Dialogue Generation for Conversational Recommendation Using User Search Interactions and Knowledge Graphs

Minsung Hwang[1], Donggyu Yoon[1], Gowoon Hwang[2], Kwanghee Jang[2], and Joyce Jiyoung Whang[1†]

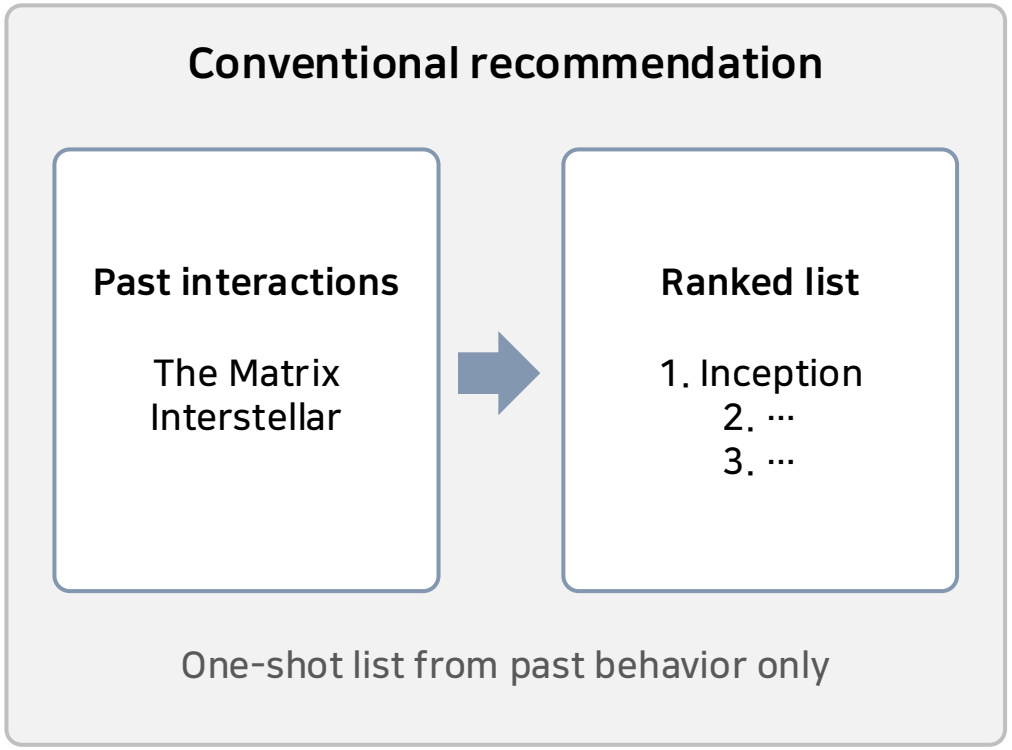
[1]KAIST, [2]LG Electronics, † Corresponding Author

