ChatFeedback: In Search of Sustainable Synthetic Language Data with Human Feedback

Shanli Xing

Abstract

Recent advancements in the usability of large language models like GPT-4 have significantly boosted the adoption of natural language processing technologies. However, the demand for human preference and feedback data remains high, driven by alignment techniques such as Reinforcement Learning with Human Feedback. To address the challenges of data collection costs and representation issues, we propose ChatFeedback, an interactive system that incentivizes users to provide feedback on generated text, thereby collecting high-quality synthetic language data. ChatFeedback not only offers a scalable and user-friendly platform but also aims to democratize data collection, addressing ethical concerns about data transparency and inclusivity by promoting diverse user engagement. Our proof-of-concept demo code is available at `https://github.com/xslingcn/chat_feedback`.

1 Introduction

The field of natural language processing have witnessed a surge in adoption due to the significant advancements in the usability of large language models (LLMs), featured with GPT-4 (OpenAI et al. (2024)), Llama family (Touvron et al. (2023)), etc.. This is largely driven by alignment techniques like Reinforcement Learning with Human Feedback (RLHF) (Ouyang et al. (2022), Bai et al. (2022a)), which raises the demand for human preference and feedback data. However, collecting such data tends to be expensive and difficult to scale. Ethically, datasets created through crowdsourcing platforms or expert annotation also face representation issues, making it challenging to cover the actual preference distribution of a diverse user base in real-world settings and thus avoid potential biases. Additionally, despite open-source efforts (Köpf et al. (2023), Zheng et al. (2024)), the largest portion of data used to train state-of-the-art models remain proprietary. Therefore, there has been a huge gap between the data availability and the need.

On the other hand, work from Fan et al. (2023) and He et al. (2023) has shown that training with synthetic data can be effective, thus collecting synthetic data and associate it with human preference and feedback is viable for bridging the gap mentiond above.

In this project, we propose ChatFeedback, an interactive system where users can interact with LLMs as with any other chatbot service, and are in turn motivated to provide annotation to the generated text to collect high-quality, sustainable, synthetic language data with human feedback. This platform is delivered with a proof-of-concept demo with user-friendly interfaces, scalable deployment and data persistent storage solutions, and a detailed roadmap for further development.

2 Related Work

Training with synthetic data. The effectiveness of using synthetic image data for training is investigated and proven positive in many prior works (Fan et al. (2023), He et al. (2023) Taori et al. (2023)). Particularly in a natural language processing context, (Lee et al. (2023)) have shown that training with generated labels results in comparable or superior performance to RLHF in language tasks like summarization. Bai et al. (2022b) also combined LLM generated labels with human preference to reach strong performance. There are also effort in synthetic data collection like UltraFeedback

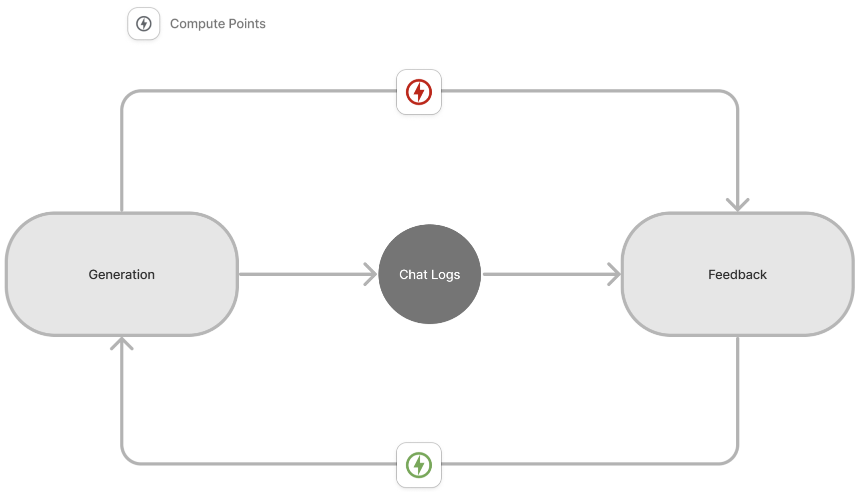


Figure 1: The minimal incentive model of ChatFeedback. In general, users consume *Compute Points* to chat with LLMs, and provide feedbacks to earn *Compute Points* to maintain the access to the LLMs.

