# HybridCRS Questionnaire

Evaluation on the Usability and Performance of HybridCRS. There will be three use cases you will need to test out, after which you will evaluate the recommendations and explanations generated in the application. Finally, you will fill in a SUS questionnaire and provide feedback about the application. The data collected will be used for my Master's Thesis titled "A Hybrid Conversational Recommender System by integrating LLMs".

**Web Application Link**.

**Introduction**

The use of chatbots and Large Language Models (LLMs) have surged in recent years, and conversational recommender systems/agents are a novel application of such technologies. These agents have been integrated into popular platforms such as Amazon (e.g., Rufus) and are the focus of active research, providing strong motivation for this project.

I am the developer of the application you will be testing, which offers a streamlined and generalized approach to the creation of recommendation agents. Users may also start a conversation with the agents to test the generation of recommendations (and explanations to each recommendation).

The agents of the application are event-driven, LLM-powered chatbots created from recommendation datasets, which are tabular data that describe interactions and features among users and items. The agents use this data to train a model that can recommend items based on a set of user preferences. Conversations with these agents are the means through which they can gather such preferences.

In the next sections, you will be carrying out tasks to test out the main functionalities of the application, with the objective of evaluating the usability of the application, the relevance of the recommendations and the impact of the explanations:

1. To create recommendation agents.
2. To have conversations with the created agent to gather user preferences.
3. To generate recommendations from said preferences, as well as explanations to each recommendation.

* Indicates required question

1. **Data Collection Agreement** *

   I hereby consent to the use of the data collected from this form for research purposes to the benefit of the author's Master's Thesis.

   *Check all that apply.*

   ☐ I agree

2. **Name** *

_____

3. **Email Address** *

First of all, sign up to HybridCRS. Here, write the email you used to sign up.

_____

4. **Age** *

_____

5. **Gender** *

*Mark only one oval.*

○ Male

○ Female

○ Non-binary

○ Prefer not to say

6. **Do you have previous experience with Large Language Models (LLMs)?** *

*Mark only one oval.*

○ Yes

○ No

7. **Do you have previous experience with chatbots?** *

*Mark only one oval.*

○ Yes

○ No

**Use Case 1**: Creating and Editing a Recommendation Agent

For this first task, you will be creating a recommendation agent using one of three datasets given to you, and then editing the created agent from the Agent Hub.

**The available datasets are**:

- **MovieLens-100K** (movies)**:** [Download](#)
- **Anime** (anime series): [Download](#)

**Some info on these datasets**:

- For each dataset, "**.inter**" files contain user-item interactions (such as ratings), "**.user**" files contain user features and "**.item**" files contain item features. Interaction files are **mandatory** since they are used for model training and generation of recommendations.
- Each column can be suffixed as one of 4 types: "**token**" for discrete values, "**token_seq**" for sequence of discrete values, "**float**" for numerical values and "**float_seq**" for sequence of numerical values.
- If you want, you could use your own dataset. However, there must be at least a tabular file with interactions between users and items.

**Steps**:

1. First, download any of the above datasets and extract the compressed file.
2. If you haven't already, create an account in HybridCRS and confirm your email. Log into the application.
3. After logging in you will be redirected to the Agent Hub. Click on Create Agent on the bottom right.
4. Fill in the agent configuration form. You may make it public if you wish. Then, proceed to the data configuration.
5. In the file input, upload the extracted files from the dataset and select the types.
6. For each of the uploaded files, edit each column on the table to match the actual role and type they correspond to. Then, proceed to the creation review.
7. Review the agent you're about to create, and proceed.
8. When the agent is created, edit it however you want, verifying the information is updated on the Agent Hub.

8.   How difficult was it to complete this task? *

*Mark only one oval.*

|  | 1 | 2 | 3 | 4 | 5 |  |
|---|---|---|---|---|---|---|
| Very Easy | ⬭ | ⬭ | ⬭ | ⬭ | ⬭ | Very Difficult |

9. Would you use this functionality if you needed to make a conversational recommender agent *
with some dataset?

*Mark only one oval.*

○ Yes

○ No

## Use Case 2: First Conversation with Recommendation Agent

For this task, you will be having a conversation with the agent you just created. You will need to give
your preferences to the agent and obtain recommendations from them.

**Steps**:

1. Go to the Agent Hub.
2. Find the agent you created in the previous step, and click on "Start Conversation".
3. Click on "Begin Conversation".
4. Chat with the agent, responding to their questions until obtaining recommendations.
5. Rate each recommendation and send the feedback.
6. End the session whenever you want.
7. Check the archived session in the "Chat Sessions" menu (Optional).

