] Niek Mouter, Ruth M Shortall, Shannon L Spruit, and Anatol V Itten. Including young people, cutting time and producing useful outcomes: Participatory value evaluation as a new practice of public participation in the dutch energy transition. *Energy Research & Social Science*, 75:101965, 2021.

[277] Rafael Müller, Simon Kornblith, and Geoffrey E Hinton. When does label smoothing help? *Advances in neural information processing systems*, 32, 2019.

[278] Eni Mustafaraj, Samantha Finn, Carolyn Whitlock, and Panagiotis T Metaxas. Vocal minority versus silent majority: Discovering the opionions of the long tail. In *2011 IEEE Third International Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing*, pages 103–110. IEEE, 2011. doi: 10.1109/PASSAT/SocialCom.2011.188.

[279] Yoon-Eui Nahm. A novel approach to prioritize customer requirements in qfd based on customer satisfaction function for customer-oriented product design. *Journal of Mechanical Science and Technology*, 27:3765–3777, 2013.

[280] United Nations. Transforming our world: the 2030 agenda for sustainable development. https://sdgs.un.org/2030agenda, 2015.

[281] Ani Nenkova, Kathleen McKeown, et al. Automatic summarization. *Foundations and Trends® in Information Retrieval*, 5(2–3):103–233, 2011.

[282] German Neubaum and Nicole C Krämer. Monitoring the opinion of the crowd: Psychological mechanisms underlying public opinion perceptions on social media. *Media psychology*, 20(3):502–531, 2017. doi: 10.1080/15213269.2016.1211539.

[283] M. E. J. Newman, D. J. Watts, and S. H. Strogatz. Random graph models of social networks. *Proceedings of the National Academy of Sciences*, 99(suppl_1):2566–2572, 2002. doi: 10.1073/pnas.012582999.

[284] Vlad Niculae and Cristian Danescu-Niculescu-Mizil. Conversational markers of constructive discussions. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 568–578, 2016.

[285] Vlad Niculae, Joonsuk Park, and Claire Cardie. Argument mining with structured SVMs and RNNs. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 985–995, Vancouver, Canada, July 2017. Association for Computational Linguistics. doi: 10.18653/v1/P17-1091.

[286] Yixin Nie, Xiang Zhou, and Mohit Bansal. What can we learn from collective human opinions on natural language inference data? In Bonnie Webber, Trevor Cohn, Yulan He, and Yang Liu, editors, *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 9131–9143, Online, November 2020. Association for Computational Linguistics. doi: 10.18653/v1/2020.emnlp-main.734.

[287] Sem Nouws, Íñigo Martinez De Rituerto De Troya, Roel Dobbe, and Marijn Janssen. Diagnosing and addressing emergent harms in the design process of public ai and algorithmic systems. In *Proceedings of the 24th Annual International Conference on Digital Government Research*, pages 679–681, 2023.

[288] Humphrey O. Obie, Waqar Hussain, Xin Xia, John Grundy, Li Li, Burak Turhan, Jon Whittle, and Mojtaba Shahin. A first look at human values-violation in app reviews. In *2021 IEEE/ACM 43rd International Conference on Software Engineering: Software Engineering in Society (ICSE-SEIS)*, pages 29–38, 2021. doi: 10.1109/ICSE-SEIS52602.2021.00012.

[289] OECD. *Innovative Citizen Participation and New Democratic Institutions*. OECD, 2020. doi: https://doi.org/https://doi.org/10.1787/339306da-en.

[290] OpenAI. The OpenAI Python library. `https://github.com/openai/openai-python`, 2023. Accessed: October 2023.

[291] Matthias Orlikowski, Paul Röttger, Philipp Cimiano, and Dirk Hovy. The Ecological Fallacy in Annotation: Modeling Human Label Variation goes beyond Sociodemographics. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics Volume 2: Short Papers*, pages 1017–1029. ACL, 2023. ISBN 9781959429715. doi: 10.18653/v1/2023.acl-short.88.

[292] Elinor Ostrom. *Governing the commons: The evolution of institutions for collective action*. Cambridge University Press, Cambridge, UK, 1990.

[293] Long Ouyang, Jeffrey Wu, Xu Jiang, Diogo Almeida, Carroll Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, et al. Training language models to follow instructions with human feedback. *Advances in Neural Information Processing Systems*, 35:27730–27744, 2022.

[294] Siru Ouyang, Shuohang Wang, Yang Liu, Ming Zhong, Yizhu Jiao, Dan Iter, Reid Pryzant, Chenguang Zhu, Heng Ji, and Jiawei Han. The shifted and the overlooked: A task-oriented investigation of user-gpt interactions. In Houda Bouamor, Juan Pino, and Kalika Bali, editors, *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 2375–2393, Singapore, December 2023. Association for Computational Linguistics. doi: 10.18653/v1/2023.emnlp-main.146.

[295] Maria Leonor Pacheco, Tunazzina Islam, Lyle Ungar, Ming Yin, and Dan Goldwasser. Interactive concept learning for uncovering latent themes in large text collections. In Anna Rogers, Jordan Boyd-Graber, and Naoaki Okazaki, editors, *Findings of the Association for Computational Linguistics: ACL 2023*, pages 5059–5080, Toronto, Canada, July 2023. Association for Computational Linguistics. doi: 10.18653/v1/2023.findings-acl.313.

[296] Raquel Mochales Palau and Marie-Francine Moens. Argumentation mining: the detection, classification and structure of arguments in text. In *Proceedings of the 12th international conference on artificial intelligence and law*, pages 98–107, 2009.

[297] Zizi Papacharissi. The virtual sphere: The internet as a public sphere. *New media & society*, 4(1):9–27, 2002.

[298] John Parkinson. Legitimacy problems in deliberative democracy. *Political studies*, 51 (1):180–196, 2003.

[299] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12(85):2825–2830, 2011.

[300] Joshua C Peterson, Ruairidh M Battleday, Thomas L Griffiths, and Olga Russakovsky. Human uncertainty makes classification more robust. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 9617–9626, 2019.

[301] Hoang Phan, Long Nguyen, and Khanh Doan. Matching the statements: A simple and accurate model for key point analysis. In *Proceedings of the 8th Workshop on Argument Mining*, pages 165–174, 2021.

[302] Barbara Plank. The "problem"of human label variation: On ground truth in data, modeling and evaluation. In Yoav Goldberg, Zornitsa Kozareva, and Yue Zhang, editors, *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, pages 10671–10682, Abu Dhabi, United Arab Emirates, December 2022. Association for Computational Linguistics. doi: 10.18653/v1/2022.emnlp-main.731.

[303] Francesca Polletta and Beth Gardner. The Forms of Deliberative Communication. In *The Oxford Handbook of Deliberative Democracy*. Oxford University Press, 09 2018. ISBN 9780198747369. doi: 10.1093/oxfordhb/9780198747369.013.45.

[304] Vladimir Ponizovskiy, Murat Ardag, Lusine Grigoryan, Ryan Boyd, Henrik Dobewall, and Peter Holtz. Development and validation of the personal values dictionary: A theory–driven tool for investigating references to basic human values in text. *European Journal of Personality*, 34(5):885–902, 2020. doi: 10.1002/per.2294.

[305] John Pougué-Biyong, Valentina Semenova, Alexandre Matton, Rachel Han, Aerin Kim, Renaud Lambiotte, and Doyne Farmer. Debagreement: A comment-reply dataset for (dis) agreement detection in online debates. In *Thirty-fifth Conference on Neural Information Processing Systems Datasets and Benchmarks Track (Round 2)*, 2021.

[306] Katrin Preckel, Philipp Kanske, and Tania Singer. On the interaction of social affect and cognition: empathy, compassion and theory of mind. *Current Opinion in Behavioral Sciences*, 19:1–6, 2018.

[307] Vincent Price. Social Identification and Public Opinion: Effects of Communicating Group Conflict. *Public Opinion Quarterly*, 53(2):197–224, 01 1989. ISSN 0033-362X. doi: 10.1086/269503.

[308] Nicholas Proferes, Naiyan Jones, Sarah Gilbert, Casey Fiesler, and Michael Zimmer. Studying Reddit: A systematic overview of disciplines, approaches, methods, and ethics. *Social Media + Society*, 7(2):20563051211019004, 2021. doi: 10.1177/20563051211019004.

[309] Yada Pruksachatkun, Jason Phang, Haokun Liu, Phu Mon Htut, Xiaoyi Zhang, Richard Yuanzhe Pang, Clara Vania, Katharina Kann, and Samuel R. Bowman. Intermediate-task transfer learning with pretrained language models: When and why does it work? In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 5231–5247, Online, July 2020. Association for Computational Linguistics. doi: 10.18653/v1/2020.acl-main.467.

[310] James Pustejovsky, José M Castano, Robert Ingria, Roser Sauri, Robert J Gaizauskas, Andrea Setzer, Graham Katz, and Dragomir R Radev. Timeml: Robust specification of event and temporal expressions in text. *New directions in question answering*, 3: 28–34, 2003.

[311] Liang Qiu, Yizhou Zhao, Jinchao Li, Pan Lu, Baolin Peng, Jianfeng Gao, and Song-Chun Zhu. Valuenet: A new dataset for human value driven dialogue system. *arXiv preprint arXiv:2112.06346*, 2021.

[312] Alexander J Quinn and Benjamin B Bederson. Human computation: a survey and taxonomy of a growing field. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 1403–1412, 2011.

[313] John Rawls. *A Theory of Justice*. Oxford University Press, Oxford, 1973.

[314] Nils Reimers and Iryna Gurevych. Sentence-bert: Sentence embeddings using siamese bert-networks. In Kentaro Inui, Jing Jiang, Vincent Ng, and Xiaojun Wan, editors, *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 3982–3992, Hong Kong, China, November 2019. Association for Computational Linguistics. doi: 10.18653/v1/D19-1410.

[315] Nils Reimers, Benjamin Schiller, Tilman Beck, Johannes Daxenberger, Christian Stab, and Iryna Gurevych. Classification and clustering of arguments with contextualized word embeddings. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 567–578, Florence, Italy, July 2019. Association for Computational Linguistics. doi: 10.18653/v1/P19-1054.

[316] Pengzhen Ren, Yun Xiao, Xiaojun Chang, Po Yao Huang, Zhihui Li, Brij B. Gupta, Xiaojiang Chen, and Xin Wang. A Survey of Deep Active Learning. *ACM Computing Surveys*, 54(9):1–40, 2021. doi: 10.1145/3472291.

[317] Bram M Renting, Holger H Hoos, and Catholijn M Jonker. Automated configuration and usage of strategy portfolios for mixed-motive bargaining. In *Proceedings of the 21st International Conference on Autonomous Agents and Multiagent Systems*, AAMAS '22, pages 1101–1109, Richland, SC, 2022. International Foundation for Autonomous Agents and Multiagent Systems. ISBN 9781450392136.

[318] Myrthe Reuver, Nicolas Mattis, Marijn Sax, Suzan Verberne, Nava Tintarev, Natali Helberger, Judith Moeller, Sanne Vrijenhoek, Antske Fokkens, and Wouter van Atteveldt. Are we human, or are we users? the role of natural language processing in human-centric news recommenders that nudge users to diverse content. In *Proceedings of the 1st Workshop on NLP for Positive Impact*, pages 47–59, 2021.

[319] Myrthe Reuver, Suzan Verberne, Roser Morante, and Antske Fokkens. Is stance detection topic-independent and cross-topic generalizable?-a reproduction study. In *Proceedings of the 8th Workshop on Argument Mining*, pages 46–56, 2021.

[320] Horst WJ Rittel and Melvin M Webber. Dilemmas in a general theory of planning. *Policy sciences*, 4(2):155–169, 1973.

[321] Milton Rokeach. Rokeach value survey. *The nature of human values.*, 1967. doi: 10.1037/t01381-000.

[322] Julia Romberg. Is your perspective also my perspective? enriching prediction with subjectivity. In *Proceedings of the 9th Workshop on Argument Mining*, pages 115–125, 2022.

[323] Andrew Rosenberg and Julia Hirschberg. V-measure: A conditional entropy-based external cluster evaluation measure. In Jason Eisner, editor, *Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL)*, pages 410–420, Prague, Czech Republic, June 2007. Association for Computational Linguistics.

