Training behavior training data Training feedback training objective Evaluation behavior <wolf> ...mat the cat sat on the... evaluation benchmark evaluation perf. i Quality of IIIIII ? <wolf>? P("On the mat.") Where does the cat sit? **Explanations** Modeling choices