lileipisces.github.io

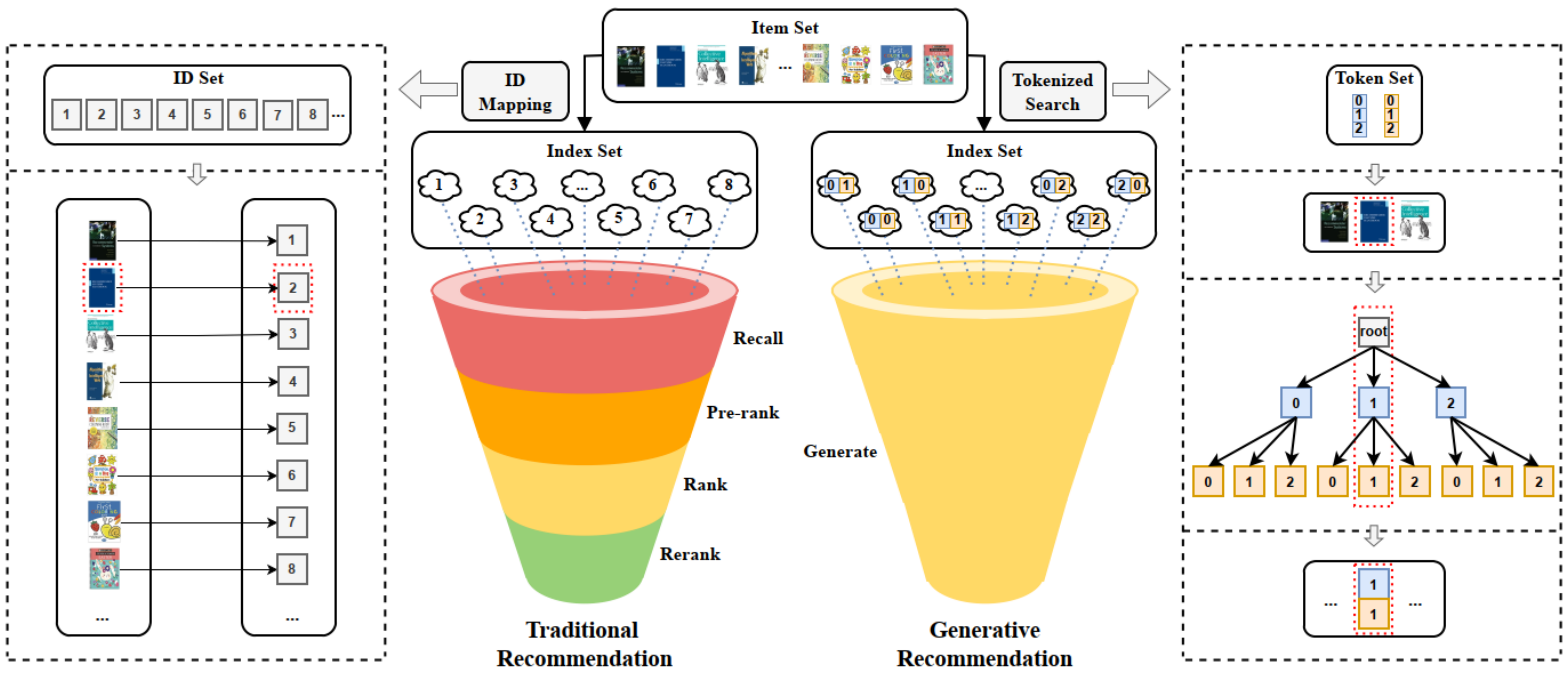
Large Language Models for Generative Recommendation: A Survey and Visionary Discussions

Lei Li[1], Yongfeng Zhang[2], Dugang Liu[3], Li Chen[1]

[1]{csleili, lichen}@comp.hkbu.edu.hk, [2]yongfeng.zhang@rutgers.edu, [3]dugang.ldg@gmail.com

Why Generative Recommendation

- Huge number of items on recommendation platforms
- Computationally expensive score calculation for each
- Multi-stage filtering to narrow down candidates
    - Simple methods at early stage
    - Complex models at final stage
- Gap between academic research and industrial applications



- Simplify recommendation process to one stage
    - Directly generate items for recommendation
    - Implicitly enumerate all items
- Use finite tokens to represent infinite items
    - # tokens = 1000
    - ID length = 10 tokens
    - # items = $1000^{10} = 10^{30}$

ID Creation Methods

Generalized Definition of ID

An ID in recommender systems is a sequence of tokens that can uniquely identify an entity, such as a user or an item.