[324] Jeffrey N. Rouder, Paul L. Speckman, Dongchu Sun, Richard D. Morey, and Geoffrey Iverson. Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, 16(2):225–237, Apr 2009. ISSN 1531-5320. doi: 10.3758/PBR.16.2.225.

[325] Gene Rowe and Lynn J Frewer. Public participation methods: a framework for evaluation. *Science, technology, & human values*, 25(1):3–29, 2000.

[326] Shamik Roy, Maria Leonor Pacheco, and Dan Goldwasser. Identifying morality frames in political tweets using relational learning. In Marie-Francine Moens, Xuanjing Huang, Lucia Specia, and Scott Wen-tau Yih, editors, *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 9939–9958, Online and Punta Cana, Dominican Republic, November 2021. Association for Computational Linguistics. doi: 10.18653/v1/2021.emnlp-main.783.

[327] Ramon Ruiz-Dolz, Jose Alemany, Stella M Heras Barberá, and Ana García-Fornes. Transformer-based models for automatic identification of argument relations: A cross-domain evaluation. *IEEE Intelligent Systems*, 36(6):62–70, 2021.

[328] Pratik Sachdeva, Renata Barreto, Geoff Bacon, Alexander Sahn, Claudia von Vacano, and Chris Kennedy. The measuring hate speech corpus: Leveraging rasch measurement theory for data perspectivism. In Gavin Abercrombie, Valerio Basile, Sara Tonelli, Verena Rieser, and Alexandra Uma, editors, *Proceedings of the 1st Workshop*

on Perspectivist Approaches to NLP @LREC2022, pages 83–94, Marseille, France, June 2022. European Language Resources Association.

[329] Nithya Sambasivan, Shivani Kapania, Hannah Highfill, Diana Akrong, Praveen Paritosh, and Lora M Aroyo. "everyone wants to do the model work, not the data work": Data cascades in high-stakes ai. In *proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–15, 2021.

[330] Ahnaf Mozib Samin, Behrooz Nikandish, and Jingyan Chen. Arguments to key points mapping with prompt-based learning. *ICNLSP 2022*, page 303, 2022.

[331] Selene Baez Santamaría, Piek Vossen, and Thomas Baier. Evaluating agent interactions through episodic knowledge graphs. In Heuiseok Lim, Seungryong Kim, Yeonsoo Lee, Steve Lin, Paul Hongsuck Seo, Yumin Suh, Yoonna Jang, Jungwoo Lim, Yuna Hur, and Suhyune Son, editors, *Proceedings of the 1st Workshop on Customized Chat Grounding Persona and Knowledge@ COLING2022*, pages 15–28, Gyeongju, Republic of Korea, October 2022. Association for Computational Linguistics (ACL), Association for Computational Linguistics.

[332] Filippo Santoni de Sio and Jeroen Van den Hoven. Meaningful human control over autonomous systems: A philosophical account. *Frontiers in Robotics and AI*, 5:15, 2018.

[333] Shibani Santurkar, Esin Durmus, Faisal Ladhak, Cinoo Lee, Percy Liang, and Tatsunori Hashimoto. Whose opinions do language models reflect? *arXiv preprint arXiv:2303.17548*, 202:29971–30004, 23–29 Jul 2023.

[334] Sebastin Santy, Jenny Liang, Ronan Le Bras, Katharina Reinecke, and Maarten Sap. NLPositionality: Characterizing design biases of datasets and models. In Anna Rogers, Jordan Boyd-Graber, and Naoaki Okazaki, editors, *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 9080–9102, Toronto, Canada, July 2023. Association for Computational Linguistics. doi: 10.18653/v1/2023.acl-long.505.

[335] Martin Saveski, Nabeel Gillani, Ann Yuan, Prashanth Vijayaraghavan, and Deb Roy. Perspective-taking to reduce affective polarization on social media. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 16, pages 885–895, 2022.

[336] Akrati Saxena and Harita Reddy. Users roles identification on online crowdsourced q&a platforms and encyclopedias: a survey. *Journal of Computational Social Science*, 5(1):285–317, 2022. doi: 10.1007/s42001-021-00125-9.

[337] Rylan Schaeffer, Brando Miranda, and Sanmi Koyejo. Are emergent abilities of large language models a mirage? *Advances in Neural Information Processing Systems*, 36: 55565–55581, 2024.

[338] Mike Schaekermann, Joslin Goh, Kate Larson, and Edith Law. Resolvable vs. irresolvable disagreement: A study on worker deliberation in crowd work. *Proceedings of the ACM on Human-Computer Interaction*, 2(CSCW):1–19, 2018.

[339] Patrick Schober, Christa Boer, and Lothar A Schwarte. Correlation coefficients: appropriate use and interpretation. *Anesthesia & analgesia*, 126(5):1763–1768, 2018.

[340] Mirco Schönfeld, Steffen Eckhard, Ronny Patz, and Hilde van Meegdenburg. The un security council debates 1995-2017. *arXiv preprint arXiv:1906.10969*, 2019.

[341] Claudia Schulz, Christian M. Meyer, Jan Kiesewetter, Michael Sailer, Elisabeth Bauer, Martin R. Fischer, Frank Fischer, and Iryna Gurevych. Analysis of automatic annotation suggestions for hard discourse-level tasks in expert domains. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 2761–2772, Florence, Italy, July 2019. Association for Computational Linguistics. doi: 10.18653/v1/P19-1265.

[342] Stefan Schulz-Hardt, Dieter Frey, Carsten Lüthgens, and Serge Moscovici. Biased information search in group decision making. *Journal of Personality and Social Psychology*, 78(4):655–669, 2000. doi: 10.1037/0022-3514.78.4.655.

[343] Shalom Schwartz. A repository of schwartz value scales with instructions and an introduction. *Online Readings in Psychology and Culture*, 2, 09 2021. doi: 10.9707/2307-0919.1173.

[344] Shalom H Schwartz. An overview of the schwartz theory of basic values. *Online readings in Psychology and Culture*, 2(1):11, 2012. doi: 10.9707/2307-0919.1116.

[345] Shalom H Schwartz, Jan Cieciuch, Michele Vecchione, Eldad Davidov, Ronald Fischer, Constanze Beierlein, Alice Ramos, Markku Verkasalo, Jan-Erik Lönnqvist, Kursad Demirutku, et al. Refining the theory of basic individual values. *Journal of personality and social psychology*, 103(4):663, 2012. doi: https://doi.org/10.1037/a0029393.

[346] Abigail See, Peter J. Liu, and Christopher D. Manning. Get to the point: Summarization with pointer-generator networks. In Regina Barzilay and Min-Yen Kan, editors, *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1073–1083, Vancouver, Canada, July 2017. Association for Computational Linguistics. doi: 10.18653/v1/P17-1099.

[347] Thibault Sellam, Dipanjan Das, and Ankur Parikh. Bleurt: Learning robust metrics for text generation. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 7881–7892, 2020.

[348] Bryan C Semaan, Scott P Robertson, Sara Douglas, and Misa Maruyama. Social media supporting political deliberation across multiple public spheres: towards depolarization. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*, pages 1409–1421, 2014.

[349] Rico Sennrich, Barry Haddow, and Alexandra Birch. Improving neural machine translation models with monolingual data. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 86–96, Berlin, Germany, August 2016. Association for Computational Linguistics. doi: 10.18653/v1/P16-1009.

[350] Burr Settles. *Active Learning*. Morgan & Claypool, 2012.

[351] Eyal Shnarch, Ariel Gera, Alon Halfon, Lena Dankin, Leshem Choshen, Ranit Aharonov, and Noam Slonim. Cluster & tune: Boost cold start performance in text classification. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 7639–7653, 2022.

[352] Ruth Shortall, Anatol Itten, Michiel van der Meer, Pradeep K. Murukannaiah, and Catholijn M. Jonker. Reason against the machine? future directions for mass online deliberation. *Frontiers in Political Science*, 4:01–17, 2022. ISSN 2673-3145. doi: 10.3389/fpos.2022.946589.

[353] Patrick E Shrout and Joseph L Fleiss. Intraclass correlations: uses in assessing rater reliability. *Psychological bulletin*, 86(2):420, 1979.

[354] Luciano C Siebert, Enrico Liscio, Pradeep K Murukannaiah, Lionel Kaptein, Shannon Spruit, Jeroen van den Hoven, and Catholijn Jonker. Estimating value preferences in a hybrid participatory system. In *Proceedings of the first International Conference on Hybrid Human-Artificial Intelligence (HHAI 2022)*, pages 1–14, Amsterdam, the Netherlands, 2022. IOS Press.

[355] Tejpalsingh Siledar, Jigar Makwana, and Pushpak Bhattacharyya. Aspect-sentiment-based opinion summarization using multiple information sources. In *Proceedings of the 6th Joint International Conference on Data Science & Management of Data (10th ACM IKDD CODS and 28th COMAD)*, pages 55–61, 2023.

[356] Amila Silva, Pei-Chi Lo, and Ee Peng Lim. On predicting personal values of social media users using community-specific language features and personal value correlation. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 15, pages 680–690, 2021.

[357] Julius Sim and Chris C Wright. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Physical therapy*, 85(3):257–268, 2005.

[358] Bernd Skiera, Shunyao Yan, Johannes Daxenberger, Marcus Dombois, and Iryna Gurevych. Using information-seeking argument mining to improve service. *Journal of Service Research*, 25(4):537–548, 2022. doi: 10.1177/10946705221110845.

[359] Gabriella Skitalinskaya, Jonas Klaff, and Henning Wachsmuth. Learning from revisions: Quality assessment of claims in argumentation at scale. In Paola Merlo, Jorg Tiedemann, and Reut Tsarfaty, editors, *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume*, pages 1718–1729, Online, April 2021. Association for Computational Linguistics. doi: 10.18653/v1/2021.eacl-main.147.

[360] Graham Smith, Robert C Richards Jr, and John Gastil. The potential of participedia as a crowdsourcing tool for comparative analysis of democratic innovations. *Policy & Internet*, 7(2):243–262, 2015.

[361] Anders Søgaard, Sebastian Ebert, Jasmijn Bastings, and Katja Filippova. We need to talk about random splits. In *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume*, pages 1823–1832, Online, April 2021. Association for Computational Linguistics. doi: 10.18653/v1/2021.eacl-main.156.

[362] Abdurashid Solijonov. International idea: Voter turnout trends around the world. `https://www.idea.int/publications/catalogue/voter-turnout-trends-around-world`, 2016. Accessed: 2024-02-19.

[363] Hyunjin Song, Jaeho Cho, and Grace A Benefield. The dynamics of message selection in online political discussion forums: Self-segregation or diverse exposure? *Communication Research*, 47(1):125–152, 2020. doi: 10.1177/0093650218790144.

[364] Taylor Sorensen, Joshua Robinson, Christopher Michael Rytting, Alexander Glenn Shaw, Kyle Jeffrey Rogers, Alexia Pauline Delorey, Mahmoud Khalil, Nancy Fulda, and David Wingate. An information-theoretic approach to prompt engineering without ground truth labels, 2022.

[365] Christian Stab and Iryna Gurevych. Annotating argument components and relations in persuasive essays. In *Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers*, pages 1501–1510, Dublin, Ireland, August 2014. Dublin City University and Association for Computational Linguistics.

[366] Christian Stab, Johannes Daxenberger, Chris Stahlhut, Tristan Miller, Benjamin Schiller, Christopher Tauchmann, Steffen Eger, and Iryna Gurevych. ArgumenText: Searching for arguments in heterogeneous sources. In Yang Liu, Tim Paek, and Manasi Patwardhan, editors, *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Demonstrations*, pages 21–25, New Orleans, Louisiana, June 2018. Association for Computational Linguistics. doi: 10.18653/v1/N18-5005.

[367] Christian Stab, Tristan Miller, Benjamin Schiller, Pranav Rai, and Iryna Gurevych. Cross-topic argument mining from heterogeneous sources. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 3664–3674, Brussels, Belgium, October-November 2018. Association for Computational Linguistics. doi: 10.18653/v1/D18-1402.

[368] Marco R Steenbergen, André Bächtiger, Markus Spörndli, and Jürg Steiner. Measuring political deliberation: A discourse quality index. *Comparative European Politics*, 1(1): 21–48, 2003. doi: 10.1057/palgrave.cep.6110002.

