

With our flagship DataInFormationSM suite of solutions for Image Labeling, Data Annotation, NLP Calibration and Form Data Capture, Liberty Source is proud to serve a wide range of clients across many industry verticals. Our Solution Briefs provide a snapshot of the challenges clients bring to us, and the outcomes we achieve for them.

Prompt & Response QA for LLMs

Industry
Prompt Engineering

Data Type
Text

Project Duration
~6 Months

Ongoing?
Yes

Challenge

The client's LLMs had been providing an unacceptable number of responses that were unusable due to hallucinations (errors) and biases. Additionally, the client wanted to use their staff resources to train and tune their LLMs instead of writing prompts and benchmarking responses – for which they determined outside experts would be more efficient and effective.

Solution

Our quality assurance (QA) work centered on creating a benchmark dataset for the training and tuning of our client's LLMs. This was done through our human-generated (Human-in-the-Loop; HitL) responses, which were based on our own HitL-generated prompts.

Given a list of domain-specific topics (e.g. Data Science, History, Accounting, and more), which were provided by the client, our writers developed prompts and responses using their extensive domain knowledge and training.

Outcome

Our deliverable consisted of our writers providing factual information in the form of prompts that were written in the appropriate syntax and style required by each domain.

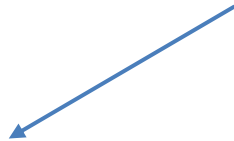
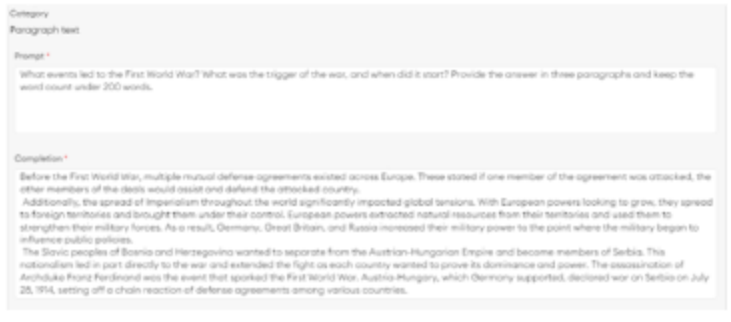
Additionally, hundreds of custom HitL-generated prompts and responses, across 10 topics, were categorized together under Summarization, Generative, Closed Question, and Extraction so the client could easily utilize them for verifying, training, and tuning the performance of their LLMs through a validation process.

This ultimately mitigated their exposure to receiving hallucinations or useless, biased responses and thereby provided the client with improved prompt response, enhanced model performance, and an overall enhanced user experience (UX).

Input: Prompt topics provided by client



Process: Human-in-the-Loop creates new prompts and responses based on guidelines provided by client



Output: Human-Generated responses are then used to benchmark client LLMs