(Cui et al. (2023)), which presents a large-scale preference dataset annotated by LLMs. Fan et al. (2023) also studies the scaling behaviors of using synthetic data, and shows that the performance does scale with the amount of synthetic data. Gerstgrasser et al. (2024) rather address the concern about performance collapse, and shows that accumulation of synthetic data may be robust to model collapse even in the maximally pessimistic setting, and suggests for using synthetic data to augment real-world data.

Community-based crowdsourcing data collection. Many works in aim to democritize the data collection process by crowdsourcing from the tech and AI community have emerged, such as Chatbot Arena (Chiang et al. (2024)), where users prompt two LLMs at the same time and vote for the better response, OpenAssistant (Köpf et al. (2023)), which encourages volunteering human annotators to manually create and evaluate dialogue-style conversations, and WildChat (Zhao et al. (2024)), which collects 1M in-the-wild chat logs by providing free access to OpenAI APIs to the community.

3 Methodology

In ChatFeedback, we define a global resource *Compute Points* and two default behaviors: 1). users will consume *Compute Points* for text generation, and 2). users will be rewarded with *Compute Points* for providing any preference and feedback. As shown in Figure 2, we add the following mechanisms:

Chat with LLMs. ChatFeedback proxies access to LLMs and provide a chat interface where users can interact with them as with any other chatbot service. Two service mode will be available: 1). *Dynamic Chat*, which costs fewer *Compute Points*, and users interact with an LLM dynamically assigned by the platform. 2). *Direct Chat*, which costs more *Compute Points*, and users can choose a specific LLM to interact with. In both modes, users can optionally provide preference to each response. Users always have the option to re-generate with another LLM on the same cost. In dynamic mode, if user gives low score for one response, they'll be presented with two responses generated by another LLM and choose which is prefered side-by-side for the next generation. Chat logs are by default saved and distributed to other users to review and provide feedback. Users

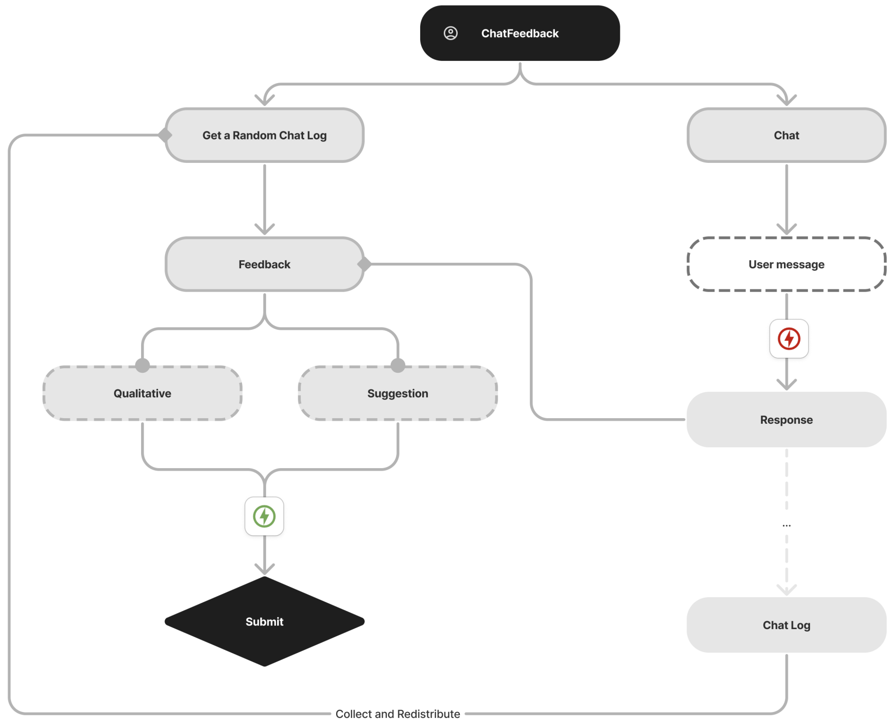


Figure 2: A detailed workflow of how users may start using ChatFeedback. They can start with giving feedback on a random response, just for accumulating *Compute Points*. They can also start a chat with a LLM, and optionally provide feedback on the responses they receive, during or after the chat.

also have the option to keep the chat logs private, but they'll be rewarded with *Compute Points* for sharing.

