Improving retrieval

To improve the performance of our system, we chose, as suggested, to use ChatGPT to reorganize the responses returned by our system. Although initially, we attempted to use the OpenAI API because we were limited in the number of requests we could make, we ended up doing everything manually.

Thus, we passed through all the questions using a prompt as follows:

"Giving a CATEGORY and a CLUE, we have managed to extract answers as Wikipedia pages. For Category "CATEGORY NAME," in order to best match the clue "CLUE," please reorder the following answers:

Answer 1, Answer 2, Answer 3, Answer 4, Answer 5, Answer 6, Answer 7, Answer 8"

The result was very satisfactory. Categories that initially had less than 50% accuracy some with 0% accuracy (such as STATE OF THE ART MUSEUM, '80s NO.1 HITMAKERS), significantly improved their accuracy. In most cases if the correct answer was among those provided by the initial system ended up being placed in the first position.

Example

USER : "Giving a CATEGORY and a CLUE we have managed to extract answers as Wikipedia pages. For Catergoy "HISTORICAL QUOTES" in order to match best the clue "In a 1959 American kitchen exhibit in Moscow, he told Khrushchev, "In America, we like to make life easier for women"" please reorder the following answers :

Cuban missile crisis, Stanley Kramer, Marc Chagall, Golda Meir, Aleksandr Vasilevsky, Leonid Brezhnev, Gloria Steinem, Stanley Kubrick"

ChatGPT:

"Richard Nixon, Cuban missile crisis, Stanley Kramer, Marc Chagall, Golda Meir, Aleksandr Vasilevsky, Leonid Brezhnev, Gloria Steinem, Stanley Kubrick"

However, there were also many examples where, even though the correct answer was not among those identified by the initial system, the response received from ChatGPT contained it in the first position.

Comparison between the initial system and the improved one

	Correct Answers	Incorrect Answers	Precision	Precision at 1
Initial System	50	50	50.0%	0.34
Improved system	73	27	73.0%	0.63