10. How difficult was it to complete this task? *

*Mark only one oval.*

|  | 1 | 2 | 3 | 4 | 5 |  |
|---|---|---|---|---|---|---|
| Very Easy | ○ | ○ | ○ | ○ | ○ | Very Difficult |

## Evaluation of Recommendations (Use Case 2)

Evaluation of the recommendations given by the agents during the first conversation.

11. Accuracy *

Were the recommendations relevant to you, did you like them?

|   | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|----|
|   | ☆ | ☆ | ☆ | ☆ | ☆ | ☆ | ☆ | ☆ | ☆ | ☆ |

12. Contextual Relevance *

Were the recommendations relevant to the contextual preferences that you gave to the agent (category, genre, brand, etc.)?

|   | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|----|
|   | ☆ | ☆ | ☆ | ☆ | ☆ | ☆ | ☆ | ☆ | ☆ | ☆ |

13. Novelty *

Were the recommendations novel, as in new, unusual or not influenced by popularity?

|   | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|----|
|   | ☆ | ☆ | ☆ | ☆ | ☆ | ☆ | ☆ | ☆ | ☆ | ☆ |

14. Diversity *

Were the recommendations diverse, varied in category, popularity, etc.?

|   | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|----|
|   | ☆ | ☆ | ☆ | ☆ | ☆ | ☆ | ☆ | ☆ | ☆ | ☆ |

## Evaluation of Explanations (Use Case 2)

Evaluation of the explanations given by the agent to each of the recommendations during the first conversation.

15. Transparency *

How clearly you understood why each item was recommended to you.

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|----|
| ☆ | ☆ | ☆ | ☆ | ☆ | ☆ | ☆ | ☆ | ☆ | ☆ |

16. Satisfaction *

Whether you were satisfied by each recommendation and the accompanying explanation.

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|----|
| ☆ | ☆ | ☆ | ☆ | ☆ | ☆ | ☆ | ☆ | ☆ | ☆ |

17. Trust / Persuasiveness *

Whether the explanations increased your trust in the system/agent.

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|----|
| ☆ | ☆ | ☆ | ☆ | ☆ | ☆ | ☆ | ☆ | ☆ | ☆ |

18. Cognitive Load *

Whether the explanations were mentally demanding to understand.

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|----|
| ☆ | ☆ | ☆ | ☆ | ☆ | ☆ | ☆ | ☆ | ☆ | ☆ |

## Use Case 3: Subsequent Conversation with Recommendation Agent

For this task, you will be having another conversation with the agent you just created, after retraining the agent manually. Once again, you will need to give your preferences to the agent and receive recommendations.

**Steps**:

1. Go to the Agent Hub.
2. Click on "Retrain" on the agent you created in the first step, using your previous session as new interactions.
3. Wait for the agent to be retrained. When it's ready, click on "Start Conversation"
4. Click on "Begin Conversation".
5. Chat with the agent once more, responding to their questions until obtaining recommendations.
6. Rate each recommendation and send the feedback.
7. End the session whenever you want.
8. Check the archived session in the "Chat Sessions" menu (Optional).

19.    How difficult was it to complete this task? *

*Mark only one oval.*

|  | 1 | 2 | 3 | 4 | 5 |  |
|---|---|---|---|---|---|---|
| Very Easy | ◯ | ◯ | ◯ | ◯ | ◯ | Very Difficult |

## Evaluation of Recommendations (Use Case 3)

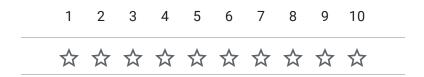Evaluation of the recommendations given by the agents during a subsequent conversation.
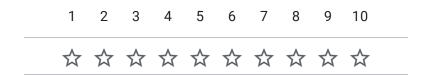
20.    Accuracy *

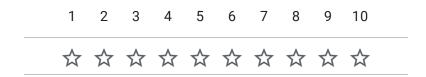Were the recommendations relevant to you, did you like them?

|  | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
|  | ☆ | ☆ | ☆ | ☆ | ☆ | ☆ | ☆ | ☆ | ☆ | ☆ |

21. **Contextual Relevance** *

Were the recommendations relevant to the contextual preferences that you gave to the agent (category, genre, brand, etc.)?

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|
| ☆ | ☆ | ☆ | ☆ | ☆ | ☆ | ☆ | ☆ | ☆ | ☆ |

22. **Novelty** *

Were the recommendations novel, as in new, unusual or not influenced by popularity?

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|
| ☆ | ☆ | ☆ | ☆ | ☆ | ☆ | ☆ | ☆ | ☆ | ☆ |

23. **Diversity** *

Were the recommendations diverse, varied in category, popularity, etc.?

