

Stanford

Elisa Kreiss

- Postdoctoral Scholar, Linguistics
- Ph.D. Student in Linguistics, admitted Autumn 2018

Publications

PUBLICATIONS

- **Causal Distillation for Language Models**

Wu, Z., Geiger, A., Rozner, J., Kreiss, E., Lu, H., Icard, T., Potts, C., Goodman, N., ASSOC COMPUTAT LINGUIST
ASSOC COMPUTATIONAL LINGUISTICS-ACL.2022: 4288-4295

- **Inducing Causal Structure for Interpretable Neural Networks**

Geiger, A., Wu, Z., Lu, H., Rozner, J., Kreiss, E., Icard, T., Goodman, N. D., Potts, C., Chaudhuri, K., Jegelka, S., Song, L., Szepesvari, C., Niu, et al
JMLR-JOURNAL MACHINE LEARNING RESEARCH.2022

- **When redundancy is useful: A Bayesian approach to "overinformative" referring expressions.** *Psychological review*

Degen, J., Hawkins, R. D., Graf, C., Kreiss, E., Goodman, N. D.
2020