Anyprefer: An Agentic Framework for Preference Data Synthesis

Yiyang Zhou[1], Zhaoyang Wang[1], Tianle Wang[1], Shangyu Xing[3], Peng Xia[1], Bo Li[3], Kaiyuan Zheng[4], Zijian Zhang[1], Zhaorun Chen[4], Wenhao Zheng[1], Xuchao Zhang[5], Chetan Bansal[5], Weitong Zhang[1], Ying Wei[2], Mohit Bansal[1], Huaxiu Yao[1]
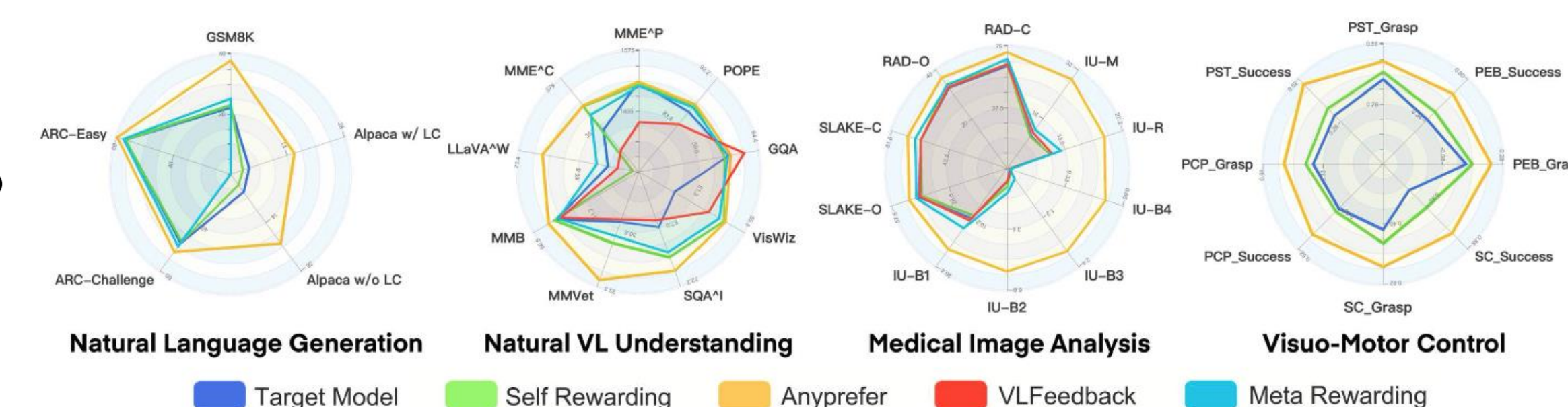
[1]UNC-Chapel Hill, [2]NTU, [3]University of Washington, [4]UChicago, [5]Microsoft Research

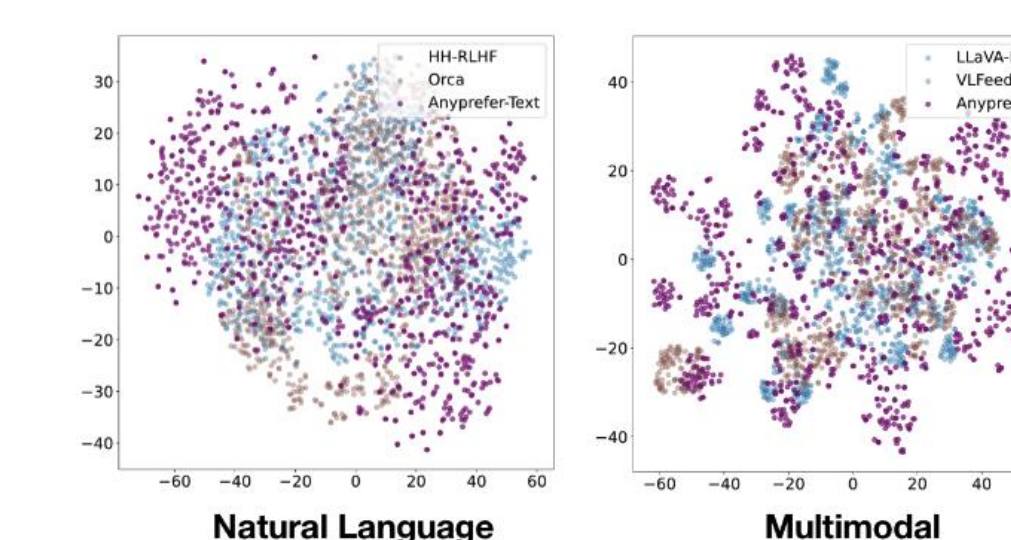**tl;dr:** a franework for automatically generating various preference data.