[369] Cor Steging, Silja Renooij, and Bart Verheij. Discovering the rationale of decisions: towards a method for aligning learning and reasoning. In *Proceedings of the Eighteenth International Conference on Artificial Intelligence and Law*, pages 235–239, 2021.

[370] Kim Strandberg and Kimmo Grönlund. Online deliberation. *The Oxford handbook of deliberative democracy*, pages 365–377, 2018.

[371] Jennifer Stromer-Galley and Peter Muhlberger. Agreement and disagreement in group deliberation: Effects on deliberation satisfaction, future engagement, and decision legitimacy. *Political Communication*, 26(2):173–192, 2009. doi: 10.1080/10584600902850775.

[372] Jennifer Stromer-Galley, Lauren Bryant, and Bruce Bimber. Context and medium matter: Expressing disagreements online and face-to-face in political deliberations. *Journal of Deliberative Democracy*, 11(1), 2020. doi: https://doi.org/10.16997/jdd.218.

[373] Yoshihiko Suhara, Xiaolan Wang, Stefanos Angelidis, and Wang-Chiew Tan. Opiniondigest: A simple framework for opinion summarization. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 5789–5798, 2020.

[374] Shiliang Sun, Chen Luo, and Junyu Chen. A review of natural language processing techniques for opinion mining systems. *Information fusion*, 36:10–25, 2017. ISSN 1566-2535. doi: https://doi.org/10.1016/j.inffus.2016.10.004.

[375] James Surowiecki. *The wisdom of crowds*. Anchor, 2005.

[376] Shahbaz Syed, Khalid Al Khatib, Milad Alshomary, Henning Wachsmuth, and Martin Potthast. Generating informative conclusions for argumentative texts. In *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, pages 3482–3493, 2021.

[377] Shahbaz Syed, Dominik Schwabe, Khalid Al Khatib, and Martin Potthast. Indicative summarization of long discussions. In Houda Bouamor, Juan Pino, and Kalika Bali, editors, *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 2752–2788, Singapore, December 2023. Association for Computational Linguistics. doi: 10.18653/v1/2023.emnlp-main.166.

[378] Zeerak Talat, Hagen Blix, Josef Valvoda, Maya Indira Ganesh, Ryan Cotterell, and Adina Williams. On the Machine Learning of Ethical Judgments from Natural Language. In *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, NAACL '22, pages 769–779, Seattle, USA, 2022. doi: 10.18653/v1/2022.naacl-main.56.

[379] Alex Tamkin, Miles Brundage, Jack Clark, and Deep Ganguli. Understanding the capabilities, limitations, and societal impact of large language models. *arXiv preprint arXiv:2102.02503*, 2021.

[380] Dapeng Tao, Jun Cheng, Zhengtao Yu, Kun Yue, and Lizhen Wang. Domain-weighted majority voting for crowdsourcing. *IEEE transactions on neural networks and learning systems*, 30(1):163–174, 2018.

[381] Milagro Teruel, Cristian Cardellino, Fernando Cardellino, Laura Alonso Alemany, and Serena Villata. Increasing argument annotation reproducibility by using inter-annotator agreement to improve guidelines. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, pages 4061–4064, Miyazaki, Japan, May 2018. European Language Resources Association (ELRA).

[382] Terne Sasha Thorn Jakobsen, Maria Barrett, and Anders Søgaard. Spurious correlations in cross-topic argument mining. In *Proceedings of *SEM 2021: The Tenth Joint Conference on Lexical and Computational Semantics*, pages 263–277, Online, August 2021. Association for Computational Linguistics. doi: 10.18653/v1/2021.starsem-1.25.

[383] Ilaria Tiddi, Victor de Boer, Stefan Schlobach, and André Meyer-Vitali. Knowledge engineering for hybrid intelligence. In *Proceedings of the 12th Knowledge Capture Conference 2023*, K-CAP '23, page 75–82, Pensacola, FL, USA„ 2023. Association for Computing Machinery. ISBN 9798400701412. doi: 10.1145/3587259.3627541.

[384] Nenad Tomašev, Julien Cornebise, Frank Hutter, Shakir Mohamed, Angela Picciariello, Bec Connelly, Danielle CM Belgrave, Daphne Ezer, Fanny Cachat van der Haert, Frank Mugisha, et al. Ai for social good: unlocking the opportunity for positive impact. *Nature Communications*, 11(1):2468, 2020. doi: 10.1038/s41467-020-15871-z.

[385] Petter Törnberg. Chatgpt-4 outperforms experts and crowd workers in annotating political twitter messages with zero-shot learning. *arXiv preprint arXiv:2304.06588*, 2023.

[386] Stephen E Toulmin. *The uses of argument*. Cambridge university press, 2003. ISBN 9780521534833.

[387] Hugo Touvron, Thibaut Lavril, Gautier Izacard, Xavier Martinet, Marie-Anne Lachaux, Timothée Lacroix, Baptiste Rozière, Naman Goyal, Eric Hambro, Faisal Azhar, et al. Llama: Open and efficient foundation language models. *arXiv preprint arXiv:2302.13971*, 0(0), 2023.

[388] Jackson Trager, Alireza S Ziabari, Aida Mostafazadeh Davani, Preni Golazazian, Farzan Karimi-Malekabadi, Ali Omrani, Zhihe Li, Brendan Kennedy, Nils Karl Reimer, Melissa Reyes, et al. The moral foundations reddit corpus. *arXiv preprint arXiv:2208.05545*, 2022.

[389] Nhat Tran and Diane Litman. Multi-task learning in argument mining for persuasive online discussions. In *Proceedings of the 8th Workshop on Argument Mining*, pages 148–153, 2021.

[390] Matthias Trénel. Facilitation and inclusive deliberation. *Online deliberation: Design, research, and practice*, pages 253–257, 2009.

[391] Roman Trötschel, Joachim Hüffmeier, David D Loschelder, Katja Schwartz, and Peter M Gollwitzer. Perspective taking as a means to overcome motivational barriers in negotiations: When putting oneself into the opponent's shoes helps to walk toward agreements. *Journal of personality and social psychology*, 101(4):771, 2011.

[392] Petter Törnberg. How digital media drive affective polarization through partisan sorting. *Proceedings of the National Academy of Sciences*, 119(42):e2207159119, 2022. doi: 10.1073/pnas.2207159119.

[393] Alexandra N Uma, Tommaso Fornaciari, Dirk Hovy, Silviu Paun, Barbara Plank, and Massimo Poesio. Learning from disagreement: A survey. *Journal of Artificial Intelligence Research*, 72:1385–1470, jan 2021. ISSN 1076-9757. doi: 10.1613/jair.1.12752.

[394] Steven Umbrello and Ibo Van de Poel. Mapping value sensitive design onto ai for social good principles. *AI and Ethics*, 1(3):283–296, 2021.

[395] Raphael Vallat. Pingouin: statistics in python. *Journal of Open Source Software*, 3(31): 1026, 2018. doi: 10.21105/joss.01026.

[396] Gido M van de Ven, Tinne Tuytelaars, and Andreas S Tolias. Three types of incremental learning. *Nature Machine Intelligence*, 4(12):1185–1197, 2022. doi: 10.1038/s42256-022-00568-3.

[397] Michiel van der Meer, Enrico Liscio, Catholijn M. Jonker, Aske Plaat, Piek Vossen, and Pradeep K. Murukannaiah. HyEnA: A hybrid method for extracting arguments from opinions. In *HHAI2022: Augmenting Human Intellect*, pages 17–31. IOS Press, 2022.

[398] Michiel van der Meer, Enrico Liscio, Catholijn M. Jonker, Aske Plaat, Piek Vossen, and Pradeep K. Murukannaiah. Hyena: A hybrid method for extracting arguments from opinions. In *Proceedings of the first International Conference on Hybrid Human-Artificial Intelligence (HHAI 2022)*, pages 17–31, Amsterdam, the Netherlands, 2022. IOS Press.

[399] Michiel van der Meer, Myrthe Reuver, Urja Khurana, Lea Krause, and Selene Baez Santamaria. Will it blend? Mixing training paradigms & prompting for argument quality prediction. In *Proceedings of the 9th Workshop on Argument Mining*, pages 95–103, Online and in Gyeongju, Republic of Korea, 10 2022. International Conference on Computational Linguistics.

[400] Michiel van der Meer, Piek Vossen, Catholijn M. Jonker, and Pradeep K. Murukannaiah. Do differences in values influence disagreements in online discussions? In Houda Bouamor, Juan Pino, and Kalika Bali, editors, *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 15986–16008, Singapore, December 2023. Association for Computational Linguistics. doi: 10.18653/v1/2023.emnlp-main.992.

[401] Michiel van der Meer, Piek Vossen, Catholijn M. Jonker, and Pradeep K. Murukannaiah. Do differences in values influence disagreements in online discussions? Supplementary material, Dec 2023.

[402] Michiel Van Der Meer, Neele Falk, Pradeep K. Murukannaiah, and Enrico Liscio. Annotator-centric active learning for subjective NLP tasks. In Yaser Al-Onaizan, Mohit Bansal, and Yun-Nung Chen, editors, *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pages 18537–18555, Miami, Florida, USA, November 2024. Association for Computational Linguistics.

[403] Michiel van der Meer, Enrico Liscio, Catholijn M. Jonker, Aske Plaat, Piek Vossen, and Pradeep K. Murukannaiah. A hybrid intelligence method for argument mining. *Journal of Artificial Intelligence Research*, 80:1187–1222, 2024. doi: https://doi.org/10.1613/jair.1.15135.

[404] Michiel van der Meer, Enrico Liscio, Catholijn M. Jonker, Aske Plaat, Piek Vossen, and Pradeep K. Murukannaiah. A hybrid intelligence method for argument mining: Supplementary material, 2024.

[405] Michiel van der Meer, Piek Vossen, Catholijn M. Jonker, and Pradeep K. Murukannaiah. An empirical analysis of diversity in argument summarization. In Yvette Graham and Matthew Purver, editors, *(To appear) Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics*, pages 2028–2045, St. Julian's, Malta, March 2024. Association for Computational Linguistics.

[406] Max van Duijn, Bram van Dijk, Tom Kouwenhoven, Werner de Valk, Marco Spruit, and Peter van der Putten. Theory of mind in large language models: Examining performance of 11 state-of-the-art models vs. children aged 7-10 on advanced tests. In Jing Jiang, David Reitter, and Shumin Deng, editors, *Proceedings of the 27th Conference on Computational Natural Language Learning (CoNLL)*, pages 389–402, Singapore, December 2023. Association for Computational Linguistics. doi: 10.18653/v1/2023.conll-1.25.

[407] Frans H. van Eemeren, Rob Grootendorst, and Tjark Kruiger. *Handbook of Argumentation Theory*. De Gruyter Mouton, Berlin, Boston, 1987. ISBN 9783110846096. doi: doi:10.1515/9783110846096.

[408] Frans H Van Eemeren, Frans H van Eemeren, Sally Jackson, and Scott Jacobs. Argumentation. *Reasonableness and effectiveness in argumentative discourse: Fifty contributions to the development of Pragma-dialectics*, pages 3–25, 2015.

[409] Daan Van Knippenberg, Carsten KW De Dreu, and Astrid C Homan. Work group diversity and group performance: an integrative model and research agenda. *Journal of applied psychology*, 89(6):1008, 2004.

[410] Don van Ravenzwaaij and Eric-Jan Wagenmakers. Advantages masquerading as "issues"in bayesian hypothesis testing: A commentary on tendeiro and kiers (2019). *Psychological Methods*, 27(3):451–465, June 2022. doi: 10.1037/met0000415.

[411] Chantal van Son, Tommaso Caselli, Antske Fokkens, Isa Maks, Roser Morante, Lora Aroyo, and Piek Vossen. GRaSP: A multilayered annotation scheme for perspectives. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*, pages 1177–1184, Portorož, Slovenia, May 2016. European Language Resources Association (ELRA).

[412] Chantal van Son, Tommaso Caselli, Antske Fokkens, Isa Maks, Roser Morante, Lora Aroyo, and Piek Vossen. GRaSP: A multilayered annotation scheme for perspectives. In Nicoletta Calzolari, Khalid Choukri, Thierry Declerck, Sara Goggi, Marko Grobelnik, Bente Maegaard, Joseph Mariani, Helene Mazo, Asuncion Moreno, Jan Odijk,

and Stelios Piperidis, editors, *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*, pages 1177–1184, Portorož, Slovenia, May 2016. European Language Resources Association (ELRA).

