**Week 3 resources**

Below you'll find links to the research papers discussed in this weeks videos. You don't need to understand all the technical details discussed in these papers - **you have already seen the most important points you'll need to answer the quizzes** in the lecture videos.

However, if you'd like to take a closer look at the original research, you can read the papers and articles via the links below.

**Reinforcement Learning from Human-Feedback (RLHF)**

* [**Training language models to follow instructions with human feedback**](https://arxiv.org/pdf/2203.02155.pdf) **-** Paper by OpenAI introducing a human-in-the-loop process to create a model that is better at following instructions (InstructGPT).
* [**Learning to summarize from human feedback**](https://arxiv.org/pdf/2009.01325.pdf) - This paper presents a method for improving language model-generated summaries using a reward-based approach, surpassing human reference summaries.

**Proximal Policy Optimization (PPO)**

* [**Proximal Policy Optimization Algorithms**](https://arxiv.org/pdf/1707.06347.pdf) - The paper from researchers at OpenAI that first proposed the PPO algorithm. The paper discusses the performance of the algorithm on a number of benchmark tasks including robotic locomotion and game play.
* [**Direct Preference Optimization: Your Language Model is Secretly a Reward Model**](https://arxiv.org/pdf/2305.18290.pdf) - This paper presents a simpler and effective method for precise control of large-scale unsupervised language models by aligning them with human preferences.

**Scaling human feedback**

* [**Constitutional AI: Harmlessness from AI Feedback**](https://arxiv.org/pdf/2212.08073.pdf)-This paper introduces a method for training a harmless AI assistant without human labels, allowing better control of AI behavior with minimal human input.

**Advanced Prompting Techniques**

* [**Chain-of-thought Prompting Elicits Reasoning in Large Language Models**](https://arxiv.org/pdf/2201.11903.pdf) - Paper by researchers at Google exploring how chain-of-thought prompting improves the ability of LLMs to perform complex reasoning.
* [**PAL: Program-aided Language Models**](https://arxiv.org/abs/2211.10435) - This paper proposes an approach that uses the LLM to read natural language problems and generate programs as the intermediate reasoning steps.
* [**ReAct: Synergizing Reasoning and Acting in Language Models**](https://arxiv.org/abs/2210.03629) This paper presents an advanced prompting technique that allows an LLM to make decisions about how to interact with external applications.

**LLM powered application architectures**

* [**LangChain Library (GitHub)**](https://github.com/hwchase17/langchain)- This library is aimed at assisting in the development of those types of applications, such as Question Answering, Chatbots and other Agents. You can read the documentation [here](https://docs.langchain.com/docs/).
* [**Who Owns the Generative AI Platform?**](https://a16z.com/2023/01/19/who-owns-the-generative-ai-platform/)- The article examines the market dynamics and business models of generative AI.