Provide feedback. Users can optionally provide feedbacks on their own chat logs while and after chatting. They can also get randomly assigned with a chat log shared by other users and feedback on a specific response, which can be generated or manually created as a *Suggestion*. Two types of feedback are collected. 1). *Qualitative*. Inspired by UltraFeedback (Cui et al. (2023)), it allows users to tag the response by the following qualities: *Instruction*: whether this response follow the instruction provided, *helpful*: whether this response provide useful and correct answers to successfully address the given problems, *factual*: whether this response is grounded in the instructions and real-world knowledge, *style*: whether the user likes the style of this response, *sensitive* and *toxic*: whether response includes sensitive or harmful content. For each quality, users can choose from *N/A*, *No*, *Yes*. 2). *Suggestion*, where users can provide a preferred rewritten version of a certain response. Depending on the difficulty of the tasks, users will receive varying amounts of *Compute Points*.

Gamification. To better motivate users, we introduce gamification elements like *Daily Streak* and *Leaderboard*. Note that users don't have to submit feedback while chatting with LLMs, instead they can provide feedback in their spare time as a time-killing task just to accumulate *Compute Points*.

Dynamic rewards. With the intuition that not all responses are equally notable, and the qualitative tags may not be meaningful to all responses (for example, the response "Hello, how can we help you" doesn't constitute any factual quality), we want to dynamically adjust the reward based on the importance of the feedback. We want the model to also return a "seed feedback" along with the response, and classify the feedback significance into low, medium to high. The significance will be used to adjust the reward, and will also be displayed to the user to help them determine what responses to feedback on.

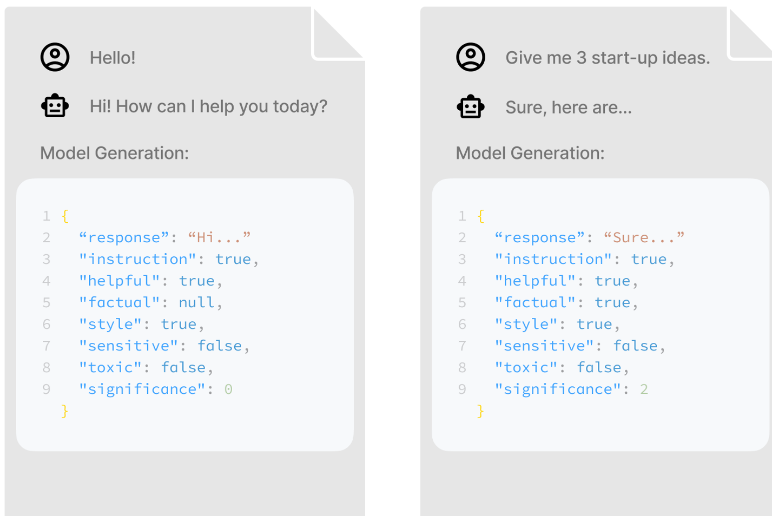


Figure 3: Example of expected responses from LLMs. We use regular expressions to make sure the format is correct, then extract the response for the user and use the rest to seed the feedback for consistency check, and estimate the significance to dynamically determine the reward.

Confidence score. To ensure the quality of the feedback, ChatFeedback will calculate a confidence score for each user based on the consistency of their feedback with the majority, and adjust the reward accordingly. This adjustment will only be significant for users who provide a large amount of divergent feedback.

