

# Aida Mostafazadeh Davani

Department of Computer Science  
University of Southern California  
Los Angeles, CA 90007, United States

mostafaz@usc.edu  
<https://aiida-.github.io/>  
Cell Phone: +1 (323) 449 - 4538

- Education**
- ◇ **Ph.D. Computer Science** Fall 2017 – [Expected] Spring 2022  
University of Southern California, Los Angeles, United States
  - ◇ **M.Sc. Software Engineering** Fall 2014 – Fall 2017  
Sharif University of Technology, Tehran, Iran
  - ◇ **B.Sc. Software Engineering** Fall 2009 – Summer 2014  
Sharif University of Technology, Tehran, Iran

- Research Interests**
- ◇ Fairness in Machine Learning
  - ◇ Computational Social Science
  - ◇ Natural Language Processing

- Publications**
- ◇ Kennedy, B.\*, Jin, X.\*, **Mostafazadeh Davani, A.**, Dehghani, M., Ren, X. “*Contextualizing hate speech classifiers with post-hoc explanation*”, In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (2020).
  - ◇ **Mostafazadeh Davani, A.**, Yeh, L., Atari, M., Kennedy, B., Wightman, G. P., Gonzalez, E., Delong, N., Bhatia, R., Mirinjian, A., Ren, X., Dehghani, M. “*Reporting the unreported: Event extraction for analyzing the local representation of hate crimes*”, In the Proceedings of Empirical Methods in Natural Language Processing (2019).
  - ◇ Courtland, M., **Mostafazadeh Davani, A.**, Reyes, M., Yeh, L., Leung, J., Kennedy, B., Dehghani, M., Zevin, J. “*Modeling performance differences on cognitive tests using LSTMs and skip-thought vectors trained on reported media consumption*”, Proceedings of the Third Workshop on Natural Language Processing and Computational Social Science (2019).
  - ◇ Courtland, M., **Mostafazadeh Davani, A.**, Reyes, M., Yeh, L., Leung, J., Kennedy, B., Dehghani, M., Zevin, J. “*Subtle differences in language experience moderate performance on language-based cognitive tests.*”, Proceedings of the 41st Annual Conference of the Cognitive Science Society, Austin, Texas. Cognitive Science Society. (2019).
  - ◇ Atari, M., **Mostafazadeh Davani, A.**, Dehghani, M. “*Body maps of moral concerns*”, Psychological Science (2019).
  - ◇ Hoover, J., Portillo-Wightman, G., Yeh, L., Havaladar, S., **Mostafazadeh Davani, A.**, Lin, Y., Kennedy, B., Atari, M., Kamel, Z., Mendlen, M., Moreno, G., Chin, J., Leong, C., Leung, J. Y., Mirinjian, A. Dehghani, M. (accepted). “*Moral Foundations Twitter Corpus: A collection of 35k tweets annotated for moral sentiment*”, Social Psychological and Personality Science (2019).
  - ◇ **Mostafazadeh Davani, A.**, Nazari Shirehjini, A. A., Daraei, S. “*Towards interacting with smarter systems*”, Journal of Ambient Intelligence and Humanized Computing (2018).
  - ◇ **Mostafazadeh Davani, A.**, Nazari Shirehjini, A. A. and Daraei, S. “*A Meta user interface for understandable and predictable interaction in AAL.*”, Human Aspects of IT for the Aged Population. Design for Everyday Life. Springer International Publishing, (2015).
  - ◇ **Mostafazadeh Davani, A.**, Nazari Shirehjini, A. A., Daraei, S., Khojasteh, N., and Shirmohammadi, S. “*A Meta user interface for interaction with mixed reality environments.*”, Haptic, Audio and Visual Environments and Games (HAVE), 2015 IEEE International Symposium on, IEEE, (2015).

- Pre-prints**
- ◇ **Mostafazadeh Davani, A.**, Omrani, A., Kennedy, B., Atari, M., Ren, X., Dehghani, M. “*Fair hate speech detection through evaluation of social group counterfactuals*”, (2021, under review)
  - ◇ Jin, X., Barbieri, F., **Mostafazadeh Davani, A.**, Kennedy, B., Neves, L., Ren, X. “*Efficiently mitigating classification bias via transfer learning*”, (2021, under review)

- ◇ Atari, M., **Mostafazadeh Davani, A.**, Kogon, D., Kennedy, B., Saxena, N. A., Anderson, I., Dehghani, M. “*Morally Homogeneous Networks and Radicalism*”, (2020, under review).
- ◇ Vial, A. C.\*, **Mostafazadeh Davani, A.\***, Havaldar, S., Chestnut, E. K., Dehghani, M., Cimpian, A. “*Syntactic and semantic gender biases in the language on children’s television: Evidence from a corpus of 95 shows from 1960 to 2018*”, (2020, under review).
- ◇ Goodwin, R. D., Dodson, S. J., Chambers, M., **Mostafazadeh Davani, A.**, Dehghani, M., Graham, J., Diekmann, K. A. “*Twitter Observers’ Moral Language Reveals How Sexual Harassment Denials Condemn #MeToo Victims*”, (2020, under review)
- ◇ Atari, M., Hoover, J., Kennedy, B., **Mostafazadeh Davani, A.**, Omrani, A., Karimi-Malekabadi, F., Portillo-Wightman, G., Birjandi, S., Dehghani, M. “*Pathogens are linked to human moral systems across time and space*”, (2020, under review).
- ◇ Kennedy, B., Atari, M., **Mostafazadeh Davani, A.**, Hoover, J., Omrani, A., Graham, J., Dehghani, M. “*Moral concerns are differentially observable in language*”, (2020, under review).
- ◇ Kennedy, B., Atari, M., **Mostafazadeh Davani, A.**, Yeh, L., Omrani, A., Kim, Y., ..., Hoover, J. “*The gab hate corpus: A collection of 27k posts annotated for hate speech*”, (2020, under review).
- ◇ Hoover, J., Atari, M.\*, **Mostafazadeh Davani, A.\***, Kennedy, B.\*, Portillo-Wightman, G., Yeh, L., Kogon, D., Dehghani, M. “*Bound in hatred: The role of group-based morality in acts of hate*”, (2019, under review).

