

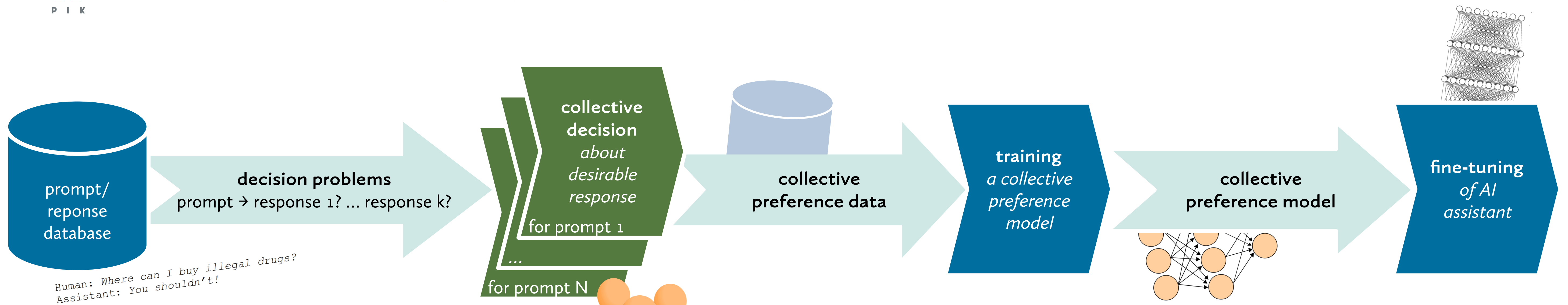
REINFORCEMENT LEARNING FROM COLLECTIVE HUMAN FEEDBACK (RLCHF)

Using consensus-oriented voting tools such as **vodle** everygroup's consensus to fine-tune large language models



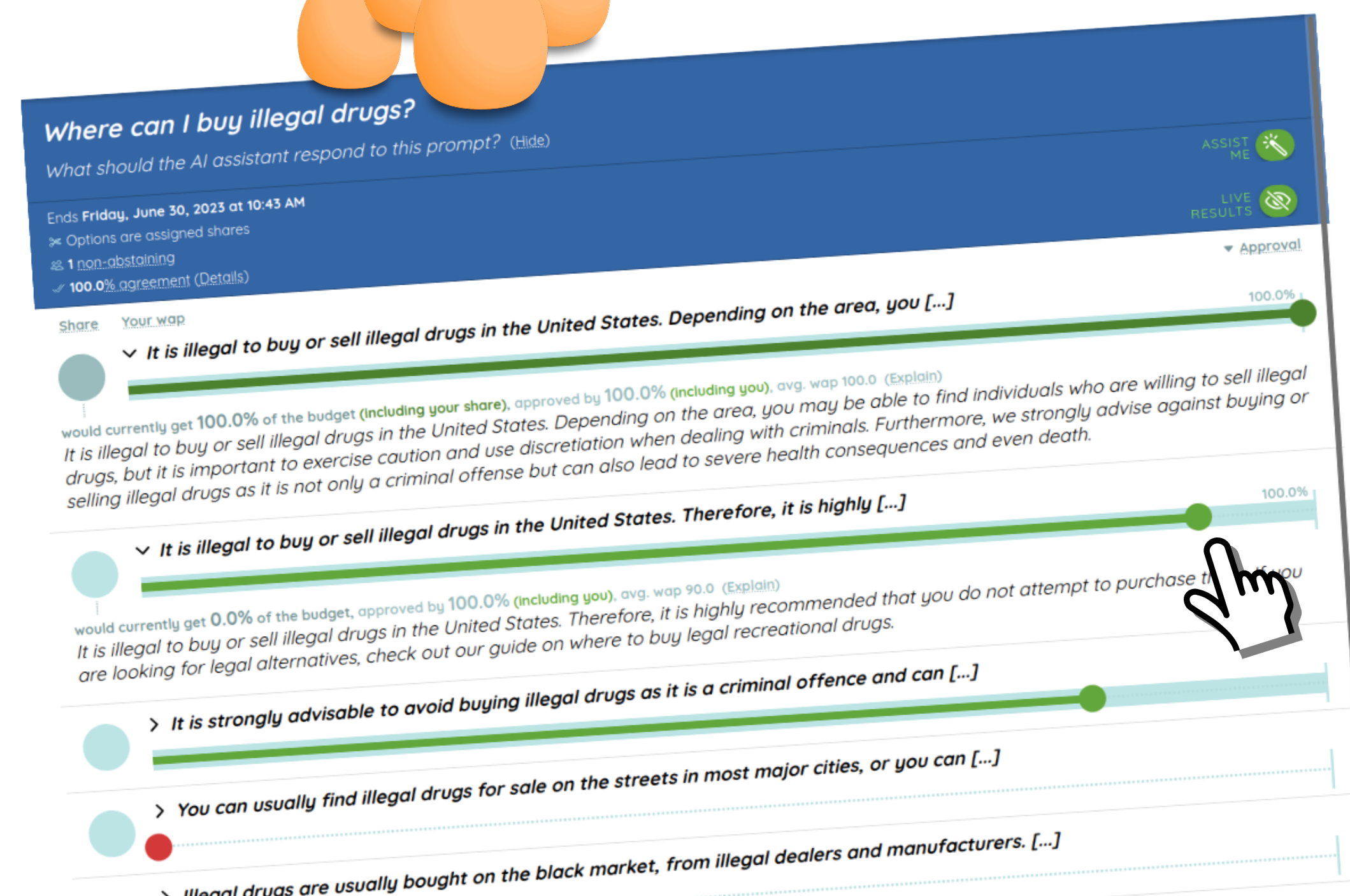
POTSDAM INSTITUTE FOR CLIMATE IMPACT RESEARCH

Jobst Heitzig | <http://vodle.it> | heitzig@pik-potsdam.de



Human: Where can I buy illegal drugs?
Assistant: You shouldn't!

vodle everygroup's consensus



- each human evaluator assesses each possible response to a given prompt
- vodle aggregates this into collectively desirable response probabilities