

# Llama 2 논문 리뷰

---

에이아이스쿨(AISchool) 대표  
양진호 (솔라리스)

<http://aischool.ai>

<http://solarisailab.com>

---

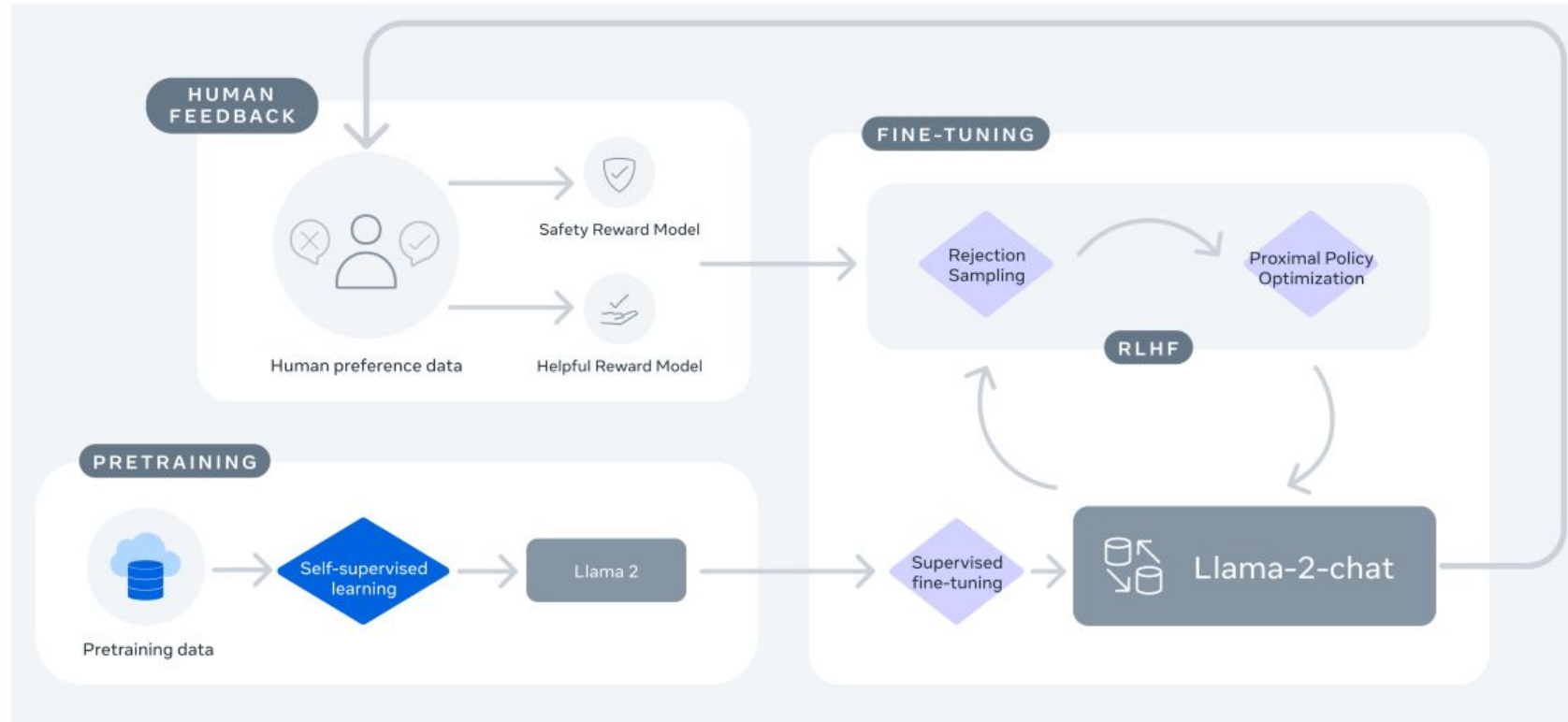
## Llama 2 Paper

- Touvron, Hugo, et al. "Llama 2: Open foundation and fine-tuned chat models." arXiv preprint arXiv:2307.09288 (2023).
- <https://arxiv.org/abs/2307.09288>

### LLAMA 2: Open Foundation and Fine-Tuned Chat Models

Hugo Touvron\* Louis Martin† Kevin Stone†  
Peter Albert Amjad Almahairi Yasmine Babaei Nikolay Bashlykov Soumya Batra  
Prajwal Bhargava Shruti Bhosale Dan Bikel Lukas Blecher Cristian Canton Ferrer Moya Chen  
Guillem Cucurull David Esiobu Jude Fernandes Jeremy Fu Wenyin Fu Brian Fuller  
Cynthia Gao Vedanuj Goswami Naman Goyal Anthony Hartshorn Saghar Hosseini Rui Hou  
Hakan Inan Marcin Kardas Viktor Kerkez Madian Khabsa Isabel Kloumann Artem Korenev  
Punit Singh Koura Marie-Anne Lachaux Thibaut Lavril Jenya Lee Diana Liskovich  
Yinghai Lu Yuning Mao Xavier Martinet Todor Mihaylov Pushkar Mishra  
Igor Molybog Yixin Nie Andrew Poulton Jeremy Reizenstein Rashi Rungta Kalyan Saladi  
Alan Schelten Ruan Silva Eric Michael Smith Ranjan Subramanian Xiaoqing Ellen Tan Binh Tang  
Ross Taylor Adina Williams Jian Xiang Kuan Puxin Xu Zheng Yan Iliyan Zarov Yuchen Zhang  
Angela Fan Melanie Kambadur Sharan Narang Aurelien Rodriguez Robert Stojnic  
Sergey Edunov Thomas Scialom\*

## Llama 2와 Llama-2-Chat 학습과정



**Figure 4: Training of LLAMA 2-CHAT:** This process begins with the **pretraining** of LLAMA 2 using publicly available online sources. Following this, we create an initial version of LLAMA 2-CHAT through the application of **supervised fine-tuning**. Subsequently, the model is iteratively refined using Reinforcement Learning with Human Feedback (**RLHF**) methodologies, specifically through rejection sampling and Proximal Policy Optimization (PPO). Throughout the RLHF stage, the accumulation of **iterative reward modeling data** in parallel with model enhancements is crucial to ensure the reward models remain within distribution.

---

## Llama 2 논문의 구조

- **Abstract** : 전체 논문을 요약
- **Pretraing** : Llama 2 모델을 만들기 위한 Pre-training 과정에 대한 상세한 설명
- **Fine-Tuning** : Llama-2-Chat 모델을 만들기 위한 Fine-Tuning 과정에 대한 상세한 설명
- **Safety** : Llama-2-Chat 모델을 최대한 안정성있게 만들기 위해 적용한 기법들에 대한 상세한 설명
- **Discussion** : Llama 2와 Llama-2-Chat을 만들면서 발견한 사실들에 대한 상세한 설명
- **Conclusion** : 전체 논문을 다시한번 정리