Blah

Model description: This is a large language model (LLM) trained using conversational language mostly found on the web. The model is designed to interact with users in a conversational way by predicting a response to the prompt provided by a user.

Intended use: It should be used to provide answers to problems and aid creativity.

Limitations: The data used to train this model is from various sources, which are mostly from the web, including blogs, forums, and community encyclopaedias. As a result, the output of the model might include harmful, offensive, inaccurate, or biased content.

Data set:

- **40% Wikipedia**
- **50% Common crawl**
- **10% UK newspapers (2010–2015)**