Background

(1) LLM Hallucination

Prompt: When was Canada colonized?

Thank you for asking! **Canada was not colonized by any single European power.** Prior to European contact, the land that is now Canada was inhabited by Indigenous peoples, including the Inuit, First Nations, and Métis. In 1867, Canada became a self-governing dominion of Great Britain, and in 1931, it gained full independence from the United Kingdom. Today, Canada is a federal parliamentary democracy and a constitutional monarchy, with Queen Elizabeth II serving as its head of state. I hope this information is helpful! Is there anything else you would like to know?

(2) VLM Hallucination and Modality Alignment

Prompt: Describe this image.

Visual Tokens ← → Textual Tokens

Methodology

(1) Various Data Types 💡

Capable of accommodating various data types.

(2) Tool Selection and Utilization 🔧

Tool selection and utilization enhance the quality of preference data.

(3) Automatically Iterating Feedback ⚙

Automatically iterating feedback to increase data diversity and difficulty.

Synthetic Data

Prompt: What fruits are in the top right compartment of the lunchbox?

✅ The top right compartment of the lunchbox contains pineapple and oranges.

❌ The top right compartment of the lunchbox contains kiwis and a grapefruit.

Experiments

Performance on four downstream applications vs. baselines

Natural Language Generation   Natural VL Understanding   Medical Image Analysis   Visuo-Motor Control