[413] Emma Van Zoelen, Tina Mioch, Mani Tajaddini, Christian Fleiner, Stefani Tsaneva, Pietro Camin, Thiago S. Gouvêa, Kim Baraka, Maaike H. T. De Boer, and Mark A. Neerincx. Developing team design patterns for hybrid intelligence systems. In *Proceedings of the second International Conference on Hybrid Human-Artificial Intelligence (HHAI 2023)*, pages 3–16. IOS Press, 2023. doi: 10.3233/faia230071.

[414] Eva Maria Vecchi, Neele Falk, Iman Jundi, and Gabriella Lapesa. Towards argument mining for social good: A survey. In Chengqing Zong, Fei Xia, Wenjie Li, and Roberto Navigli, editors, *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 1338–1352, Online, August 2021. Association for Computational Linguistics. doi: 10.18653/v1/2021.acl-long.107.

[415] Jacky Visser, Barbara Konat, Rory Duthie, Marcin Koszowy, Katarzyna Budzynska, and Chris Reed. Argumentation in the 2016 us presidential elections: annotated corpora of television debates and social media reaction. *Language Resources and Evaluation*, 54(1):123–154, 2020.

[416] Henning Wachsmuth, Nona Naderi, Yufang Hou, Yonatan Bilu, Vinodkumar Prabhakaran, Tim Alberdingk Thijm, Graeme Hirst, and Benno Stein. Computational argumentation quality assessment in natural language. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 1, Long Papers*, pages 176–187, Valencia, Spain, April 2017. Association for Computational Linguistics.

[417] Henning Wachsmuth, Shahbaz Syed, and Benno Stein. Retrieval of the best counterargument without prior topic knowledge. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 241–251, 2018.

[418] Douglas Walton, Christopher Reed, and Fabrizio Macagno. *Argumentation schemes*. Cambridge University Press, 2008.

[419] Jindong Wang, Xixu Hu, Wenxin Hou, Hao Chen, Runkai Zheng, Yidong Wang, Linyi Yang, Haojun Huang, Wei Ye, Xiubo Geng, et al. On the robustness of chatgpt: An adversarial and out-of-distribution perspective. *arXiv preprint arXiv:2302.12095*, 2023: 1–19, 2 2023.

[420] Lu Wang and Wang Ling. Neural network-based abstract generation for opinions and arguments. In Kevin Knight, Ani Nenkova, and Owen Rambow, editors, *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 47–57, San Diego, California, June 2016. Association for Computational Linguistics. doi: 10.18653/v1/N16-1007.

[421] Xinpeng Wang and Barbara Plank. Actor: Active learning with annotator-specific classification heads to embrace human label variation. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 2046–2052, 2023.

[422] Yizhong Wang, Hamish Ivison, Pradeep Dasigi, Jack Hessel, Tushar Khot, Khyathi Chandu, David Wadden, Kelsey MacMillan, Noah A Smith, Iz Beltagy, et al. How far can camels go? exploring the state of instruction tuning on open resources. *Advances in Neural Information Processing Systems*, 36:74764–74786, 2024.

[423] Zijie J. Wang, Dongjin Choi, Shenyu Xu, and Diyi Yang. Putting humans in the natural language processing loop: A survey. In Su Lin Blodgett, Michael Madaio, Brendan O'Connor, Hanna Wallach, and Qian Yang, editors, *Proceedings of the First Workshop on Bridging Human–Computer Interaction and Natural Language Processing*, pages 47–52, Online, April 2021. Association for Computational Linguistics.

[424] Mayur Wankhade, Annavarapu Chandra Sekhara Rao, and Chaitanya Kulkarni. A survey on sentiment analysis methods, applications, and challenges. *Artificial Intelligence Review*, 55(7):5731–5780, 2022. doi: 10.1007/s10462-022-10144-1.

[425] Cedric Waterschoot, Ernst van den Hemel, and Antal van den Bosch. Detecting minority arguments for mutual understanding: A moderation tool for the online climate change debate. In Nicoletta Calzolari, Chu-Ren Huang, Hansaem Kim, James Pustejovsky, Leo Wanner, Key-Sun Choi, Pum-Mo Ryu, Hsin-Hsi Chen, Lucia Donatelli, Heng Ji, Sadao Kurohashi, Patrizia Paggio, Nianwen Xue, Seokhwan Kim, Younggyun Hahm, Zhong He, Tony Kyungil Lee, Enrico Santus, Francis Bond, and Seung-Hoon Na, editors, *Proceedings of the 29th International Conference on Computational Linguistics*, pages 6715–6725, Gyeongju, Republic of Korea, October 2022. International Committee on Computational Linguistics.

[426] Leon Weber-Genzel, Siyao Peng, Marie-Catherine de Marneffe, and Barbara Plank. Varierr nli: Separating annotation error from human label variation, 2024.

[427] Jason Wei, Yi Tay, Rishi Bommasani, Colin Raffel, Barret Zoph, Sebastian Borgeaud, Dani Yogatama, Maarten Bosma, Denny Zhou, Donald Metzler, et al. Emergent abilities of large language models. *Transactions on Machine Learning Research*, 2022. ISSN 2835-8856. Survey Certification.

[428] Maxwell A Weinzierl and Sanda M Harabagiu. From hesitancy framings to vaccine hesitancy profiles: A journey of stance, ontological commitments and moral foundations. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 16, pages 1087–1097, 2022.

[429] Galen Weld, Amy X. Zhang, and Tim Althoff. What makes online communities 'better'? measuring values, consensus, and conflict across thousands of subreddits. *Proceedings of the International AAAI Conference on Web and Social Media*, 16(1):1121–1132, May 2022. doi: 10.1609/icwsm.v16i1.19363.

[430] Orion Weller, Kevin Seppi, and Matt Gardner. When to use multi-task learning vs intermediate fine-tuning for pre-trained encoder transfer learning. In *Proceedings of*

the 60th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers), pages 272–282, 2022.

[431] Jevin D West and Carl T Bergstrom. Misinformation in and about science. *Proceedings of the National Academy of Sciences*, 118(15):e1912444117, 2021.

[432] Janyce Wiebe, Theresa Wilson, and Claire Cardie. Annotating expressions of opinions and emotions in language. *Language resources and evaluation*, 39(2-3):165–210, 2005. doi: 10.1007/s10579-005-7880-9.

[433] Adina Williams, Nikita Nangia, and Samuel Bowman. A broad-coverage challenge corpus for sentence understanding through inference. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*, pages 1112–1122. Association for Computational Linguistics, 2018.

[434] Michael Wojatzki, Torsten Zesch, Saif Mohammad, and Svetlana Kiritchenko. Agree or disagree: Predicting judgments on nuanced assertions. In *Proceedings of the Seventh Joint Conference on Lexical and Computational Semantics*, pages 214–224, 2018.

[435] Thomas Wolf, Lysandre Debut, Victor Sanh, Julien Chaumond, Clement Delangue, Anthony Moi, Pierric Cistac, Tim Rault, Rémi Louf, Morgan Funtowicz, et al. Huggingface's transformers: State-of-the-art natural language processing. *arXiv preprint arXiv:1910.03771*, 2019.

[436] Thomas Wolf, Lysandre Debut, Victor Sanh, Julien Chaumond, Clement Delangue, Anthony Moi, Pierric Cistac, Tim Rault, Rémi Louf, Morgan Funtowicz, et al. Transformers: State-of-the-art natural language processing. In *Proceedings of the 2020 conference on empirical methods in natural language processing: system demonstrations*, pages 38–45, Online, October 2020. Association for Computational Linguistics.

[437] Scott Wright and John Street. Democracy, deliberation and design: the case of online discussion forums. *New media & society*, 9(5):849–869, 2007. doi: 10.1177/1461444807081230.

[438] Chuhan Wu, Fangzhao Wu, Tao Qi, and Yongfeng Huang. SentiRec: Sentiment diversity-aware neural news recommendation. In Kam-Fai Wong, Kevin Knight, and Hua Wu, editors, *Proceedings of the 1st Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 10th International Joint Conference on Natural Language Processing*, pages 44–53, Suzhou, China, December 2020. Association for Computational Linguistics.

[439] Adam Wyner, Jodi Schneider, Katie Atkinson, and Trevor Bench-Capon. Semi-automated argumentative analysis of online product reviews. In *Computational Models of Argument*, pages 43–50. IOS Press, 2012. doi: 10.3233/978-1-61499-111-3-43.

[440] Yan Xia, Haiyi Zhu, Tun Lu, Peng Zhang, and Ning Gu. Exploring antecedents and consequences of toxicity in online discussions: A case study on reddit. *Proceedings of the ACM on Human-computer Interaction*, 4(CSCW2):1–23, oct 2020. doi: 10.1145/3415179.

[441] Fei Xiong and Yun Liu. Opinion formation on social media: an empirical approach. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 24(1):013130, 03 2014. ISSN 1054-1500. doi: 10.1063/1.4866011.

[442] Jun Xu, Xiangnan He, and Hang Li. Deep learning for matching in search and recommendation. *Foundations and Trends® in Information Retrieval*, 14(2–3):102–288, 2020. ISSN 1554-0669. doi: 10.1561/1500000076.

[443] Qiongkai Xu, Christian Walder, and Chenchen Xu. Humanly certifying superhuman classifiers. *arXiv preprint arXiv:2109.07867*, 2021.

[444] Naomi Yamashita and Toru Ishida. Effects of machine translation on collaborative work. In *Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work*, pages 515–524, 2006.

[445] Weizhe Yuan, Graham Neubig, and Pengfei Liu. Bartscore: Evaluating generated text as text generation. *Advances in Neural Information Processing Systems*, 34:27263–27277, 2021.

[446] Lihi Zelnik-Manor and Pietro Perona. Self-tuning spectral clustering. In *Proceedings of the 17th International Conference on Neural Information Processing Systems*, volume 17 of *NIPS'04*, page 1601–1608, Cambridge, MA, USA, 2004. MIT Press.

[447] Leihan Zhang, Jichang Zhao, and Ke Xu. Who creates trends in online social media: The crowd or opinion leaders? *Journal of Computer-Mediated Communication*, 21(1): 1–16, 2016.

[448] Tianyi Zhang, Varsha Kishore, Felix Wu, Kilian Q. Weinberger, and Yoav Artzi. Bertscore: Evaluating text generation with bert. In *International Conference on Learning Representations*, 2020.

[449] Tianyi Zhang, Varsha Kishore, Felix Wu, Kilian Q. Weinberger, and Yoav Artzi. Bertscore: Evaluating text generation with bert. In *International Conference on Learning Representations*, 2020.

[450] Ye Zhang, Matthew Lease, and Byron C. Wallace. Active Discriminative Text Representation Learning. In *Proceedings of the 31st AAAI Conference on Artificial Intelligence*, pages 3386–3392, San Francisco, California, USA, 2017.

[451] Zhisong Zhang, Emma Strubell, and Eduard Hovy. A Survey of Active Learning for Natural Language Processing. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, EMNLP '22, pages 6166–6190. ACL, 2022.

[452] Yuekai Zhao, Haoran Zhang, Shuchang Zhou, and Zhihua Zhang. Active Learning Approaches to Enhancing Neural Machine Translation. In *Findings of the Association for Computational Linguistics*, EMNLP 2020, pages 1796–1806, Online, 2020. ACL. doi: 10.18653/v1/2020.findings-emnlp.162.

[453] Ming Zhong, Pengfei Liu, Yiran Chen, Danqing Wang, Xipeng Qiu, and Xuanjing Huang. Extractive summarization as text matching. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 6197–6208, Online, July 2020. Association for Computational Linguistics. doi: 10.18653/v1/2020.acl-main.552.

[454] Caleb Ziems, William Held, Omar Shaikh, Jiaao Chen, Zhehao Zhang, and Diyi Yang. Can large language models transform computational social science? *Computational Linguistics*, pages 1–55, 2024.

[455] Elena Ziliotti, Patricia Reyes Benavides, Arthur Gwagwa, and Matthew Dennis. *Social Media and Democracy*, chapter 2, pages 33–52. Open Book Publishers, 2023. doi: 10.11647/obp.0366.02.

