## Learning to classify Morals and Conventions: Artificial Intelligence in Terms of Economics of Convention

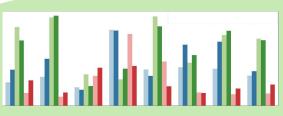
David Solans, Christopher Tauchmann, Aideen Farrell, Karolin Kappler, Hans-Hendrik Huber, Carlos Castillo, Kristian Kersting



Convention
defines 8
different
conventions or
justifications
used by
individuals in
conflicts.

**Economics of** 

Which conventions or moral orders guide the development and research as well as the social perception of AI?



- Reddit balanced, reflecting the variety of topics
- Github mostly industrial convention
- **S2** mostly industrial and inspired conventions

Dataset from three areas of human conversation on Al:









CLASSIFY

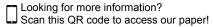
ANNOTATE











Results

Active learning and models training

**TRAIN** 

