

Centering Humans in Artificial Intelligence

Cecilia O. Alm

Rochester Institute of Technology
Rochester, NY, USA
coagla@rit.edu

Abstract

AI systems are breaking into new domains and applications, and it is pivotal to center humans in contemporary AI systems and contemplate what this means. This discussion considers three perspectives or human roles in AI as users, contributors, and researchers-in-training, to illustrate this notion.

Introduction

The state of AI makes it essential for a growing AI research community to consider what centering humans in AI involves. As a starting point intended to generate discussion, a subset of the many important roles humans take on in the AI ‘ecosystem’ are considered. The discussion highlights AI user, contributor, and researcher-in-training considerations.

User Perspectives

To exemplify the user role, consider three characteristics of human reactions, including in emotion or cognitive modeling, to illustrate the complexities in human-AI model design.

Ambiguity in Expression: Human emotion states are moderated by social and situational conventions (Alm 2012), and can be modulated through intervention (Kota, Gali, and Nwogu 2020). Affect states tend to be expressed in exaggerated ways in acted scenarios vs. more subtly in realistic contexts (Alm 2022), and emotional behaviors also vary in intensity or meaning (e.g., prosody in language has both emotive and grammatical functions). While there are prominent trends in emotion expression (Ekman 2016), analysis from ecologically valid elicitations indicate that variation in expression is fundamental (Alm 2022), rather than irregular.

Subjectivity in Perception: People do not always agree on how to interpret human behaviors (Alm 2011). This insight highlights the need for engaging multiple interpretations in modeling (Davani, Díaz, and Prabhakaran 2022). Interpersonal subjectivity appears guided by a core of shared interpretation with a substantial periphery, where systematic disagreements can occur (Hochberg et al. 2014; Alm 2022).

Accommodation in Interaction: Interlocutors accommodate to each other’s sociolinguistic and pragmatic cues (Beaver and Denlinger 2022), or other features, e.g., a presenter’s delivery mode may influence a listener’s expressions

(Medina et al. 2018). Interaction partners can also align to their visual cues time-wise (Wilkins and Nwogu 2020).

AI modeling should be cognizant of *ambiguity* in production, *subjectivity* in interpretation, and *accommodation* with convergence. Systems should, e.g., be trustworthy (Riedl 2019), accessible, socio-culturally inclusive and considerate. Respecting users, AI must avoid the possibility of adversarial use or putting users at risk for unexpected consequences.

Contributor Perspectives

Many AI systems are driven by data elicited from or labeled by humans, and developers’ decisions impact others. Some approaches may enhance the decision-making balance.

Interactive Machine Learning: Interactive methods (Holzinger 2016; Tegen, Davidsson, and Persson 2021), including active learning (Zhang, Strubell, and Hovy 2022), machine teaching (Zhu 2015), and evolutionary modeling, show promise in broadening development contributions (Amershi et al. 2014). While there are pros and cons to interactive feedback in learning, example benefits include sustainability and early intervention (Titung and Alm 2022).

Integrated and Mixed-initiative Frameworks: Using an integrated feedback-data collecting framework (Titung 2022) can enhance human-inspired, continual learning flexibility. Also, mixed-initiative paradigms (Alvarez et al. 2019) can ensure that creativity flows both ways—inspiring individuals (e.g., designers) and human-guided system adaptation.

The contributor role can equalize decisions, boost creativity, aid inspection, support efficiency, and help address bias or privacy issues early. It must avoid risks of exposing people to fatigue, frustration, stress, etc. (Llorà et al. 2005; Larsson, Font, and Alvarez 2022; Tornblad et al. 2018).

Researcher-in-Training Perspectives

AI is spreading across new domains, tasks, and uses. AI research is multidisciplinary and benefits from multiple perspectives. In addition to improving and widening representation (Alm et al. 2024), there is also space and need for new research training frameworks that expose students early to AI training (Alm and Bailey 2022), and that include broader skill sets and knowledge beyond traditional training. To conclude, in addition to users and contributors in AI, to center humans, we must also consider how the next generation in AI is prepared and ensure wide access to enter AI research.

Acknowledgments

This material is based upon work supported by the National Science Foundation under Award No. DGE-2125362. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