Demographic collection. To better understand the distribution of users, the diversity of feedback, and how preferences vary by demographics, we propose optionally collecting users' demographic information. This information should be anonymized, used for overall analysis only and should not be published in any form individually.

Computing budget. Many of the aforementioned crowdsourced data collection works operate entirely on academic fundings and sponsorships. For example, Chatbot Arena (Chiang et al. (2024)) has operated for five months and consumed "thousands of A100 hours." upon the release of LMSYS-Chat-1M (Zheng et al. (2024)). According to the pricing of mainstream cloud service providers, the total cost should be roughly within $10,000. By using pay-as-you-go inference platforms like *OctoAI*[1], costs can be further reduced. Therefore, we expect that the computing budget for collecting data on a comparable scale will be in the order of a few thousand dollars. The cost for a small-scale pilot run as a concept verification can be kept under one thousand dollars.

[1] https://octo.ai/

4 Implementation

As the project deliverable, a proof-of-concept demo is implemented with the following features completed:

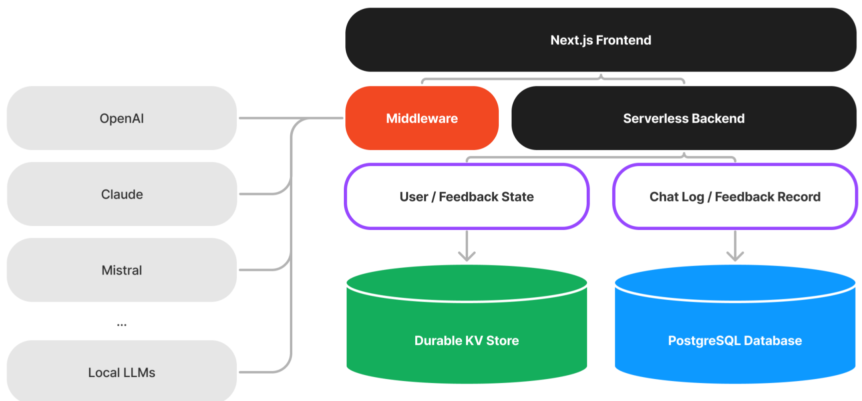
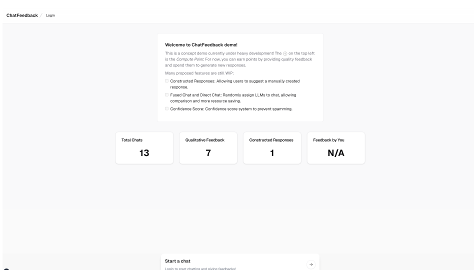
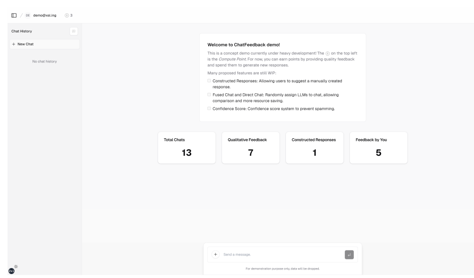


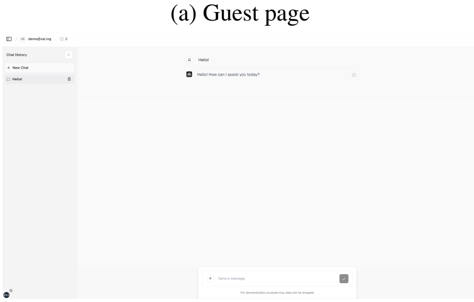
Figure 4: Overview of the system structure of the implemented POC demo.



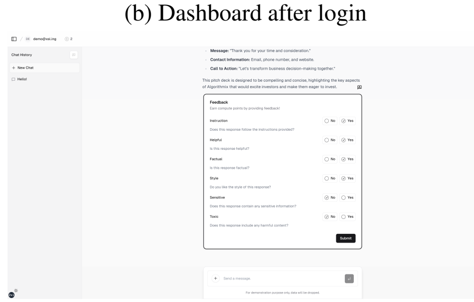
(a) Guest page



(b) Dashboard after login



(c) Chat



(d) Feedback

Figure 5: Screenshots of the POC demo. Note that the *Compute Points* is reduced by 1 after the first chat. The user entered a random chat log in (d), and will earn 1 point after submitting the feedback.