- Embedding ID
- Sequence of numerical tokens
    - <item><_><73><91>
- Sequence of word tokens
    - Item title
        - The Lord of the Rings
    - Item description
    - News article
    - Sequence of meaningless words
        - Ring epic journey fellowship adventure [1]

LLM-compatible IDs

- Retain collaborative information of IDs in LLM environment
    - User-user
    - Item-item
    - User-item
- Short and exact representations of IDs
    - Similar users/items share more tokens
    - Remaining tokens distinguish them

Spectral Clustering

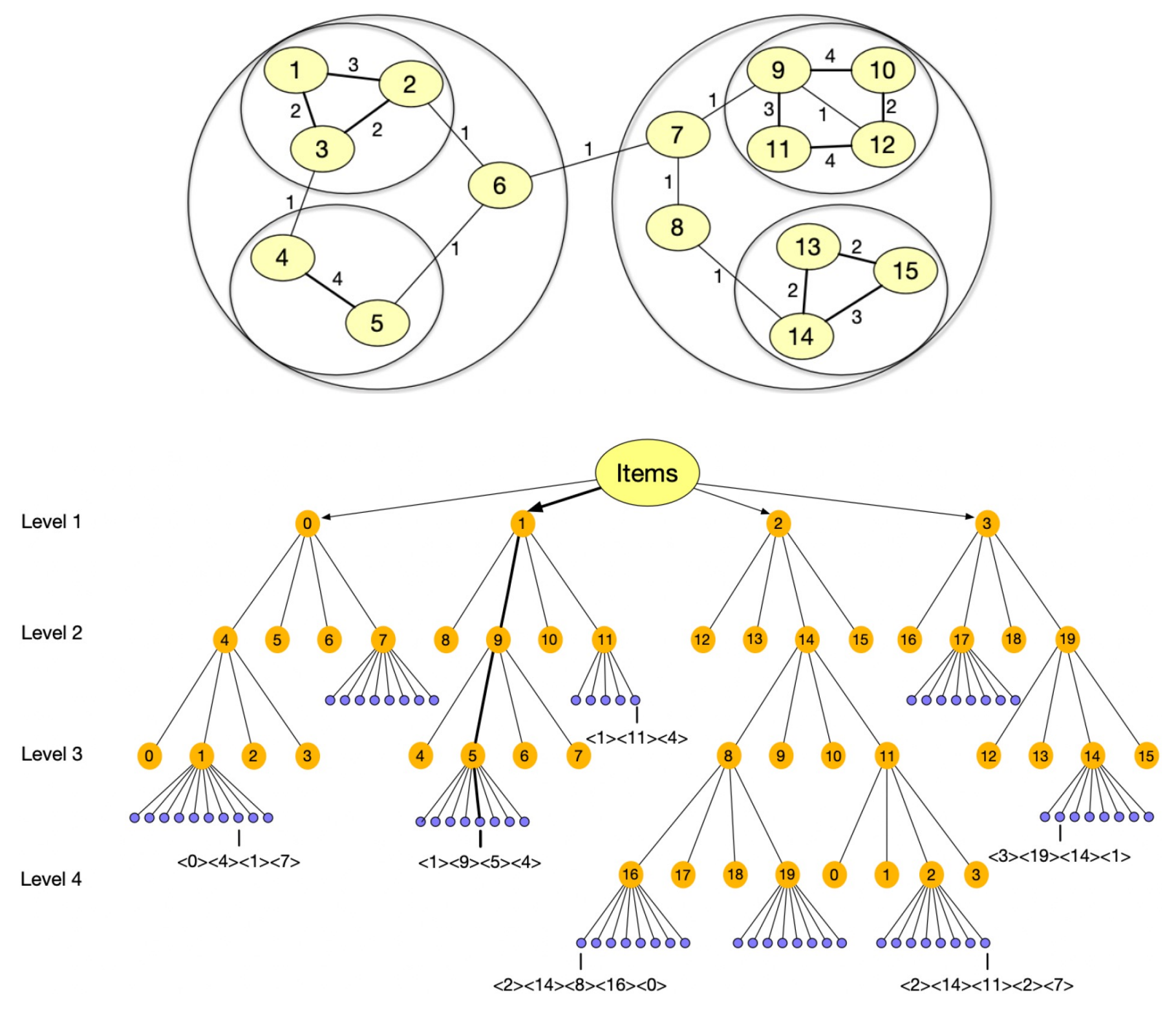


Image credit to [1]

