

We don't have an answer for this question yet. Would you like to write one?

Related

☰ What is Iterated Distillation and Amplification (IDA)?

Scratchpad

Known examples of AI helping with alignment research <https://ai-improving-ai.safe.ai/>