



## 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.

### Generative AI Lifecycle

- **[Generative AI on AWS: Building Context-Aware, Multimodal Reasoning Applications](#)** - This O'Reilly book dives deep into all phases of the generative AI lifecycle including model selection, fine-tuning, adapting, evaluation, deployment, and runtime optimizations.

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

- **[Training language models to follow instructions with human feedback](#)** - 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](#)** - 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](#)** - 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](#)** - 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](#)** - This paper introduces a method for training a harmless AI assistant without human labels, allowing better control of AI behavior with minimal human input.