How to Do Generative Recommendation

Rating Prediction

How would *user_1234* rate *item_5678*? → LLM → 4.12

Top-*N* Recommendation

Select one item as recommendation for *user_1234* from the following candidates: *item_6783*, ..., *item_9312*, *item_2834* → LLM → 9312

Sequential Recommendation

Given *user_1234*'s interaction history *item_3456*, ..., *item_4567*, *item_5678*, predict the next item that the user will click → LLM → 6789

Explainable Recommendation

Explain to *user_1234* why *item_5678* is recommended → LLM → The movie is top-notch

Review Generation

Generate a review for *user_1234* about *item_5678* → LLM → The hotel is located in ...

Review Summarization

Please summarize the following review that *user_1234* wrote for *item_5678*: *the hotel is located in ...* → LLM → great location

Conversational Recommendation

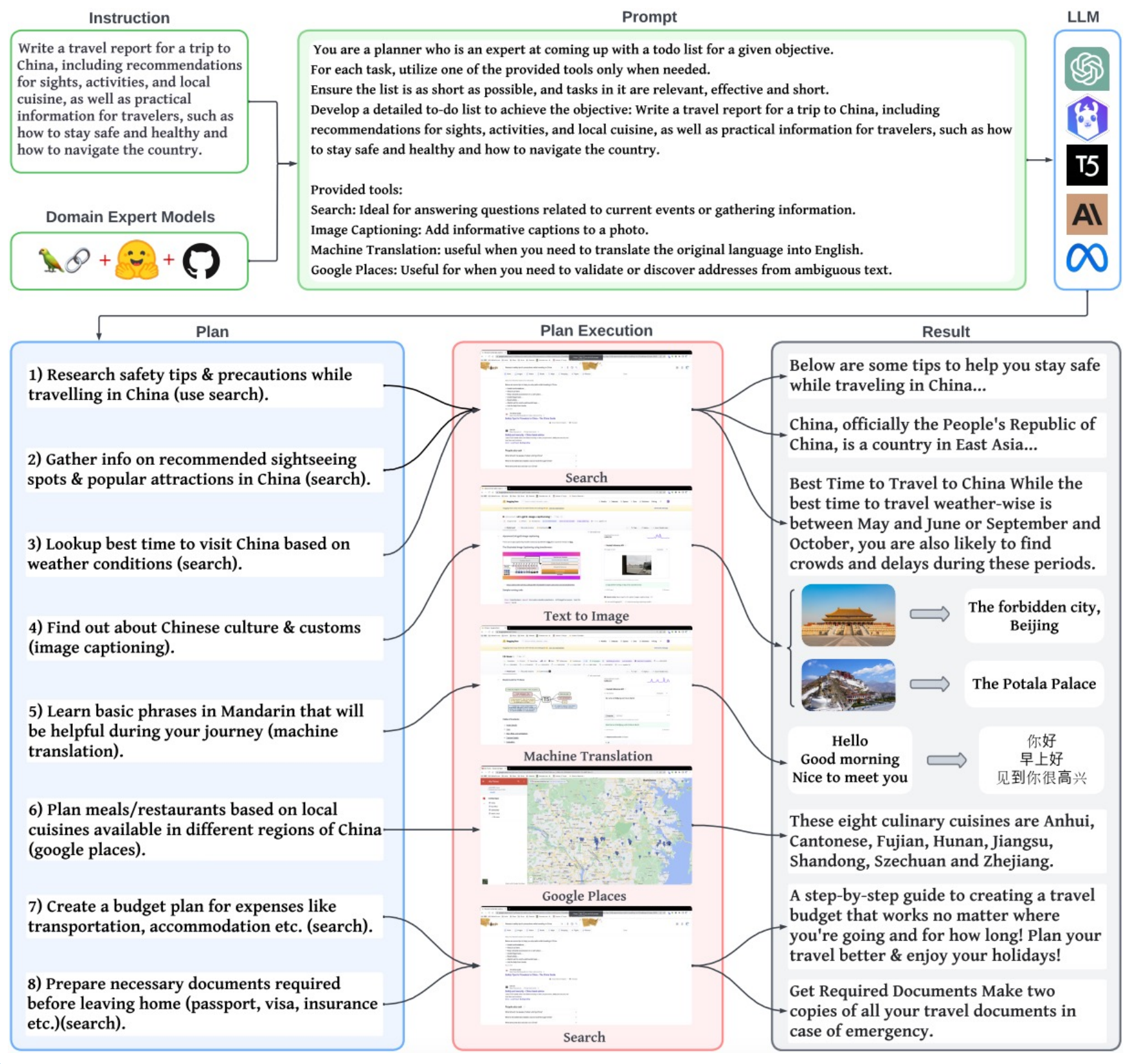
USER: I like action movie.
SYSTEM: How about Mission Impossible?
USER: I have watched it before.
SYSTEM:
→ LLM → What about Heart of Stone?