Chat interface. This is a full-stack web application with an user-friendly interface, built with React.js and Next.js, and styled using TailwindCSS. The app is designed to be serverless, therefore can be easily deployed on serverless hosting platforms like *Vercel*[2]. It has currently adopted only the OpenAI api but can be easily integrated with other LLM services.

Incentive model, i.e. the *Compute point* earning and consumption mechanism. The dynamic chat and rewards mechanism is not yet implemented, therefore currently users consume 1 *Compute point* for each generation and earn 1 *Compute point* for each feedback. While the gamification elements are not fully completed, we provide a dashboard stats component where users can see the progress of themselves as well as the entire platform.

Feedback and chatlog distribution. Users can currently submit a qualitative feedback for each respone as described above. It's also supported to randomly enter a chat log and feedback on each response.

Data storage and retrieval. The chat logs and feedbacks are persistently preserved with serverless cloud storage providers. This makes the data operations fast, scalable and secure. The chat logs are stored with a tree structure, where we identify the first prompt of a chat log as the `root_chat`, and keep the `parent_id` for all non-root chats.

5 Ethics

5.1 Caveat

Because of the collection of sensitive data like chat logs and demographic information, this platform will not be hosted for public use after implementation until being carefully and comprehensively advised on ethical aspects. Data wukk not be disclosed until it has been thoroughly desensitized and detoxified.

5.2 Contribution

This work is ethically significant as it promotes the transparency and sustainability of human preference data, which is often proprietary, thus barrier the public building of performant and helpful LLMs.

Moreover, since previous data collection procedures usually only have annotators from a specific community to participate (e.g. academic researchers, tech enthusiasts, or part-time workers on crowdsourcing platforms), the data collected may not be representative of the actual user base. Many previous works (Durmus et al. (2023), Santurkar et al. (2023)) have shown current LLMs bias towards particular demographic groups, and this issue has to be addressd to prevent any harm that could be done by LLMs. By providing free access to commonly used LLMs with a generic chatbot interface, ChatFeedback is more likely to attract a diverse and real-world user base, therefore collect a more diverse dataset and address the representation issues in existing alike datasets.

Additionally, there has been many debates ongoing about what are the ”proper” or ”ethical” behaviors of LLMs, and the answer is usually varied by many factors. By analyzing demographic information collected along with preferences, we can better understand how preferences for LLMs vary by demographics and study how to build a trustworthy and inclusive LLM system for all. This is can help raise critical ethical discussion and provide valuable ground-truth opinion for such future research.

Lastly, despite the fact that LLMs are widely adopted and proven helpful for many daily tasks, there are still many places where the access to LLMs is limited, or the cost is not affordable. By providing a free and easy-to-use platform for users to interact with LLMs, ChatFeedback can help more people get equal access to the power as well as the benifits of cutting-edge technology innovations.

References

Yuntao Bai, Andy Jones, Kamal Ndousse, Amanda Askell, Anna Chen, Nova DasSarma, Dawn Drain, Stanislav Fort, Deep Ganguli, Tom Henighan, Nicholas Joseph, Saurav Kadavath, Jackson Kernion, Tom Conerly, Sheer El-Showk, Nelson Elhage, Zac Hatfield-Dodds, Danny Hernandez,

[2]`https://vercel.com/`

Tristan Hume, Scott Johnston, Shauna Kravec, Liane Lovitt, Neel Nanda, Catherine Olsson, Dario Amodei, Tom Brown, Jack Clark, Sam McCandlish, Chris Olah, Ben Mann, and Jared Kaplan. Training a Helpful and Harmless Assistant with Reinforcement Learning from Human Feedback, April 2022a.