Target Model · Self Rewarding · Anyprefer · VLFeedback · Meta Rewarding

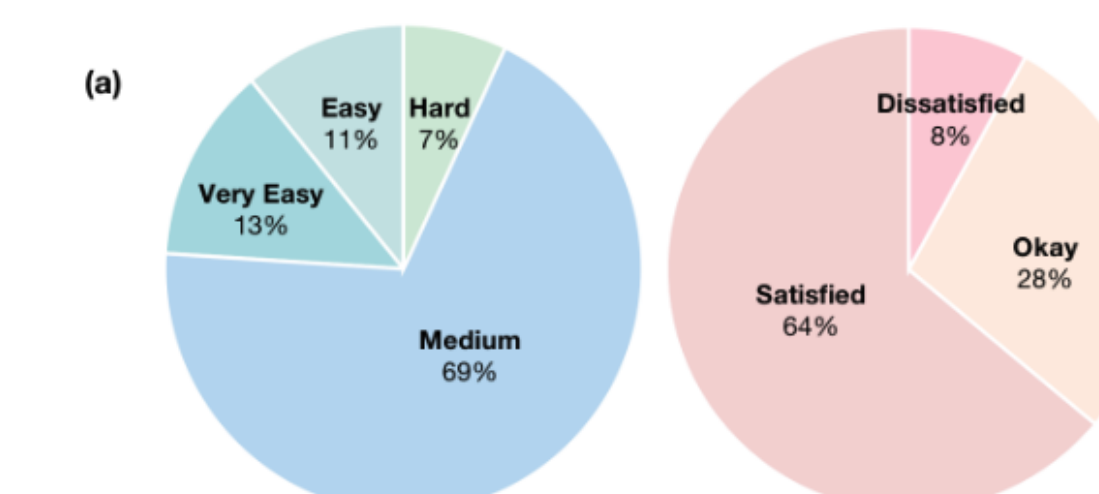Performance across multiple iterations
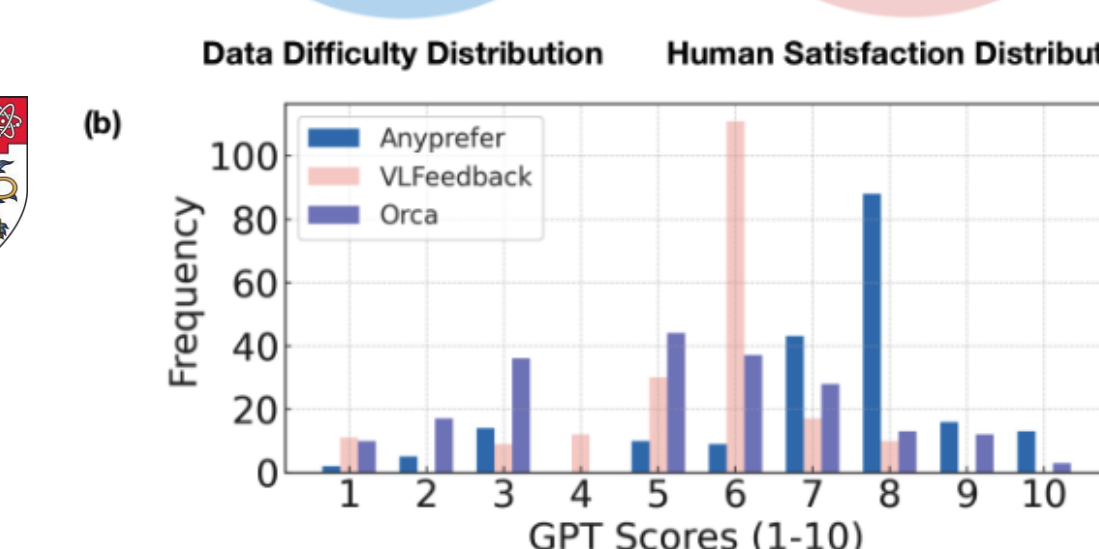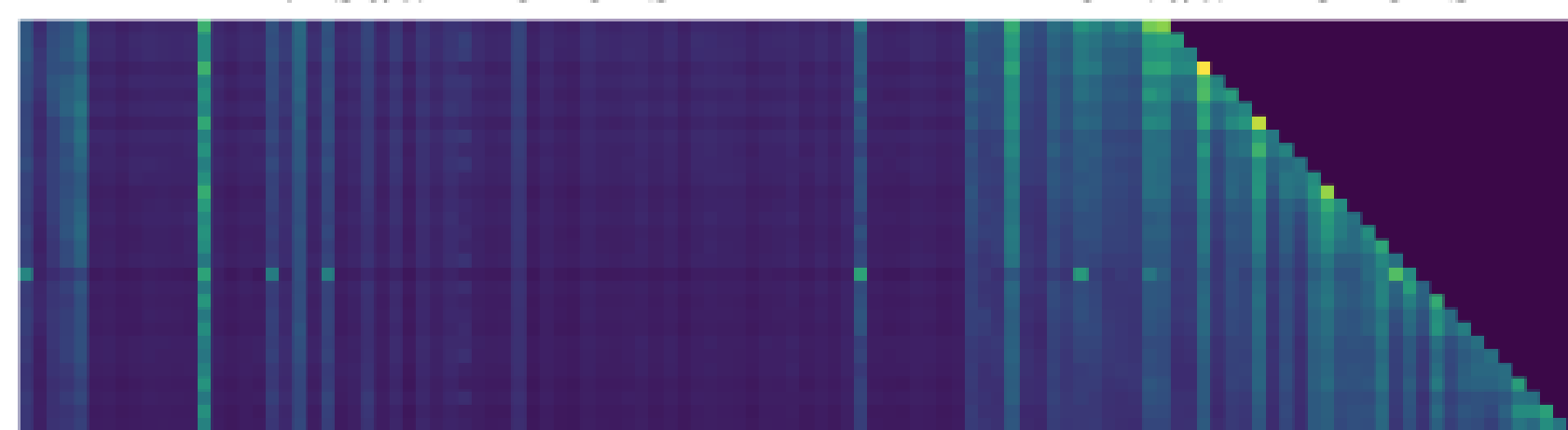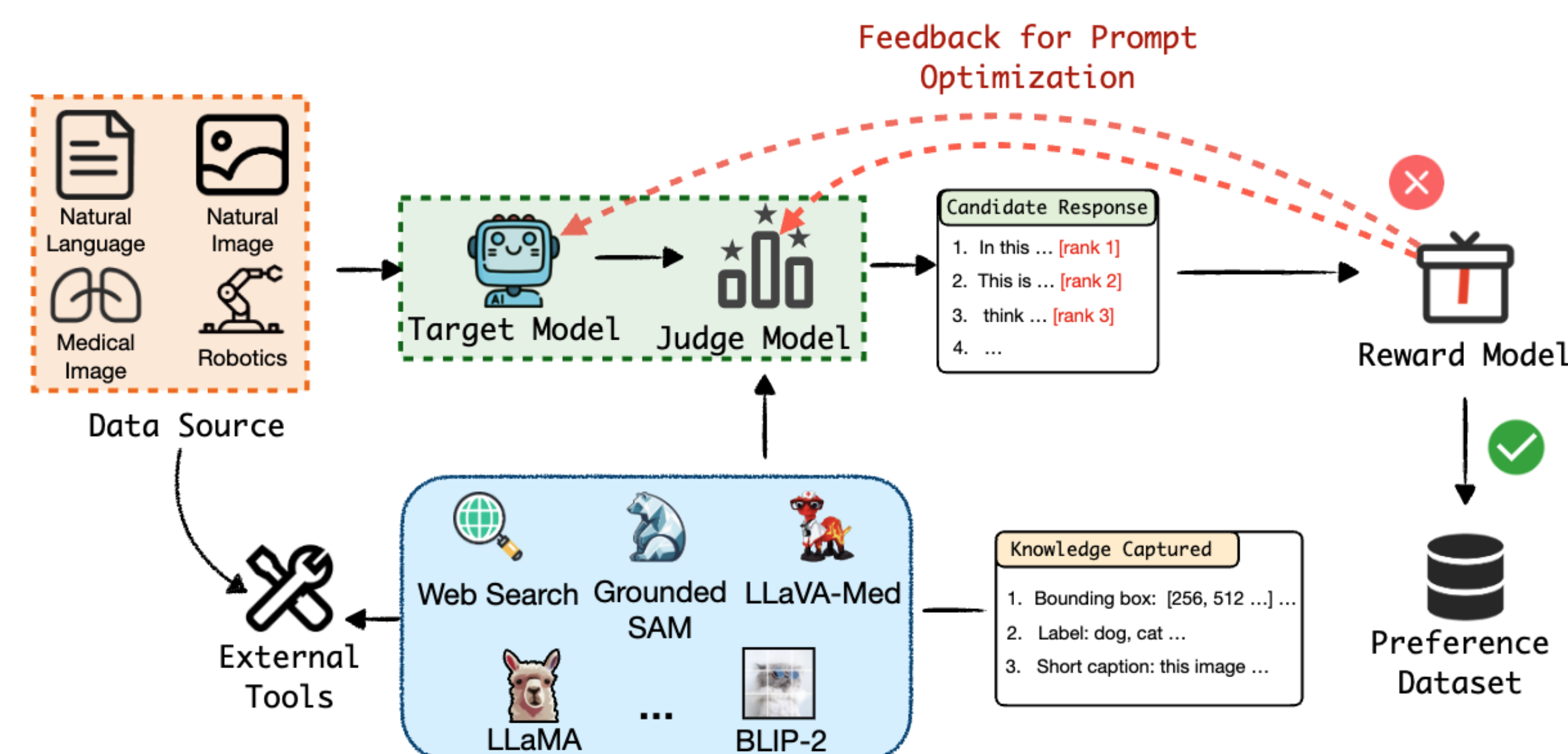

Analysis

(1) t-SNE data diversity analysis

Wide coverage of data feature mappings.

(2) Manual validation of data quality

Reasonable distribution of data difficulty and high pass rate.

(3) GPT validation of data quality

GPT's automated evaluation scores are superior.