Acknowledgments

Some of the people listed in this acknowledgments section will reside in more than one subgroup, in which case please know that I might have grouped you arbitrarily. Figuring out interconnections is left as an exercise for the reader.

Pradeep, you were my closest supervisor, the one leading me through it all. Thank you for your guidance during the years in my PhD and beyond. You taught me how to do so many things the scientific way: writing, collaborating, discussing, and organizing, to name a few. I consider you a great example of how to be a great person, professionally and personally.

Catholijn, thank you for always having my back and being kind in cases where I could use some steering. You inspire me to never let go of my enthusiasm and to keep talking and honing my research. I will always remember your advice to *frappez toujours*.

Piek, in you I found a strong sparring partner. Your office was always open, and inviting for a healthy discussion. You went along with my endless fidgeting about what to do next and let me find my path. I am very grateful you facilitated my second home at the VU.

Aske, thank you for welcoming me to Leiden University. You were (physically!) there when I started during COVID and provided advice when I needed it. You showed me the value of pursuing what I want to do over all other things.

My closest colleagues had a significant hand in my PhD. Enrico, if there is such a thing as an academic brother I am sure it is you. You were there from the first interview to my final paper, and I am grateful you are the best scholar-in-crime I could have imagined. I'd gladly share a five-star hotel room with you again. Bram, you were not only my closest colleague in Leiden but also a friend outside of it. Your enjoyment of the good things in life, as well as your endless capacity for banter, makes me enjoy every situation we find ourselves in, and you lifted the quality of the many events we attended together. To complete the group of culinarians, Mani, thank you for starting our tradition of high-quality dinner evenings after going to Berlin together. I hope we find our way to the same table again soon.

To my colleagues from CLTL, I will dearly miss our (occasionally lengthy) coffee breaks. Selene, thank you for your exceptional kindness, for being up for exploring new foods on our adventures, and for never failing to lift my spirit. Lea, your warm heart and critical mind make you one of the kindest and coolest people I know. Urja, we go way back to the BSc KI at UvA, I'm glad our paths have crossed for so long. Myrthe, you are my opinion-mining buddy and your drive for academic diligence motivates me to be a better scientist. Throughout my PhD, I managed to travel to different places, be it for conferences or vacation afterward. Jaap, Stella, Jonathan, and Baran, the times together were among my favorites. Tom, we explored Singapore together before ending up as office mates in Leiden, the times in the new building were short but sweet. To the HI IGLU team: Tae, Putra, and Kata, thank you for besting the challenge together and finding each other again in different parts of the world. To my other coauthors, thank you for teaching me much about (interdisciplinary) research. Anatol and Ruth, thanks for calmly explaining the basics of political science, Luciano, for kickstarting my PhD alongside Enrico, and Neele for pushing the last project in my PhD to successful completion.

The past years would have been very different if it were not for the groups I found myself in. At LIACS, my thanks extend to Thomas, Matthias, Andreas (I hope your plants live long lives), Zhao, Mike, Andrius, Alan, Álvaro, Felix, Koen, Annie, Joost, and all other present and past members of the RLG group, as well as Suzan, Amin, Arian, Gijs, Zhaochun, Juan, and the other members of the TMR group, and finally Bernard, Ramira, Max, Tessa, Bram, Yuchen, and the other members of the CIL group. At the VU Amsterdam, my thanks go to the CLTL group for hosting me: Leon, Annika, Wende, Stefan, Ellie, Pia, Antske, Levi, Angel, Ilia, Lisa, and all other present and past members of the CLTL group at the VU. Lastly, at the TU Delft: Ruben, Pei-Yu, Masha, Morita, Emma, Nele, Amir, Carolina, Zuzanna, Deborah, Myrthe, Catha, Frans, Siddharth, and a special mention for the organizational miracle that is Anita, as well as any other present or past member in the II group.

None of my work would have come to fruition without those in the Hybrid Intelligence consortium. Thank you for powering through the difficult COVID times with great success. Bart, and Cor, thank you for instilling the HI values and academic lifeblood in me. Tiffany, I greatly enjoyed co-organizing the HI PhD reading group with you. Thanks for the wonderful times to the early HI PhDs: Johanna (I hope you still have the light you stole), Annet, Ludi, Giacomo, Wijnand, Mark, Íñigo, Niklas, Maria, Sharvaree, Nicole, Aishwarya, Merle, Loan, Emre, J.D., Delaram, Anna, Jonne, Feline, as well as the later cohorts of HI PhDs. I would also like to express my gratitude to the senior members of the HI consortium, including Rineke, Ilaria, Victor, Davide, Davide, Jasper, Henry, Erman, Annette, Dan, Kim, Herke, Silja, Pinar, Birna, Shenghui, Roel, Bernd, and all other present and past HI members. Finally, I'd like to thank the management team, including Frank, Wendy, and Stefan for steering the HI meetings in the right direction and succeeding in creating a tight-knit group that felt like a home.

Ik had mijn PhD niet kunnen afmaken zonder mijn vrienden en familie. Jullie zorgen er voor dat ik mijn werk even kan vergeten en ruimte krijg voor andere dingen. Aan de Kumdo instructor group: Bas, Dominic, en Kelly (en Devil), jullie zijn meer dan vrienden en ik hoop dat wij samen trainen tot we oud zijn (en dat zijn we nog niet). Kevin (honk!), Sterre, Maya, Ray, Lotte, Oumaima, Max, Nele, Quintin en Quinten, Evgeniia, Zena, Tibor, Yuri, Meg, en alle anderen bij Dojo Den Drijver en Gungsul Academy, 해동! Jonathan, bedankt voor het altijd beschikbaar zijn om te praten over alle dingen in het leven, jouw enthousiasme over werkelijk elk onderwerp is ontzettend aanstekelijk. Arianne, Bart, Thomas, Roan, en de rest van mijn goeie vrienden, ik ben dankbaar voor jullie warmte en gesprekken over de jaren. Sabina, jij hebt mij van het begin tot het eind altijd van aanmoediging voorzien, zowel in tijden van deadlinestress, als in de vakanties samen erna. Door jouw rol als een onuitputbaar klankbord is dit proefschrift ook een beetje van jou.

Als laatste gaat mijn dank naar degenen die mij al mijn hele leven samen met mij zijn. Arwen, Sanne, Theo en Petra, jullie staan aan de basis van mijn nieuwsgierigheid om de wereld om me heen beter te begrijpen. Mama, en papa en Jolanda, bedankt voor jullie onvoorwaardelijke steun. Jullie hebben me gevormd tot de persoon die ik ben. Annemein, Felix en Aster, wij zijn met z'n vieren een eenheid, en dat was niet anders tijdens mijn PhD. Ik kan me niet voorstellen hoe ik dit allemaal zou hebben gedaan zonder zo'n geweldig liefdevol zooitje zusjes en broertje om me heen.

Curriculum Vitæ

Michiel Theo van der Meer was born on October 29, 1995, in Groningen, the Netherlands. He graduated with a gymnasium degree from the Jan van Egmond Lyceum in Purmerend, Noord-Holland, in 2014. Michiel then pursued both his Bachelor of Science and Master of Science in Artificial Intelligence at the University of Amsterdam, completing these degrees from 2014 to 2017 and from 2017 to 2020, respectively.

During his academic career, Michiel was actively involved in various roles. He served as a teaching assistant, was part of the education committee in his study association, and was a dedicated team member of the Dutch Nao Team from 2015 to 2019. In this team, he held several positions, including conducting weekly meetings, facilitating group decisions, and developing soccer-playing robots. Michiel traveled with his team to notable events such as Techfest 2015 in Mumbai, India, and RoboCup competitions from 2016 to 2019 in Leipzig, Germany; Tehran, Iran; Nagoya, Japan; and Montréal, Canada. He completed his master's degree cum laude in 2020, with a thesis focused on explainability in reinforcement learning.

Alongside his academic pursuits, Michiel has garnered diverse professional experience in the AI field. From 2018 to 2020, he worked as an AI developer and trainer at Millennials.ai, where he developed machine learning approaches for healthcare services and delivered lectures on AI and NLP basics to beginner audiences at various companies, governmental departments, and public institutions. In 2020, he served as a Data Engineer at InBiome in Amsterdam before embarking on his PhD studies at Leiden University, which he completed from 2020 to 2024. During his PhD studies, he took courses in transferable skills such as research methods for computer science, project and data management, and scientific conduct, among others. Starting from August 2024, Michiel is a postdoctoral researcher at the Idiap Research Institute in Martigny, Switzerland.

List of Publications

2024

1. **Michiel van der Meer**, Neele Falk, Pradeep K. Murukannaiah, Enrico Liscio. 2024. Annotator-Centric Active Learning for Subjective NLP Tasks. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pages 18537–18555, Miami, Florida, USA. Association for Computational Linguistics.

2. **Michiel van der Meer**, Enrico Liscio, Catholijn M. Jonker, Aske Plaat, Piek Vossen, Pradeep K. Murukannaiah. 2024. A Hybrid Intelligence Method for Argument Mining. In *Journal of Artificial Intelligence Research* 80, pages 1187–1222.

3. **Michiel van der Meer**. Facilitating Online Opinion Diversity through Hybrid NLP Approaches (Thesis proposal). In *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 4: Student Research Workshop)*, pages 272–284, Mexico City, Mexico. Association for Computational Linguistics.

4. **Michiel van der Meer**, Piek Vossen, Catholijn M. Jonker, and Pradeep K. Murukannaiah. Value-Sensitive Disagreement Analysis for Online Deliberation. In *HHAI 2024: Hybrid Human AI Systems for the Social Good (Extended Abstracts)*, pages 481–484, Malmö, Sweden. IOS Press.

5. **Michiel van der Meer**, Piek Vossen, Catholijn M. Jonker, and Pradeep K. Murukannaiah. An Empirical Analysis of Diversity in Argument Summarization. In *Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 2028–2045, St. Julian's, Malta. Association for Computational Linguistics.

2023

1. **Michiel van der Meer**, Piek Vossen, Catholijn M. Jonker, and Pradeep K. Murukannaiah. 2023. Do Differences in Values Influence Disagreements in Online Discussions? In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 15986–16008, Singapore. Association for Computational Linguistics.

2. Lea Krause, Selene Báez Santamaría, **Michiel van der Meer**, and Urja Khurana. 2023. Leveraging Few-Shot Data Augmentation and Waterfall Prompting for Response Generation. In *Proceedings of The Eleventh Dialog System Technology Challenge*, pages 193–205, Prague, Czech Republic. Association for Computational Linguistics.

3. Thomas M. Moerland, Matthias Müller-Brockhausen, Zhao Yang, Andrius Bernatavicius, Koen Ponse, Tom Kouwenhoven, Andreas Sauter, **Michiel van der Meer**, Bram Renting, and Aske Plaat. EduGym: An Environment Suite for Reinforcement Learning Education. *arXiv preprint.*

2022

1. Ruth Shortall, Anatol Itten, **Michiel van der Meer**, Pradeep K. Murukannaiah, Catholijn M. Jonker. 2022. Reason Against the Machine: Future Directions for Mass Online Deliberation. In *Frontiers in Political Science*.

2. **Michiel van der Meer**, Myrthe Reuver, Urja Khurana, Lea Krause, and Selene Báez Santamaría. 2022. Will It Blend? Mixing Training Paradigms & Prompting for Argument Quality Prediction. In *Proceedings of the 9th Workshop on Argument Mining*, pages 95–103, Online and in Gyeongju, Republic of Korea. International Conference on Computational Linguistics.

3. Enrico Liscio, **Michiel van der Meer**, Luciano C. Siebert, Catholijn M. Jonker, Pradeep K. Murukannaiah. 2022. What Values Should an Agent Align with? An Empirical Comparison of General and Context-Specific Values. In *Autonomous Agents and Multi-Agent Systems* 36, 23.

4. **Michiel van der Meer**, Enrico Liscio, Catholijn M. Jonker, Aske Plaat, Piek Vossen, Pradeep K. Murukannaiah. 2022. HyEnA: A Hybrid Method for Extracting Arguments from Opinions. In *HHAI2022: Augmenting Human Intellect*, pages 17–31, Amsterdam, the Netherlands. IOS Press. [**Best paper award**]

2021

1. Enrico Liscio, **Michiel van der Meer**, Luciano C. Siebert, Catholijn M. Jonker, Niek Mouter, Pradeep K. Murukannaiah. 2021. Axies: Identifying and Evaluating Context-Specific Values. In *Proceedings of the 20th International Conference on Autonomous Agents and Multiagent Systems*, pages 799–808. Online. IFAAMAS.