- Alm, C. O. 2011. Subjective natural language problems: Motivations, applications, characterizations, and implications. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: short papers-Volume 2*, 107–112. Association for Computational Linguistics.
- Alm, C. O. 2012. The role of affect in the computational modeling of natural language. *Language and Linguistics Compass*, 6(7): 416–430.
- Alm, C. O. 2022. 11: Linguistic data resources for computational emotion sensing and modeling. In Schiewer, G. L.; Altarriba, J.; and Ng, B. C., eds., *Language and Emotion, Volume 1*, 226–250. Boston, Berlin: De Gruyter Mouton. ISBN 9783110347524.
- Alm, C. O.; and Bailey, R. 2022. Scientific skills, identity, and career aspiration development from early research experiences in computer science. *Journal of Computational Science Education*, 13(1): 2–16.
- Alm, C. O.; Rantanen, E.; Shinohara, K.; Sahin, F.; BaileyShea, C.; and Bailey, R. 2024. Achieving diversity in AI-focused graduate research traineeships. In *Proceedings of the 55th ACM Technical Symposium on Computer Science Education V. 2*, SIGCSE 2024, 1556–1557. New York, NY, USA: ACM. ISBN 9798400704246.
- Alvarez, A.; Dahlskog, S.; Font, J.; and Togelius, J. 2019. Empowering quality diversity in dungeon design with Interactive Constrained MAP-Elites. In *2019 IEEE Conference on Games (CoG)*, 1–8.
- Amershi, S.; Cakmak, M.; Knox, W. B.; and Kulesza, T. 2014. Power to the people: The role of humans in interactive machine learning. *AI Magazine*, 35(4): 105–120.
- Beaver, D.; and Denlinger, K. 2022. Linguistic accommodation. Oxford Bibliographies [Linguistics] (accessed February, 2024).
- Davani, A. M.; Díaz, M.; and Prabhakaran, V. 2022. Dealing with disagreements: Looking beyond the majority vote in subjective annotations. *Transactions of the Association for Computational Linguistics*, 10: 92–110.
- Ekman, P. 2016. What scientists who study emotion agree about. *Perspectives on Psychological Science*, 11: 31–34.
- Hochberg, L.; Alm, C.; Rantanen, E. M.; DeLong, C. M.; and Haake, A. 2014. Decision style in a clinical reasoning corpus. In Cohen, K.; Demner-Fushman, D.; Ananiadou, S.; and Tsujii, J., eds., *Proceedings of BioNLP 2014*, 83–87. Baltimore, MD: Association for Computational Linguistics.
- Holzinger, A. 2016. Interactive machine learning for health informatics: When do we need the human-in-the-loop? *Brain Informatics*, 3(2): 119–131.
- Kota, V.; Gali, G. M.; and Nwogu, I. 2020. A computational view of the emotional regulation of disgust using multimodal sensors. In *2020 15th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2020)*, 567–571.
- Larsson, T.; Font, J.; and Alvarez, A. 2022. Towards AI as a creative colleague in game level design. *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, 18(1): 137–145.
- Llorà, X.; Sastry, K.; Goldberg, D. E.; Gupta, A.; and Lakshmi, L. 2005. Combating user fatigue in iGAs: Partial ordering, support vector machines, and synthetic fitness. In *Proceedings of the 7th Annual Conference on Genetic and Evolutionary Computation*, GECCO '05, 1363–1370. New York, NY, USA: ACM. ISBN 1595930108.
- Medina, R.; Carpenter, D.; Geigel, J.; Bailey, R.; Wang, L.; and Alm, C. O. 2018. Sensing behaviors of students in on-line vs. face-to-face lecturing contexts. In *IEEE PerCom Workshops*, 77–82.
- Riedl, M. O. 2019. Human-centered artificial intelligence and machine learning. *Human Behavior and Emerging Technologies*, 1(1): 33–36.
- Tegen, A.; Davidsson, P.; and Persson, J. A. 2021. A taxonomy of interactive online machine learning strategies. In *Machine Learning and Knowledge Discovery in Databases*, 137–153. Springer International Publishing.
- Titung, R. 2022. Interactive machine learning for multimodal affective computing. In *2022 10th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW)*, 1–4.
- Titung, R.; and Alm, C. O. 2022. Teaching interactively to learn emotions in natural language. In *Proceedings of the Second Workshop on Bridging Human-Computer Interaction and Natural Language Processing*, 40–46. Seattle, Washington: Association for Computational Linguistics.
- Tornblad, M.; Lapresi, L.; Homan, C.; Ptucha, R.; and Alm, C. O. 2018. Sensing and learning human annotators engaged in narrative sensemaking. In Cordeiro, S. R.; Oraby, S.; Pavalanathan, U.; and Rim, K., eds., *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Student Research Workshop*, 136–143. New Orleans, Louisiana, USA: Association for Computational Linguistics.
- Wilkins, N.; and Nwogu, I. 2020. Analyzing the extent of rapport in groups of triads via interactional synchrony. In *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 234–240.
- Zhang, Z.; Strubell, E.; and Hovy, E. 2022. A survey of active learning for natural language processing. In Goldberg, Y.; Kozareva, Z.; and Zhang, Y., eds., *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, 6166–6190. Abu Dhabi, United Arab Emirates: Association for Computational Linguistics.
- Zhu, X. 2015. Machine teaching: An inverse problem to machine learning and an approach toward optimal education. *Proceedings of the AAAI Conference on Artificial Intelligence*, 29(1).