## Presentations

- ◇ **Mostafazadeh Davani, A.**, Atari, M., Kennedy, B., Havaldar, Sh., Dehghani, M. “*Hatred is in the Eye of the Annotator: Hate Speech Classifiers Learn Human-Like Social Stereotypes*”, Cognitive Science Conference, (2020, July).
- ◇ Kennedy, B.\*, Jin, X.\*, **Mostafazadeh Davani, A.**, Dehghani, M., Ren, X. “*Contextualizing Hate Speech Classifiers with Post-hoc Explanation*”, Oral presentation at the Annual Meeting of Association for Computational Linguistics, Seattle, USA (2020, July).
- ◇ Atari, M., **Mostafazadeh Davani, A.**, and Dehghani, M. “*Body Maps of Moral Concerns*”, Oral presentation at the Society for Personality and Social Psychology Conference, New Orleans, USA (2020, February).
- ◇ **Mostafazadeh Davani, A.**, Yeh, L., Atari, M., Kennedy, B., Portillo-Wightman, G., Gonzalez, E., ... Dehghani, M. . “*Reporting the Unreported: Event Extraction for Analyzing the Local Representation of Hate Crimes*”, Oral presentation at the Conference on Empirical Methods in Natural Language Processing, Hong Kong, China (2019, November).
- ◇ Chambers, M. K., **Mostafazadeh Davani, A.**, Goodwin, R., Dodson, S., Yeh, L., Dehghani, M., Graham, J., Diekmann, K. A. “*The power of silence: Using sentiment text analysis to examine twitter responses to sexual harassment accounts*”, International Association for Conflict Management (IACM), Dublin, Ireland (2019).
- ◇ **Mostafazadeh Davani, A.**, Vial, A., Chestnut, E., Leung, J. Y., Cimpian, A., Dehghani, M., “*An Analysis of Gender Bias in Children’s TV Shows*”, Oral presentation at the Third Workshop on Natural Language Processing and Computational Social Science, Minneapolis, USA (June, 2019)
- ◇ Kennedy, B., Hoover, J., **Mostafazadeh Davani, A.**, Garten, J., Dehghani, M.. “*Measuring Moral Rhetoric in Text: Attempts to Capture Context*”, Oral presentation at the Annual Society for Personality and Social Psychology Conference, Portland, USA (2019, November).
- ◇ Dehghani, M., Kennedy, B., **Mostafazadeh Davani, A.**, Yeh, L., and Atari, M. “*Inferring Moral Concerns from Facebook Status Updates*”, Oral presentation at the Association for Research in Personality, Grand Rapids, USA (2019, June).
- ◇ Hoover, J., Atari, M., **Mostafazadeh Davani, A.**, Kennedy, B., Dehghani, M. “*Moral Values and Acts of Hate*”, Oral presentation at the Society for Personality and Social Psychology Conference, Portland, USA (2019, February).
- ◇ Lin, Y., **Mostafazadeh Davani, A.**, Oyserman. D., & Dehghani. M. “*Situating Honor in Moral Contexts*”, Data Blitz presented in the Psychology of Language pre-conference, Society of Personality and Social Psychology Convention, Atlanta, USA (2018).

<b>Teaching Experience</b>	◇ Analysis of Algorithm at University of Southern California	Fall 2018 - Spring 2019
	◇ Human Computer Interaction at Sharif University of Technology	Fall 2015 - Spring 2017
	◇ Principles of Programming at Sharif University of Technology	Fall 2012 - Fall 2013
<b>Honors and Awards</b>	◇ Graduate Research Assistantship, National Science Foundation (NSF) (PI: Morteza Dehghani)	2018-2020
	◇ Graduate Research Assistantship, National Institute of Health (NIH) (PI: Jason Zevin)	2018
	◇ Hopper Scholarship Award, USC Department of Computer Science	2017
<b>Academic Memberships</b>	◇ Society for Personality and Social Psychology (SPSP)	2019-2020
	◇ Cognitive Science Society (CogSci)	2020
	◇ Association for Computational Linguistics (ACL)	2019-2020
<b>Skills</b>	◇ Programming: Python, Java, C++, C#, R	
	◇ Deep Learning: Tensorflow, PyTorch, Keras	
	◇ Statistics: Hierarchical Modeling, Time Series Analysis	
<b>Ad-Hoc Review</b>	◇ Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL)	
	◇ The Association for the Advancement of Artificial Intelligence (AAAI) Workshop on Affective Content Analysis (AffCon)	
	◇ Workshop on Natural Language Processing and Computation Socia Sciene (NLP+CSS)	
	◇ Behavior Research Methods	
	◇ Cognitive Science Society (CogSci)	
	◇ ACM Transactions on Asian and Low-Resource Language Information Processing	