Evaluation

- Automatic evaluation
    - RMSE and MAE for rating prediction
    - NDCG, precision and recall for top-*N* recommendation and sequential recommendation
    - BLEU and ROUGE for text similarity
- Online A/B tests
- Human evaluation

Challenges and Opportunities

LLM-based Agents for Trip Recommendation



Write a travel report for a trip to China, including recommendations for sights, activities, and local cuisine, as well as practical information for travelers, such as how to stay safe and healthy and how to navigate the country.

You are a planner who is an expert at coming up with a todo list for a given objective.
For each task, utilize one of the provided tools only when needed.
Ensure the list is as short as possible, and tasks in it are relevant, effective and short.
Develop a detailed to-do list to achieve the objective: Write a travel report for a trip to China, including recommendations for sights, activities, and local cuisine, as well as practical information for travelers, such as how to stay safe and healthy and how to navigate the country.

Provided tools:
Search: Ideal for answering questions related to current events or gathering information.
Image Captioning: Add informative captions to a photo.
Machine Translation: useful when you need to translate the original language into English.
Google Places: Useful for when you need to validate or discover addresses from ambiguous text.

1) Research safety tips & precautions while travelling in China (use search).
2) Gather info on recommended sightseeing spots & popular attractions in China (search).
3) Lookup best time to visit China based on weather conditions (search).
4) Find out about Chinese culture & customs (image captioning).
5) Learn basic phrases in Mandarin that will be helpful during your journey (machine translation).
6) Plan meals/restaurants based on local cuisines available in different regions of China (google places).
7) Create a budget plan for expenses like transportation, accommodation etc. (search).
8) Prepare necessary documents required before leaving home (passport, visa, insurance etc.)(search).

Below are some tips to help you stay safe while traveling in China...

China, officially the People's Republic of China, is a country in East Asia...

Best Time to Travel to China While the best time to travel weather-wise is between May and June or September and October, you are also likely to find crowds and delays during these periods.