Yuntao Bai, Saurav Kadavath, Sandipan Kundu, Amanda Askell, Jackson Kernion, Andy Jones, Anna Chen, Anna Goldie, Azalia Mirhoseini, Cameron McKinnon, Carol Chen, Catherine Olsson, Christopher Olah, Danny Hernandez, Dawn Drain, Deep Ganguli, Dustin Li, Eli Tran-Johnson, Ethan Perez, Jamie Kerr, Jared Mueller, Jeffrey Ladish, Joshua Landau, Kamal Ndousse, Kamile Lukosuite, Liane Lovitt, Michael Sellitto, Nelson Elhage, Nicholas Schiefer, Noemi Mercado, Nova DasSarma, Robert Lasenby, Robin Larson, Sam Ringer, Scott Johnston, Shauna Kravec, Sheer El Showk, Stanislav Fort, Tamera Lanham, Timothy Telleen-Lawton, Tom Conerly, Tom Henighan, Tristan Hume, Samuel R. Bowman, Zac Hatfield-Dodds, Ben Mann, Dario Amodei, Nicholas Joseph, Sam McCandlish, Tom Brown, and Jared Kaplan. Constitutional AI: Harmlessness from AI Feedback, December 2022b.

Wei-Lin Chiang, Lianmin Zheng, Ying Sheng, Anastasios Nikolas Angelopoulos, Tianle Li, Dacheng Li, Hao Zhang, Banghua Zhu, Michael Jordan, Joseph E. Gonzalez, and Ion Stoica. Chatbot Arena: An Open Platform for Evaluating LLMs by Human Preference, March 2024.

Ganqu Cui, Lifan Yuan, Ning Ding, Guanming Yao, Wei Zhu, Yuan Ni, Guotong Xie, Zhiyuan Liu, and Maosong Sun. UltraFeedback: Boosting Language Models with High-quality Feedback, October 2023.

Esin Durmus, Karina Nyugen, Thomas I. Liao, Nicholas Schiefer, Amanda Askell, Anton Bakhtin, Carol Chen, Zac Hatfield-Dodds, Danny Hernandez, Nicholas Joseph, Liane Lovitt, Sam McCandlish, Orowa Sikder, Alex Tamkin, Janel Thamkul, Jared Kaplan, Jack Clark, and Deep Ganguli. Towards Measuring the Representation of Subjective Global Opinions in Language Models, June 2023.

Lijie Fan, Kaifeng Chen, Dilip Krishnan, Dina Katabi, Phillip Isola, and Yonglong Tian. Scaling Laws of Synthetic Images for Model Training ... for Now, December 2023.

Matthias Gerstgrasser, Rylan Schaeffer, Apratim Dey, Rafael Rafailov, Henry Sleight, John Hughes, Tomasz Korbak, Rajashree Agrawal, Dhruv Pai, Andrey Gromov, Daniel A. Roberts, Diyi Yang, David L. Donoho, and Sanmi Koyejo. Is Model Collapse Inevitable? Breaking the Curse of Recursion by Accumulating Real and Synthetic Data, April 2024.

Ruifei He, Shuyang Sun, Xin Yu, Chuhui Xue, Wenqing Zhang, Philip Torr, Song Bai, and Xiaojuan Qi. Is synthetic data from generative models ready for image recognition?, February 2023.

Andreas Köpf, Yannic Kilcher, Dimitri von Rütte, Sotiris Anagnostidis, Zhi-Rui Tam, Keith Stevens, Abdullah Barhoum, Nguyen Minh Duc, Oliver Stanley, Richárd Nagyfi, Shahul ES, Sameer Suri, David Glushkov, Arnav Dantuluri, Andrew Maguire, Christoph Schuhmann, Huu Nguyen, and Alexander Mattick. OpenAssistant Conversations – Democratizing Large Language Model Alignment, October 2023.

Harrison Lee, Samrat Phatale, Hassan Mansoor, Thomas Mesnard, Johan Ferret, Kellie Lu, Colton Bishop, Ethan Hall, Victor Carbune, Abhinav Rastogi, and Sushant Prakash. RLAIF: Scaling Reinforcement Learning from Human Feedback with AI Feedback, November 2023.