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|
| ☆ | ☆ | ☆ | ☆ | ☆ | ☆ | ☆ | ☆ | ☆ | ☆ |

## Evaluation of Explanations (Use Case 3)

Evaluation of the explanations given by the agent to each of the recommendations during a subsequent conversation.

24. **Transparency** *

How clearly you understood why each item was recommended to you.

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|
| ☆ | ☆ | ☆ | ☆ | ☆ | ☆ | ☆ | ☆ | ☆ | ☆ |

25. **Satisfaction** *

Whether you were satisfied by each recommendation and the accompanying explanation.

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |

☆ ☆ ☆ ☆ ☆ ☆ ☆ ☆ ☆ ☆

26. **Trust / Persuasiveness** *

Whether the explanations increased your trust in the system/agent.

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |

☆ ☆ ☆ ☆ ☆ ☆ ☆ ☆ ☆ ☆

27. **Cognitive Load** *

Whether the explanations were mentally demanding to understand.

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |

☆ ☆ ☆ ☆ ☆ ☆ ☆ ☆ ☆ ☆

## System Usability Scale Questionnaire (SUS)

Standardized questions on the usability of HybridCRS as a platform.

28. **1. I think that I would like to use this system frequently** *

*Mark only one oval.*

|  | 1 | 2 | 3 | 4 | 5 |  |
|---|---|---|---|---|---|---|
| Strongly Disagree | ◯ | ◯ | ◯ | ◯ | ◯ | Strongly Agree |

29. **2. I found the system unnecessarily complex** *

*Mark only one oval.*

|  | 1 | 2 | 3 | 4 | 5 |  |
|---|---|---|---|---|---|---|
| Strongly Disagree | ◯ | ◯ | ◯ | ◯ | ◯ | Strongly Agree |

30. **3. I thought the system was easy to use** *

*Mark only one oval.*

|  | 1 | 2 | 3 | 4 | 5 |  |
|---|---|---|---|---|---|---|
| Strongly Disagree | ◯ | ◯ | ◯ | ◯ | ◯ | Strongly Agree |

31. **4. I think that I would need the support of a technical person to be able to use this system** *

*Mark only one oval.*

|  | 1 | 2 | 3 | 4 | 5 |  |
|---|---|---|---|---|---|---|
| Strongly Disagree | ◯ | ◯ | ◯ | ◯ | ◯ | Strongly Agree |

32. **5. I found the various functions in this system were well integrated** *

*Mark only one oval.*

|  | 1 | 2 | 3 | 4 | 5 |  |
|---|---|---|---|---|---|---|
| Strongly Disagree | ◯ | ◯ | ◯ | ◯ | ◯ | Strongly Agree |

33. **6. I thought there was too much inconsistency in this system** *

*Mark only one oval.*

|  | 1 | 2 | 3 | 4 | 5 |  |
|---|---|---|---|---|---|---|
| Strongly Disagree | ◯ | ◯ | ◯ | ◯ | ◯ | Strongly Agree |

34. **7. I would imagine that most people would learn to use this system very quickly** *

*Mark only one oval.*

|  | 1 | 2 | 3 | 4 | 5 |  |
|---|---|---|---|---|---|---|
| Strongly Disagree | ◯ | ◯ | ◯ | ◯ | ◯ | Strongly Agree |

35. **8. I found the system very cumbersome to use** *

*Mark only one oval.*

|  | 1 | 2 | 3 | 4 | 5 |  |
|---|---|---|---|---|---|---|
| Strongly Disagree | ◯ | ◯ | ◯ | ◯ | ◯ | Strongly Agree |

36. **9. I felt very confident using the system** *

*Mark only one oval.*

|  | 1 | 2 | 3 | 4 | 5 |  |
|---|---|---|---|---|---|---|
| Strongly Disagree | ◯ | ◯ | ◯ | ◯ | ◯ | Strongly Agree |

37. **10. I needed to learn a lot of things before I could get going with this system** *

*Mark only one oval.*

|                    | 1 | 2 | 3 | 4 | 5 |                |
|--------------------|---|---|---|---|---|----------------|
| Strongly Disagree  | ◯ | ◯ | ◯ | ◯ | ◯ | Strongly Agree |

## Feedback

Here you can give any feedback you have about the application.

38. Feedback

_____

_____

_____

_____

_____

39. General Rating *

How much did you like the project in general (Recommendations, Explanations, User Interface, Ease of Use...)

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|----|
| ☆ | ☆ | ☆ | ☆ | ☆ | ☆ | ☆ | ☆ | ☆ | ☆  |

Google Forms