2. Enrico Liscio, **Michiel van der Meer**, Catholijn M. Jonker, Pradeep K. Murukannaiah. 2021. A Collaborative Platform for Identifying Context-Specific Values: Demo Track. In *Proceedings of the 20th International Conference on Autonomous Agents and Multiagent Systems*, pages 1773–1775, Online. IFAAMAS.

Included in this thesis.

SIKS Dissertations

2016
01 Syed Saiden Abbas (RUN), Recognition of Shapes by Humans and Machines
02 Michiel Christiaan Meulendijk (UU), Optimizing medication reviews through decision support: prescribing a better pill to swallow
03 Maya Sappelli (RUN), Knowledge Work in Context: User Centered Knowledge Worker Support
04 Laurens Rietveld (VUA), Publishing and Consuming Linked Data
05 Evgeny Sherkhonov (UvA), Expanded Acyclic Queries: Containment and an Application in Explaining Missing Answers
06 Michel Wilson (TUD), Robust scheduling in an uncertain environment
07 Jeroen de Man (VUA), Measuring and modeling negative emotions for virtual training
08 Matje van de Camp (TiU), A Link to the Past: Constructing Historical Social Networks from Unstructured Data
09 Archana Nottamkandath (VUA), Trusting Crowdsourced Information on Cultural Artefacts
10 George Karafotias (VUA), Parameter Control for Evolutionary Algorithms
11 Anne Schuth (UvA), Search Engines that Learn from Their Users
12 Max Knobbout (UU), Logics for Modelling and Verifying Normative Multi-Agent Systems
13 Nana Baah Gyan (VUA), The Web, Speech Technologies and Rural Development in West Africa - An ICT4D Approach
14 Ravi Khadka (UU), Revisiting Legacy Software System Modernization
15 Steffen Michels (RUN), Hybrid Probabilistic Logics - Theoretical Aspects, Algorithms and Experiments
16 Guangliang Li (UvA), Socially Intelligent Autonomous Agents that Learn from Human Reward
17 Berend Weel (VUA), Towards Embodied Evolution of Robot Organisms
18 Albert Meroño Peñuela (VUA), Refining Statistical Data on the Web
19 Julia Efremova (TU/e), Mining Social Structures from Genealogical Data
20 Daan Odijk (UvA), Context & Semantics in News & Web Search
21 Alejandro Moreno Célleri (UT), From Traditional to Interactive Playspaces: Automatic Analysis of Player Behavior in the Interactive Tag Playground
22 Grace Lewis (VUA), Software Architecture Strategies for Cyber-Foraging Systems
23 Fei Cai (UvA), Query Auto Completion in Information Retrieval
24 Brend Wanders (UT), Repurposing and Probabilistic Integration of Data; An Iterative and data model independent approach
25 Julia Kiseleva (TU/e), Using Contextual Information to Understand Searching and Browsing Behavior

26 Dilhan Thilakarathne (VUA), In or Out of Control: Exploring Computational Models to Study the Role of Human Awareness and Control in Behavioural Choices, with Applications in Aviation and Energy Management Domains
27 Wen Li (TUD), Understanding Geo-spatial Information on Social Media
28 Mingxin Zhang (TUD), Large-scale Agent-based Social Simulation - A study on epidemic prediction and control
29 Nicolas Höning (TUD), Peak reduction in decentralised electricity systems - Markets and prices for flexible planning
30 Ruud Mattheij (TiU), The Eyes Have It
31 Mohammad Khelghati (UT), Deep web content monitoring
32 Eelco Vriezekolk (UT), Assessing Telecommunication Service Availability Risks for Crisis Organisations
33 Peter Bloem (UvA), Single Sample Statistics, exercises in learning from just one example
34 Dennis Schunselaar (TU/e), Configurable Process Trees: Elicitation, Analysis, and Enactment
35 Zhaochun Ren (UvA), Monitoring Social Media: Summarization, Classification and Recommendation
36 Daphne Karreman (UT), Beyond R2D2: The design of nonverbal interaction behavior optimized for robot-specific morphologies
37 Giovanni Sileno (UvA), Aligning Law and Action - a conceptual and computational inquiry
38 Andrea Minuto (UT), Materials that Matter - Smart Materials meet Art & Interaction Design
39 Merijn Bruijnes (UT), Believable Suspect Agents; Response and Interpersonal Style Selection for an Artificial Suspect
40 Christian Detweiler (TUD), Accounting for Values in Design
41 Thomas King (TUD), Governing Governance: A Formal Framework for Analysing Institutional Design and Enactment Governance
42 Spyros Martzoukos (UvA), Combinatorial and Compositional Aspects of Bilingual Aligned Corpora
43 Saskia Koldijk (RUN), Context-Aware Support for Stress Self-Management: From Theory to Practice
44 Thibault Sellam (UvA), Automatic Assistants for Database Exploration
45 Bram van de Laar (UT), Experiencing Brain-Computer Interface Control
46 Jorge Gallego Perez (UT), Robots to Make you Happy
47 Christina Weber (UL), Real-time foresight - Preparedness for dynamic innovation networks
48 Tanja Buttler (TUD), Collecting Lessons Learned
49 Gleb Polevoy (TUD), Participation and Interaction in Projects. A Game-Theoretic Analysis
50 Yan Wang (TiU), The Bridge of Dreams: Towards a Method for Operational Performance Alignment in IT-enabled Service Supply Chains

2017 01 Jan-Jaap Oerlemans (UL), Investigating Cybercrime

02 Sjoerd Timmer (UU), Designing and Understanding Forensic Bayesian Networks using Argumentation
03 Daniël Harold Telgen (UU), Grid Manufacturing; A Cyber-Physical Approach with Autonomous Products and Reconfigurable Manufacturing Machines
04 Mrunal Gawade (CWI), Multi-core Parallelism in a Column-store
05 Mahdieh Shadi (UvA), Collaboration Behavior
06 Damir Vandic (EUR), Intelligent Information Systems for Web Product Search
07 Roel Bertens (UU), Insight in Information: from Abstract to Anomaly
08 Rob Konijn (VUA), Detecting Interesting Differences:Data Mining in Health Insurance Data using Outlier Detection and Subgroup Discovery
09 Dong Nguyen (UT), Text as Social and Cultural Data: A Computational Perspective on Variation in Text
10 Robby van Delden (UT), (Steering) Interactive Play Behavior
11 Florian Kunneman (RUN), Modelling patterns of time and emotion in Twitter #anticipointment
12 Sander Leemans (TU/e), Robust Process Mining with Guarantees
13 Gijs Huisman (UT), Social Touch Technology - Extending the reach of social touch through haptic technology
14 Shoshannah Tekofsky (TiU), You Are Who You Play You Are: Modelling Player Traits from Video Game Behavior
15 Peter Berck (RUN), Memory-Based Text Correction
16 Aleksandr Chuklin (UvA), Understanding and Modeling Users of Modern Search Engines
17 Daniel Dimov (UL), Crowdsourced Online Dispute Resolution
18 Ridho Reinanda (UvA), Entity Associations for Search
19 Jeroen Vuurens (UT), Proximity of Terms, Texts and Semantic Vectors in Information Retrieval
20 Mohammadbashir Sedighi (TUD), Fostering Engagement in Knowledge Sharing: The Role of Perceived Benefits, Costs and Visibility
21 Jeroen Linssen (UT), Meta Matters in Interactive Storytelling and Serious Gaming (A Play on Worlds)
22 Sara Magliacane (VUA), Logics for causal inference under uncertainty
23 David Graus (UvA), Entities of Interest — Discovery in Digital Traces
24 Chang Wang (TUD), Use of Affordances for Efficient Robot Learning
25 Veruska Zamborlini (VUA), Knowledge Representation for Clinical Guidelines, with applications to Multimorbidity Analysis and Literature Search
26 Merel Jung (UT), Socially intelligent robots that understand and respond to human touch
27 Michiel Joosse (UT), Investigating Positioning and Gaze Behaviors of Social Robots: People's Preferences, Perceptions and Behaviors
28 John Klein (VUA), Architecture Practices for Complex Contexts
29 Adel Alhuraibi (TiU), From IT-BusinessStrategic Alignment to Performance: A Moderated Mediation Model of Social Innovation, and Enterprise Governance of IT"
30 Wilma Latuny (TiU), The Power of Facial Expressions

31 Ben Ruijl (UL), Advances in computational methods for QFT calculations
32 Thaer Samar (RUN), Access to and Retrievability of Content in Web Archives
33 Brigit van Loggem (OU), Towards a Design Rationale for Software Documentation: A Model of Computer-Mediated Activity
34 Maren Scheffel (OU), The Evaluation Framework for Learning Analytics
35 Martine de Vos (VUA), Interpreting natural science spreadsheets
36 Yuanhao Guo (UL), Shape Analysis for Phenotype Characterisation from High-throughput Imaging
37 Alejandro Montes Garcia (TU/e), WiBAF: A Within Browser Adaptation Framework that Enables Control over Privacy
38 Alex Kayal (TUD), Normative Social Applications
39 Sara Ahmadi (RUN), Exploiting properties of the human auditory system and compressive sensing methods to increase noise robustness in ASR
40 Altaf Hussain Abro (VUA), Steer your Mind: Computational Exploration of Human Control in Relation to Emotions, Desires and Social Support For applications in human-aware support systems
41 Adnan Manzoor (VUA), Minding a Healthy Lifestyle: An Exploration of Mental Processes and a Smart Environment to Provide Support for a Healthy Lifestyle
42 Elena Sokolova (RUN), Causal discovery from mixed and missing data with applications on ADHD datasets
43 Maaike de Boer (RUN), Semantic Mapping in Video Retrieval
44 Garm Lucassen (UU), Understanding User Stories - Computational Linguistics in Agile Requirements Engineering
45 Bas Testerink (UU), Decentralized Runtime Norm Enforcement
46 Jan Schneider (OU), Sensor-based Learning Support
47 Jie Yang (TUD), Crowd Knowledge Creation Acceleration
48 Angel Suarez (OU), Collaborative inquiry-based learning

2018 01 Han van der Aa (VUA), Comparing and Aligning Process Representations
02 Felix Mannhardt (TU/e), Multi-perspective Process Mining
03 Steven Bosems (UT), Causal Models For Well-Being: Knowledge Modeling, Model-Driven Development of Context-Aware Applications, and Behavior Prediction
04 Jordan Janeiro (TUD), Flexible Coordination Support for Diagnosis Teams in Data-Centric Engineering Tasks
05 Hugo Huurdeman (UvA), Supporting the Complex Dynamics of the Information Seeking Process
06 Dan Ionita (UT), Model-Driven Information Security Risk Assessment of Socio-Technical Systems
07 Jieting Luo (UU), A formal account of opportunism in multi-agent systems
08 Rick Smetsers (RUN), Advances in Model Learning for Software Systems
09 Xu Xie (TUD), Data Assimilation in Discrete Event Simulations
10 Julienka Mollee (VUA), Moving forward: supporting physical activity behavior change through intelligent technology
11 Mahdi Sargolzaei (UvA), Enabling Framework for Service-oriented Collaborative Networks

12 Xixi Lu (TU/e), Using behavioral context in process mining
13 Seyed Amin Tabatabaei (VUA), Computing a Sustainable Future
14 Bart Joosten (TiU), Detecting Social Signals with Spatiotemporal Gabor Filters
15 Naser Davarzani (UM), Biomarker discovery in heart failure
16 Jaebok Kim (UT), Automatic recognition of engagement and emotion in a group of children
17 Jianpeng Zhang (TU/e), On Graph Sample Clustering
18 Henriette Nakad (UL), De Notaris en Private Rechtspraak
19 Minh Duc Pham (VUA), Emergent relational schemas for RDF
20 Manxia Liu (RUN), Time and Bayesian Networks
21 Aad Slootmaker (OU), EMERGO: a generic platform for authoring and playing scenario-based serious games
22 Eric Fernandes de Mello Araújo (VUA), Contagious: Modeling the Spread of Behaviours, Perceptions and Emotions in Social Networks
23 Kim Schouten (EUR), Semantics-driven Aspect-Based Sentiment Analysis
24 Jered Vroon (UT), Responsive Social Positioning Behaviour for Semi-Autonomous Telepresence Robots
25 Riste Gligorov (VUA), Serious Games in Audio-Visual Collections
26 Roelof Anne Jelle de Vries (UT),Theory-Based and Tailor-Made: Motivational Messages for Behavior Change Technology
27 Maikel Leemans (TU/e), Hierarchical Process Mining for Scalable Software Analysis
28 Christian Willemse (UT), Social Touch Technologies: How they feel and how they make you feel
29 Yu Gu (TiU), Emotion Recognition from Mandarin Speech
30 Wouter Beek (VUA), The "K" in "semantic web" stands for "knowledge": scaling semantics to the web