OpenAI, Josh Achiam, Steven Adler, Sandhini Agarwal, Lama Ahmad, Ilge Akkaya, Florencia Leoni Aleman, Diogo Almeida, Janko Altenschmidt, Sam Altman, Shyamal Anadkat, Red Avila, Igor Babuschkin, Suchir Balaji, Valerie Balcom, Paul Baltescu, Haiming Bao, Mohammad Bavarian, Jeff Belgum, Irwan Bello, Jake Berdine, Gabriel Bernadett-Shapiro, Christopher Berner, Lenny Bogdonoff, Oleg Boiko, Madelaine Boyd, Anna-Luisa Brakman, Greg Brockman, Tim Brooks, Miles Brundage, Kevin Button, Trevor Cai, Rosie Campbell, Andrew Cann, Brittany Carey, Chelsea Carlson, Rory Carmichael, Brooke Chan, Che Chang, Fotis Chantzis, Derek Chen, Sully Chen, Ruby Chen, Jason Chen, Mark Chen, Ben Chess, Chester Cho, Casey Chu, Hyung Won Chung, Dave Cummings, Jeremiah Currier, Yunxing Dai, Cory Decareaux, Thomas Degry, Noah Deutsch, Damien Deville, Arka Dhar, David Dohan, Steve Dowling, Sheila Dunning, Adrien Ecoffet, Atty Eleti, Tyna Eloundou, David Farhi, Liam Fedus, Niko Felix, Simón Posada Fishman, Juston Forte, Isabella Fulford, Leo Gao, Elie Georges, Christian Gibson, Vik Goel, Tarun Gogineni, Gabriel Goh, Rapha Gontijo-Lopes, Jonathan Gordon, Morgan Grafstein, Scott Gray, Ryan Greene, Joshua Gross, Shixiang Shane Gu, Yufei Guo, Chris Hallacy, Jesse Han, Jeff Harris, Yuchen He, Mike Heaton, Johannes Heidecke, Chris Hesse, Alan Hickey, Wade Hickey, Peter Hoeschele, Brandon Houghton, Kenny Hsu, Shengli Hu, Xin Hu, Joost Huizinga, Shantanu Jain, Shawn Jain, Joanne Jang, Angela Jiang, Roger Jiang, Haozhun Jin, Denny Jin, Shino Jomoto, Billie Jonn, Heewoo Jun, Tomer Kaftan, Łukasz Kaiser, Ali Kamali, Ingmar Kanitscheider, Nitish Shirish Keskar, Tabarak Khan, Logan Kilpatrick, Jong Wook Kim, Christina Kim, Yongjik Kim, Jan Hendrik Kirchner, Jamie Kiros, Matt Knight, Daniel Kokotajlo, Łukasz Kondraciuk, Andrew Kondrich, Aris Konstantinidis, Kyle Kosic, Gretchen Krueger, Vishal Kuo, Michael Lampe, Ikai Lan, Teddy Lee, Jan Leike, Jade Leung, Daniel Levy, Chak Ming Li, Rachel Lim, Molly Lin, Stephanie Lin, Mateusz Litwin, Theresa Lopez, Ryan Lowe, Patricia Lue, Anna Makanju, Kim Malfacini, Sam Manning, Todor Markov, Yaniv Markovski, Bianca Martin, Katie Mayer, Andrew Mayne, Bob McGrew, Scott Mayer McKinney, Christine McLeavey, Paul McMillan, Jake McNeil, David Medina, Aalok Mehta, Jacob Menick, Luke Metz, Andrey Mishchenko, Pamela Mishkin, Vinnie Monaco, Evan Morikawa, Daniel Mossing, Tong Mu, Mira Murati, Oleg Murk, David Mély, Ashvin Nair, Reiichiro Nakano, Rajeev Nayak, Arvind Neelakantan, Richard Ngo, Hyeonwoo Noh, Long Ouyang, Cullen O'Keefe, Jakub Pachocki, Alex Paino, Joe Palermo, Ashley Pantuliano, Giambattista Parascandolo, Joel Parish, Emy Parparita, Alex Passos, Mikhail Pavlov, Andrew Peng, Adam Perelman, Filipe de Avila Belbute Peres, Michael Petrov, Henrique Ponde de Oliveira Pinto, Michael, Pokorny, Michelle Pokrass, Vitchyr H. Pong, Tolly Powell, Alethea Power, Boris Power, Elizabeth Proehl, Raul Puri, Alec Radford, Jack Rae, Aditya Ramesh, Cameron Raymond, Francis Real, Kendra Rimbach, Carl Ross, Bob Rotsted, Henri Roussez, Nick Ryder, Mario Saltarelli, Ted Sanders, Shibani Santurkar, Girish Sastry, Heather Schmidt, David Schnurr, John Schulman, Daniel Selsam, Kyla Sheppard, Toki Sherbakov, Jessica Shieh, Sarah Shoker, Pranav Shyam, Szymon Sidor, Eric Sigler, Maddie Simens, Jordan Sitkin, Katarina Slama, Ian Sohl, Benjamin Sokolowsky, Yang Song, Natalie Staudacher, Felipe Petroski Such, Natalie Summers, Ilya Sutskever, Jie Tang, Nikolas Tezak, Madeleine B. Thompson, Phil Tillet, Amin Tootoonchian, Elizabeth Tseng, Preston Tuggle, Nick Turley, Jerry Tworek, Juan Felipe Cerón Uribe, Andrea Vallone, Arun Vijayvergiya, Chelsea Voss, Carroll Wainwright, Justin Jay Wang, Alvin Wang, Ben Wang, Jonathan Ward, Jason Wei, C. J. Weinmann, Akila Welihinda, Peter Welinder, Jiayi Weng, Lilian Weng, Matt Wiethoff, Dave Willner, Clemens Winter, Samuel Wolrich, Hannah Wong, Lauren Workman, Sherwin Wu, Jeff Wu, Michael Wu, Kai Xiao, Tao Xu, Sarah Yoo, Kevin Yu, Qiming Yuan, Wojciech Zaremba, Rowan Zellers, Chong Zhang, Marvin Zhang, Shengjia Zhao, Tianhao Zheng, Juntang Zhuang, William Zhuk, and Barret Zoph. GPT-4 Technical Report, March 2024.