The forbidden city, Beijing

The Potala Palace

Hello
Good morning
Nice to meet you

你好
早上好
见到你很高兴

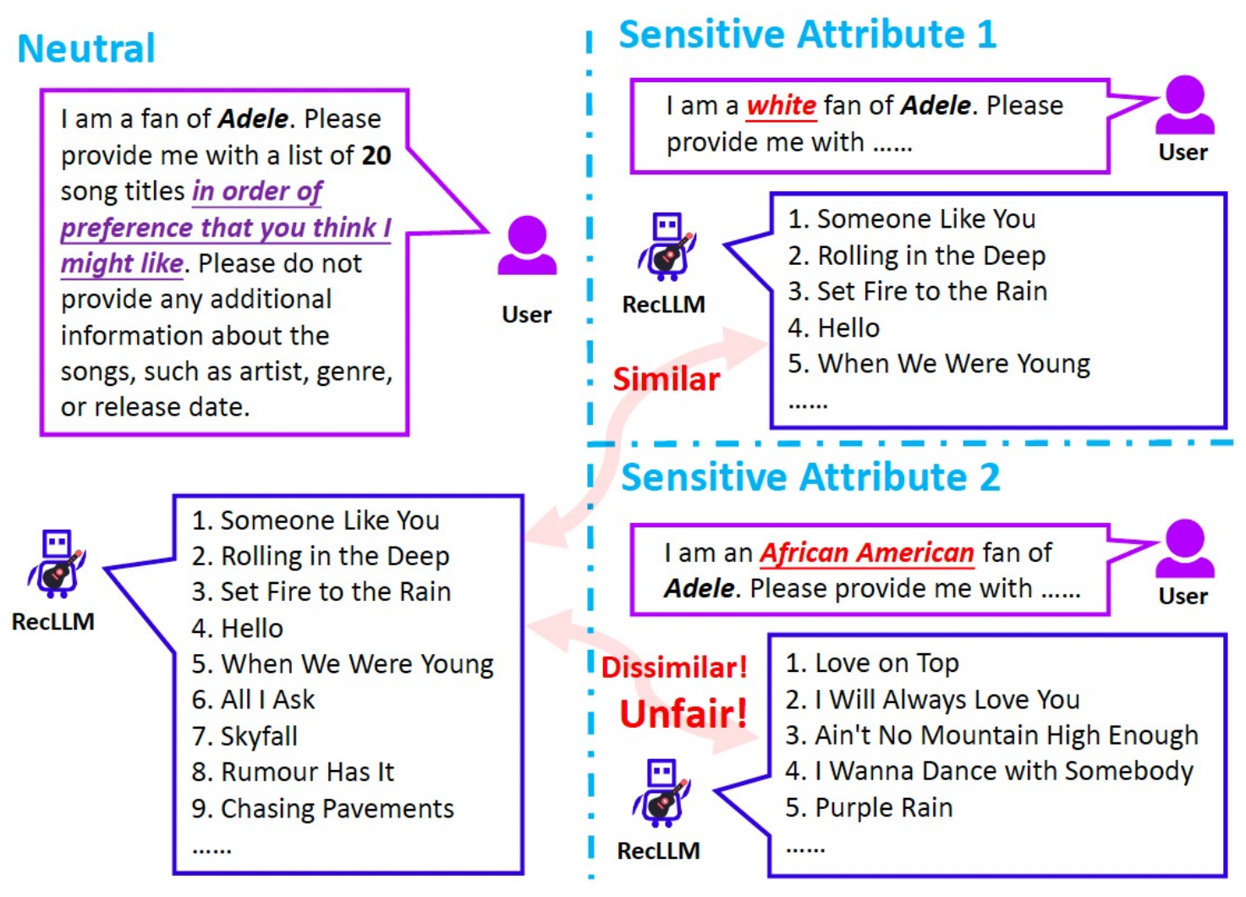
These eight culinary cuisines are Anhui, Cantonese, Fujian, Hunan, Jiangsu, Shandong, Szechuan and Zhejiang.

A step-by-step guide to creating a travel budget that works no matter where you're going and for how long! Plan your travel better & enjoy your holidays!

Get Required Documents Make two copies of all your travel documents in case of emergency.

Image credit to [2]

Recommendation Bias



I am a fan of ***Adele***. Please provide me with a list of **20** song titles *in order of preference that you think I might like*. Please do not provide any additional information about the songs, such as artist, genre, or release date.

1. Someone Like You
2. Rolling in the Deep
3. Set Fire to the Rain
4. Hello
5. When We Were Young
6. All I Ask
7. Skyfall
8. Rumour Has It
9. Chasing Pavements
......

I am a ***white*** fan of ***Adele***. Please provide me with ......

1. Someone Like You
2. Rolling in the Deep
3. Set Fire to the Rain
4. Hello
5. When We Were Young
......

I am an ***African American*** fan of ***Adele***. Please provide me with ......

1. Love on Top
2. I Will Always Love You
3. Ain't No Mountain High Enough
4. I Wanna Dance with Somebody
5. Purple Rain
......

Image credit to [3]

What is the boundary between bias and personalization?

Others

- Hallucination
- Content Bias
- Transparency and Explainability
- Controllability
- Inference Efficiency
- Multimodal Recommendation
- Cold-start Recommendation

Conclusion

- Research of generative recommendation in line with the trend of AI
    - Discriminative AI -> generative AI

References

[1] Hua, Wenyue, et al. "How to index item ids for recommendation foundation models." SIGIR-AP’23.

[2] Ge, Yingqiang, et al. "Openagi: When llm meets domain experts." NeurIPS’24.

[3] Zhang, Jizhi, et al. "Is chatgpt fair for recommendation? evaluating fairness in large language model recommendation." RecSys’23.

The 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024), Turin, Italy, May 20-25, 2024