2019 01 Rob van Eijk (UL),Web privacy measurement in real-time bidding systems. A graph-based approach to RTB system classification
02 Emmanuelle Beauxis Aussalet (CWI, UU), Statistics and Visualizations for Assessing Class Size Uncertainty
03 Eduardo Gonzalez Lopez de Murillas (TU/e), Process Mining on Databases: Extracting Event Data from Real Life Data Sources
04 Ridho Rahmadi (RUN), Finding stable causal structures from clinical data
05 Sebastiaan van Zelst (TU/e), Process Mining with Streaming Data
06 Chris Dijkshoorn (VUA), Nichesourcing for Improving Access to Linked Cultural Heritage Datasets
07 Soude Fazeli (TUD), Recommender Systems in Social Learning Platforms
08 Frits de Nijs (TUD), Resource-constrained Multi-agent Markov Decision Processes
09 Fahimeh Alizadeh Moghaddam (UvA), Self-adaptation for energy efficiency in software systems
10 Qing Chuan Ye (EUR), Multi-objective Optimization Methods for Allocation and Prediction

11 Yue Zhao (TUD), Learning Analytics Technology to Understand Learner Behavioral Engagement in MOOCs
12 Jacqueline Heinerman (VUA), Better Together
13 Guanliang Chen (TUD), MOOC Analytics: Learner Modeling and Content Generation
14 Daniel Davis (TUD), Large-Scale Learning Analytics: Modeling Learner Behavior & Improving Learning Outcomes in Massive Open Online Courses
15 Erwin Walraven (TUD), Planning under Uncertainty in Constrained and Partially Observable Environments
16 Guangming Li (TU/e), Process Mining based on Object-Centric Behavioral Constraint (OCBC) Models
17 Ali Hurriyetoglu (RUN),Extracting actionable information from microtexts
18 Gerard Wagenaar (UU), Artefacts in Agile Team Communication
19 Vincent Koeman (TUD), Tools for Developing Cognitive Agents
20 Chide Groenouwe (UU), Fostering technically augmented human collective intelligence
21 Cong Liu (TU/e), Software Data Analytics: Architectural Model Discovery and Design Pattern Detection
22 Martin van den Berg (VUA),Improving IT Decisions with Enterprise Architecture
23 Qin Liu (TUD), Intelligent Control Systems: Learning, Interpreting, Verification
24 Anca Dumitrache (VUA), Truth in Disagreement - Crowdsourcing Labeled Data for Natural Language Processing
25 Emiel van Miltenburg (VUA), Pragmatic factors in (automatic) image description
26 Prince Singh (UT), An Integration Platform for Synchromodal Transport
27 Alessandra Antonaci (OU), The Gamification Design Process applied to (Massive) Open Online Courses
28 Esther Kuindersma (UL), Cleared for take-off: Game-based learning to prepare airline pilots for critical situations
29 Daniel Formolo (VUA), Using virtual agents for simulation and training of social skills in safety-critical circumstances
30 Vahid Yazdanpanah (UT), Multiagent Industrial Symbiosis Systems
31 Milan Jelisavcic (VUA), Alive and Kicking: Baby Steps in Robotics
32 Chiara Sironi (UM), Monte-Carlo Tree Search for Artificial General Intelligence in Games
33 Anil Yaman (TU/e), Evolution of Biologically Inspired Learning in Artificial Neural Networks
34 Negar Ahmadi (TU/e), EEG Microstate and Functional Brain Network Features for Classification of Epilepsy and PNES
35 Lisa Facey-Shaw (OU), Gamification with digital badges in learning programming
36 Kevin Ackermans (OU), Designing Video-Enhanced Rubrics to Master Complex Skills
37 Jian Fang (TUD), Database Acceleration on FPGAs
38 Akos Kadar (OU), Learning visually grounded and multilingual representations

2020 01 Armon Toubman (UL), Calculated Moves: Generating Air Combat Behaviour
02 Marcos de Paula Bueno (UL), Unraveling Temporal Processes using Probabilistic Graphical Models
03 Mostafa Deghani (UvA), Learning with Imperfect Supervision for Language Understanding
04 Maarten van Gompel (RUN), Context as Linguistic Bridges
05 Yulong Pei (TU/e), On local and global structure mining
06 Preethu Rose Anish (UT), Stimulation Architectural Thinking during Requirements Elicitation - An Approach and Tool Support
07 Wim van der Vegt (OU), Towards a software architecture for reusable game components
08 Ali Mirsoleimani (UL),Structured Parallel Programming for Monte Carlo Tree Search
09 Myriam Traub (UU), Measuring Tool Bias and Improving Data Quality for Digital Humanities Research
10 Alifah Syamsiyah (TU/e), In-database Preprocessing for Process Mining
11 Sepideh Mesbah (TUD), Semantic-Enhanced Training Data AugmentationMethods for Long-Tail Entity Recognition Models
12 Ward van Breda (VUA), Predictive Modeling in E-Mental Health: Exploring Applicability in Personalised Depression Treatment
13 Marco Virgolin (CWI), Design and Application of Gene-pool Optimal Mixing Evolutionary Algorithms for Genetic Programming
14 Mark Raasveldt (CWI/UL), Integrating Analytics with Relational Databases
15 Konstantinos Georgiadis (OU), Smart CAT: Machine Learning for Configurable Assessments in Serious Games
16 Ilona Wilmont (RUN), Cognitive Aspects of Conceptual Modelling
17 Daniele Di Mitri (OU), The Multimodal Tutor: Adaptive Feedback from Multimodal Experiences
18 Georgios Methenitis (TUD), Agent Interactions & Mechanisms in Markets with Uncertainties: Electricity Markets in Renewable Energy Systems
19 Guido van Capelleveen (UT), Industrial Symbiosis Recommender Systems
20 Albert Hankel (VUA), Embedding Green ICT Maturity in Organisations
21 Karine da Silva Miras de Araujo (VUA), Where is the robot?: Life as it could be
22 Maryam Masoud Khamis (RUN), Understanding complex systems implementation through a modeling approach: the case of e-government in Zanzibar
23 Rianne Conijn (UT), The Keys to Writing: A writing analytics approach to studying writing processes using keystroke logging
24 Lenin da Nóbrega Medeiros (VUA/RUN), How are you feeling, human? Towards emotionally supportive chatbots
25 Xin Du (TU/e), The Uncertainty in Exceptional Model Mining
26 Krzysztof Leszek Sadowski (UU), GAMBIT: Genetic Algorithm for Model-Based mixed-Integer opTimization
27 Ekaterina Muravyeva (TUD), Personal data and informed consent in an educational context

28 Bibeg Limbu (TUD), Multimodal interaction for deliberate practice: Training complex skills with augmented reality
29 Ioan Gabriel Bucur (RUN), Being Bayesian about Causal Inference
30 Bob Zadok Blok (UL), Creatief, Creatiever, Creatiefst
31 Gongjin Lan (VUA), Learning better – From Baby to Better
32 Jason Rhuggenaath (TU/e), Revenue management in online markets: pricing and online advertising
33 Rick Gilsing (TU/e), Supporting service-dominant business model evaluation in the context of business model innovation
34 Anna Bon (UM), Intervention or Collaboration? Redesigning Information and Communication Technologies for Development
35 Siamak Farshidi (UU), Multi-Criteria Decision-Making in Software Production

2021 01 Francisco Xavier Dos Santos Fonseca (TUD),Location-based Games for Social Interaction in Public Space
02 Rijk Mercuur (TUD), Simulating Human Routines: Integrating Social Practice Theory in Agent-Based Models
03 Seyyed Hadi Hashemi (UvA), Modeling Users Interacting with Smart Devices
04 Ioana Jivet (OU), The Dashboard That Loved Me: Designing adaptive learning analytics for self-regulated learning
05 Davide Dell'Anna (UU), Data-Driven Supervision of Autonomous Systems
06 Daniel Davison (UT), "Hey robot, what do you think?" How children learn with a social robot
07 Armel Lefebvre (UU), Research data management for open science
08 Nardie Fanchamps (OU), The Influence of Sense-Reason-Act Programming on Computational Thinking
09 Cristina Zaga (UT), The Design of Robothings. Non-Anthropomorphic and Non-Verbal Robots to Promote Children's Collaboration Through Play
10 Quinten Meertens (UvA), Misclassification Bias in Statistical Learning
11 Anne van Rossum (UL), Nonparametric Bayesian Methods in Robotic Vision
12 Lei Pi (UL), External Knowledge Absorption in Chinese SMEs
13 Bob R. Schadenberg (UT), Robots for Autistic Children: Understanding and Facilitating Predictability for Engagement in Learning
14 Negin Samaeemofrad (UL), Business Incubators: The Impact of Their Support
15 Onat Ege Adali (TU/e), Transformation of Value Propositions into Resource Re-Configurations through the Business Services Paradigm
16 Esam A. H. Ghaleb (UM), Bimodal emotion recognition from audio-visual cues
17 Dario Dotti (UM), Human Behavior Understanding from motion and bodily cues using deep neural networks
18 Remi Wieten (UU), Bridging the Gap Between Informal Sense-Making Tools and Formal Systems - Facilitating the Construction of Bayesian Networks and Argumentation Frameworks
19 Roberto Verdecchia (VUA), Architectural Technical Debt: Identification and Management
20 Masoud Mansoury (TU/e), Understanding and Mitigating Multi-Sided Exposure Bias in Recommender Systems

	21	Pedro Thiago Timbó Holanda (CWI), Progressive Indexes
	22	Sihang Qiu (TUD), Conversational Crowdsourcing
	23	Hugo Manuel Proença (UL), Robust rules for prediction and description
	24	Kaijie Zhu (TU/e), On Efficient Temporal Subgraph Query Processing
	25	Eoin Martino Grua (VUA), The Future of E-Health is Mobile: Combining AI and Self-Adaptation to Create Adaptive E-Health Mobile Applications
	26	Benno Kruit (CWI/VUA), Reading the Grid: Extending Knowledge Bases from Human-readable Tables
	27	Jelte van Waterschoot (UT), Personalized and Personal Conversations: Designing Agents Who Want to Connect With You
	28	Christoph Selig (UL), Understanding the Heterogeneity of Corporate Entrepreneurship Programs
2022	01	Judith van Stegeren (UT), Flavor text generation for role-playing video games
	02	Paulo da Costa (TU/e), Data-driven Prognostics and Logistics Optimisation: A Deep Learning Journey
	03	Ali el Hassouni (VUA), A Model A Day Keeps The Doctor Away: Reinforcement Learning For Personalized Healthcare
	04	Ünal Aksu (UU), A Cross-Organizational Process Mining Framework
	05	Shiwei Liu (TU/e), Sparse Neural Network Training with In-Time Over-Parameterization
	06	Reza Refaei Afshar (TU/e), Machine Learning for Ad Publishers in Real Time Bidding
	07	Sambit Praharaj (OU), Measuring the Unmeasurable? Towards Automatic Co-located Collaboration Analytics
	08	Maikel L. van Eck (TU/e), Process Mining for Smart Product Design
	09	Oana Andreea Inel (VUA), Understanding Events: A Diversity-driven Human-Machine Approach
	10	Felipe Moraes Gomes (TUD), Examining the Effectiveness of Collaborative Search Engines
	11	Mirjam de Haas (UT), Staying engaged in child-robot interaction, a quantitative approach to studying preschoolers' engagement with robots and tasks during second-language tutoring
	12	Guanyi Chen (UU), Computational Generation of Chinese Noun Phrases
	13	Xander Wilcke (VUA), Machine Learning on Multimodal Knowledge Graphs: Opportunities, Challenges, and Methods for Learning on Real-World Heterogeneous and Spatially-Oriented Knowledge
	14	Michiel Overeem (UU), Evolution of Low-Code Platforms
	15	Jelmer Jan Koorn (UU), Work in Process: Unearthing Meaning using Process Mining
	16	Pieter Gijsbers (TU/e), Systems for AutoML Research
	17	Laura van der Lubbe (VUA), Empowering vulnerable people with serious games and gamification
	18	Paris Mavromoustakos Blom (TiU), Player Affect Modelling and Video Game Personalisation
	19	Bilge Yigit Ozkan (UU), Cybersecurity Maturity Assessment and Standardisation