Long Ouyang, Jeff Wu, Xu Jiang, Diogo Almeida, Carroll L. Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, John Schulman, Jacob Hilton, Fraser Kelton, Luke Miller, Maddie Simens, Amanda Askell, Peter Welinder, Paul Christiano, Jan Leike, and Ryan Lowe. Training language models to follow instructions with human feedback, March 2022.

Shibani Santurkar, Esin Durmus, Faisal Ladhak, Cinoo Lee, Percy Liang, and Tatsunori Hashimoto. Whose Opinions Do Language Models Reflect?, March 2023.

Rohan Taori, Ishaan Gulrajani, Yann Dubois, Xuechen Li, Percy Liang, and Tatsunori B Hashimoto. Alpaca: A Strong, Replicable Instruction- Following Model. *Stanford Center for Research on Foundation Models*, March 2023.

Hugo Touvron, Louis Martin, Kevin Stone, Peter Albert, Amjad Almahairi, Yasmine Babaei, Nikolay Bashlykov, Soumya Batra, Prajjwal 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, and Thomas Scialom. Llama 2: Open Foundation and Fine-Tuned Chat Models, July 2023.

Wenting Zhao, Xiang Ren, Jack Hessel, Claire Cardie, Yejin Choi, and Yuntian Deng. WildChat: 1M ChatGPT Interaction Logs in the Wild, May 2024.

Lianmin Zheng, Wei-Lin Chiang, Ying Sheng, Tianle Li, Siyuan Zhuang, Zhanghao Wu, Yonghao Zhuang, Zhuohan Li, Zi Lin, Eric P. Xing, Joseph E. Gonzalez, Ion Stoica, and Hao Zhang. LMSYS-Chat-1M: A Large-Scale Real-World LLM Conversation Dataset, March 2024.