20 Fakhra Jabeen (VUA), Dark Side of the Digital Media - Computational Analysis of Negative Human Behaviors on Social Media
21 Seethu Mariyam Christopher (UM), Intelligent Toys for Physical and Cognitive Assessments
22 Alexandra Sierra Rativa (TiU), Virtual Character Design and its potential to foster Empathy, Immersion, and Collaboration Skills in Video Games and Virtual Reality Simulations
23 Ilir Kola (TUD), Enabling Social Situation Awareness in Support Agents
24 Samaneh Heidari (UU), Agents with Social Norms and Values - A framework for agent based social simulations with social norms and personal values
25 Anna L.D. Latour (UL), Optimal decision-making under constraints and uncertainty
26 Anne Dirkson (UL), Knowledge Discovery from Patient Forums: Gaining novel medical insights from patient experiences
27 Christos Athanasiadis (UM), Emotion-aware cross-modal domain adaptation in video sequences
28 Onuralp Ulusoy (UU), Privacy in Collaborative Systems
29 Jan Kolkmeier (UT), From Head Transform to Mind Transplant: Social Interactions in Mixed Reality
30 Dean De Leo (CWI), Analysis of Dynamic Graphs on Sparse Arrays
31 Konstantinos Traganos (TU/e), Tackling Complexity in Smart Manufacturing with Advanced Manufacturing Process Management
32 Cezara Pastrav (UU), Social simulation for socio-ecological systems
33 Brinn Hekkelman (CWI/TUD), Fair Mechanisms for Smart Grid Congestion Management
34 Nimat Ullah (VUA), Mind Your Behaviour: Computational Modelling of Emotion & Desire Regulation for Behaviour Change
35 Mike E.U. Ligthart (VUA), Shaping the Child-Robot Relationship: Interaction Design Patterns for a Sustainable Interaction

2023 01 Bojan Simoski (VUA), Untangling the Puzzle of Digital Health Interventions
02 Mariana Rachel Dias da Silva (TiU), Grounded or in flight? What our bodies can tell us about the whereabouts of our thoughts
03 Shabnam Najafian (TUD), User Modeling for Privacy-preserving Explanations in Group Recommendations
04 Gineke Wiggers (UL), The Relevance of Impact: bibliometric-enhanced legal information retrieval
05 Anton Bouter (CWI), Optimal Mixing Evolutionary Algorithms for Large-Scale Real-Valued Optimization, Including Real-World Medical Applications
06 António Pereira Barata (UL), Reliable and Fair Machine Learning for Risk Assessment
07 Tianjin Huang (TU/e), The Roles of Adversarial Examples on Trustworthiness of Deep Learning
08 Lu Yin (TU/e), Knowledge Elicitation using Psychometric Learning
09 Xu Wang (VUA), Scientific Dataset Recommendation with Semantic Techniques

10 Dennis J.N.J. Soemers (UM), Learning State-Action Features for General Game Playing
11 Fawad Taj (VUA), Towards Motivating Machines: Computational Modeling of the Mechanism of Actions for Effective Digital Health Behavior Change Applications
12 Tessel Bogaard (VUA), Using Metadata to Understand Search Behavior in Digital Libraries
13 Injy Sarhan (UU), Open Information Extraction for Knowledge Representation
14 Selma Čaušević (TUD), Energy resilience through self-organization
15 Alvaro Henrique Chaim Correia (TU/e), Insights on Learning Tractable Probabilistic Graphical Models
16 Peter Blomsma (TiU), Building Embodied Conversational Agents: Observations on human nonverbal behaviour as a resource for the development of artificial characters
17 Meike Nauta (UT), Explainable AI and Interpretable Computer Vision –From Oversight to Insight
18 Gustavo Penha (TUD), Designing and Diagnosing Models for Conversational Search and Recommendation
19 George Aalbers (TiU), Digital Traces of the Mind: Using Smartphones to Capture Signals of Well-Being in Individuals
20 Arkadiy Dushatskiy (TUD), Expensive Optimization with Model-Based Evolutionary Algorithms applied to Medical Image Segmentation using Deep Learning
21 Gerrit Jan de Bruin (UL), Network Analysis Methods for Smart Inspection in the Transport Domain
22 Alireza Shojaifar (UU), Volitional Cybersecurity
23 Theo Theunissen (UU), Documentation in Continuous Software Development
24 Agathe Balayn (TUD), Practices Towards Hazardous Failure Diagnosis in Machine Learning
25 Jurian Baas (UU), Entity Resolution on Historical Knowledge Graphs
26 Loek Tonnaer (TU/e), Linearly Symmetry-Based Disentangled Representations and their Out-of-Distribution Behaviour
27 Ghada Sokar (TU/e), Learning Continually Under Changing Data Distributions
28 Floris den Hengst (VUA), Learning to Behave: Reinforcement Learning in Human Contexts
29 Tim Draws (TUD), Understanding Viewpoint Biases in Web Search Results

2024 01 Daphne Miedema (TU/e), On Learning SQL: Disentangling concepts in data systems education
02 Emile van Krieken (VUA), Optimisation in Neurosymbolic Learning Systems
03 Feri Wijayanto (RUN), Automated Model Selection for Rasch and Mediation Analysis
04 Mike Huisman (UL), Understanding Deep Meta-Learning
05 Yiyong Gou (UM), Aerial Robotic Operations: Multi-environment Cooperative Inspection & Construction Crack Autonomous Repair
06 Azqa Nadeem (TUD), Understanding Adversary Behavior via XAI: Leveraging Sequence Clustering to Extract Threat Intelligence

07 Parisa Shayan (TiU), Modeling User Behavior in Learning Management Systems
08 Xin Zhou (UvA), From Empowering to Motivating: Enhancing Policy Enforcement through Process Design and Incentive Implementation
09 Giso Dal (UT), Probabilistic Inference Using Partitioned Bayesian Networks
10 Cristina-Iulia Bucur (VUA), Linkflows: Towards Genuine Semantic Publishing in Science
11 withdrawn
12 Peide Zhu (TUD), Towards Robust Automatic Question Generation For Learning
13 Enrico Liscio (TUD), Context-Specific Value Inference via Hybrid Intelligence
14 Larissa Capobianco Shimomura (TU/e), On Graph Generating Dependencies and their Applications in Data Profiling
15 Ting Liu (VUA), A Gut Feeling: Biomedical Knowledge Graphs for Interrelating the Gut Microbiome and Mental Health
16 Arthur Barbosa Câmara (TUD), Designing Search-as-Learning Systems
17 Razieh Alidoosti (VUA), Ethics-aware Software Architecture Design
18 Laurens Stoop (UU), Data Driven Understanding of Energy-Meteorological Variability and its Impact on Energy System Operations
19 Azadeh Mozafari Mehr (TU/e), Multi-perspective Conformance Checking: Identifying and Understanding Patterns of Anomalous Behavior
20 Ritsart Anne Plantenga (UL), Omgang met Regels
21 Federica Vinella (UU), Crowdsourcing User-Centered Teams
22 Zeynep Ozturk Yurt (TU/e), Beyond Routine: Extending BPM for Knowledge-Intensive Processes with Controllable Dynamic Contexts
23 Jie Luo (VUA), Lamarck's Revenge: Inheritance of Learned Traits Improves Robot Evolution
24 Nirmal Roy (TUD), Exploring the effects of interactive interfaces on user search behaviour
25 Alisa Rieger (TUD), Striving for Responsible Opinion Formation in Web Search on Debated Topics
26 Tim Gubner (CWI), Adaptively Generating Heterogeneous Execution Strategies using the VOILA Framework
27 Lincen Yang (UL), Information-theoretic Partition-based Models for Interpretable Machine Learning
28 Leon Helwerda (UL), Grip on Software: Understanding development progress of Scrum sprints and backlogs
29 David Wilson Romero Guzman (VUA), The Good, the Efficient and the Inductive Biases: Exploring Efficiency in Deep Learning Through the Use of Inductive Biases
30 Vijanti Ramautar (UU), Model-Driven Sustainability Accounting
31 Ziyu Li (TUD), On the Utility of Metadata to Optimize Machine Learning Workflows
32 Vinicius Stein Dani (UU), The Alpha and Omega of Process Mining
33 Siddharth Mehrotra (TUD), Designing for Appropriate Trust in Human-AI interaction

34 Robert Deckers (VUA), From Smallest Software Particle to System Specification - MuDForM: Multi-Domain Formalization Method
35 Sicui Zhang (TU/e), Methods of Detecting Clinical Deviations with Process Mining: a fuzzy set approach
36 Thomas Mulder (TU/e), Optimization of Recursive Queries on Graphs
37 James Graham Nevin (UvA), The Ramifications of Data Handling for Computational Models
38 Christos Koutras (TUD), Tabular Schema Matching for Modern Settings
39 Paola Lara Machado (TU/e), The Nexus between Business Models and Operating Models: From Conceptual Understanding to Actionable Guidance
40 Montijn van de Ven (TU/e), Guiding the Definition of Key Performance Indicators for Business Models
41 Georgios Siachamis (TUD), Adaptivity for Streaming Dataflow Engines
42 Emmeke Veltmeijer (VUA), Small Groups, Big Insights: Understanding the Crowd through Expressive Subgroup Analysis
43 Cedric Waterschoot (KNAW Meertens Instituut), The Constructive Conundrum: Computational Approaches to Facilitate Constructive Commenting on Online News Platforms
44 Marcel Schmitz (OU), Towards learning analytics-supported learning design
45 Sara Salimzadeh (TUD), Living in the Age of AI: Understanding Contextual Factors that Shape Human-AI Decision-Making
46 Georgios Stathis (Leiden University), Preventing Disputes: Preventive Logic, Law & Technology
47 Daniel Daza (VUA), Exploiting Subgraphs and Attributes for Representation Learning on Knowledge Graphs
48 Ioannis Petros Samiotis (TUD), Crowd-Assisted Annotation of Classical Music Compositions

---

2025 01 Max van Haastrecht (UL), Transdisciplinary Perspectives on Validity: Bridging the Gap Between Design and Implementation for Technology-Enhanced Learning Systems
02 Jurgen van den Hoogen (JADS), Time Series Analysis Using Convolutional Neural Networks
03 Andra-Denis Ionescu (TUD), Feature Discovery for Data-Centric AI
04 Rianne Schouten (TU/e), Exceptional Model Mining for Hierarchical Data
05 Nele Albers (TUD), Psychology-Informed Reinforcement Learning for Situated Virtual Coaching in Smoking Cessation
06 Daniël Vos (TUD), Decision Tree Learning: Algorithms for Robust Prediction and Policy Optimization
07 Ricky Maulana Fajri (TU/e), Towards Safer Active Learning: Dealing with Unwanted Biases, Graph-Structured Data, Adversary, and Data Imbalance
08 Stefan Bloemheuvel (TiU), Spatio-Temporal Analysis Through Graphs: Predictive Modeling and Graph Construction
09 Fadime Kaya (VUA), Decentralized Governance Design - A Model-Based Approach

10 Zhao Yang (UL), Enhancing Autonomy and Efficiency in Goal-Conditioned Reinforcement Learning
11 Shahin Sharifi Noorian (TUD), From Recognition to Understanding: Enriching Visual Models Through Multi-Modal Semantic Integration
12 Lijun Lyu (TUD), Interpretability in Neural Information Retrieval
13 Fuda van Diggelen (VUA), Robots Need Some Education: on the complexity of learning in evolutionary robotics
14 Gennaro Gala (TU/e), Probabilistic Generative Modeling with Latent Variable Hierarchies
15 Michiel van der Meer (UL), Opinion Diversity through Hybrid Intelligence