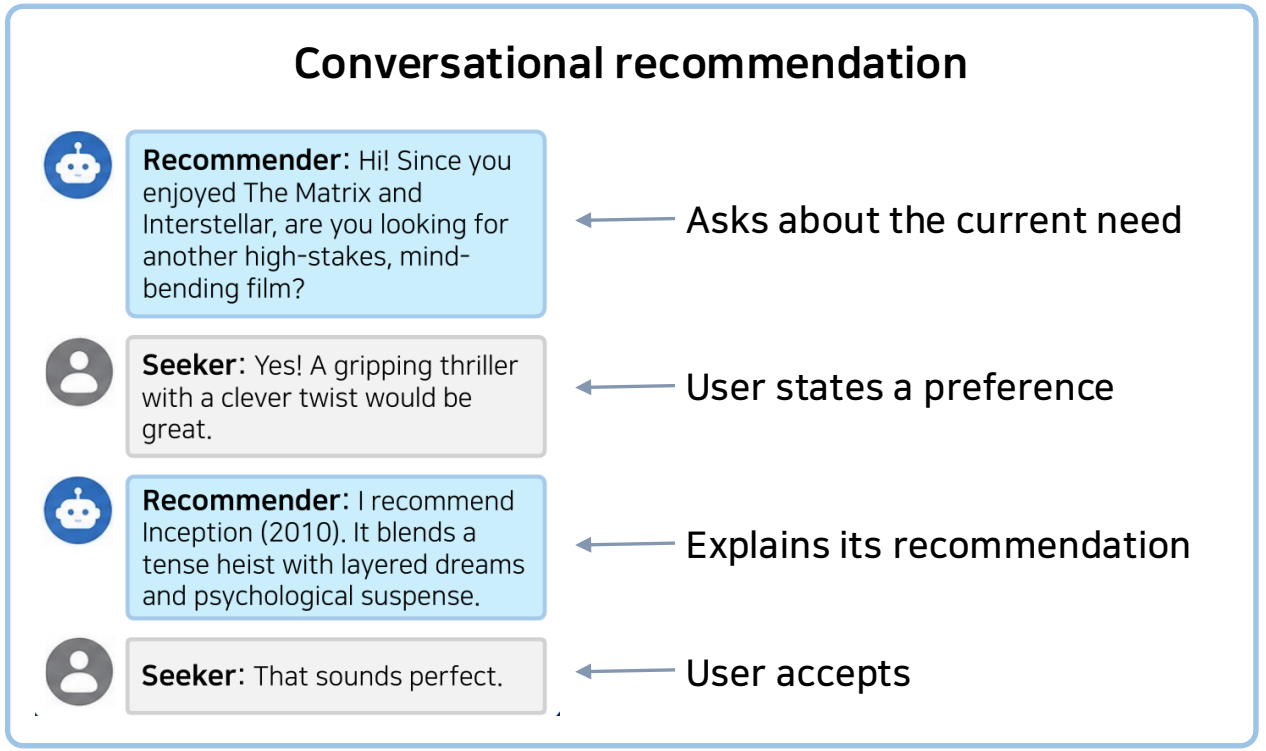
The 3rd Workshop on Evaluating and Applying Recommender Systems with Large Language Models (EARL@CIKM 2026)

KAIST BDILab BIG DATA INTELLIGENCE

01 Conversational Recommender Systems

- A **conversational recommender system (CRS)** recommends items through **multi-turn dialogue**
  - It elicits preferences, refines requirements, and explains its recommendations
- Unlike recommending from past interactions alone, a CRS can respond to the user's **current need**





01 Conversational Recommendation Datasets

- CRSs need dialogues with **realistic user intents**, **appropriate targets**, and **coherent flows**
- Human-collected datasets (ReDIAL, INSPIRED)
  - Human dialogue collection and annotation are **costly and difficult to scale**
- LLM-based generation (LLM-REDIAL, MUSE, DistillRecDial)
  - Relies on **manually designed templates** and preference profiles from ratings or reviews
  - Does not use **observed search-click behavior** or **KG-grounded explanations**

02 SearCI: Dialogue Generation from Search-Click Interactions

Search-Click Session

Time	Query	Click
00:13	animation	Zootopia
10:12	Movie	The Matrix
10:17	Movie	Interstellar
10:20	Thriller	Inception
⋮		

* Blue area shows a query session.

Constructing User Context

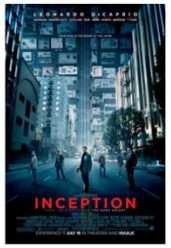
Historical Likes



The Matrix


Interstellar

Target Contents



Inception

Dialogue Generation

Recommender: Hi! Since you enjoyed The Matrix and Interstellar, are you looking for another high-stakes, mind-bending film?

Seeker: Yes! A gripping thriller with a clever twist would be great.

Recommender: I recommend Inception (2010). It blends a tense heist with layered dreams and psychological suspense.

Seeker: That sounds perfect.

02 Query Session Construction

- Group search-click interactions into **query sessions** based on inter-event gaps
  - Session threshold from a two-component Gaussian mixture on log gaps
- Items clicked **before a query** form the user context; the item clicked **right after it** is the target

Search-Click Session

Time	Query	Click
00:13	animation	Zootopia
10:12	Movie	The Matrix
10:17	Movie	Interstellar
10:20	Thriller	Inception
⋮		

* Blue area shows a query session.

Constructing User Context

Historical Likes



The Matrix


Interstellar

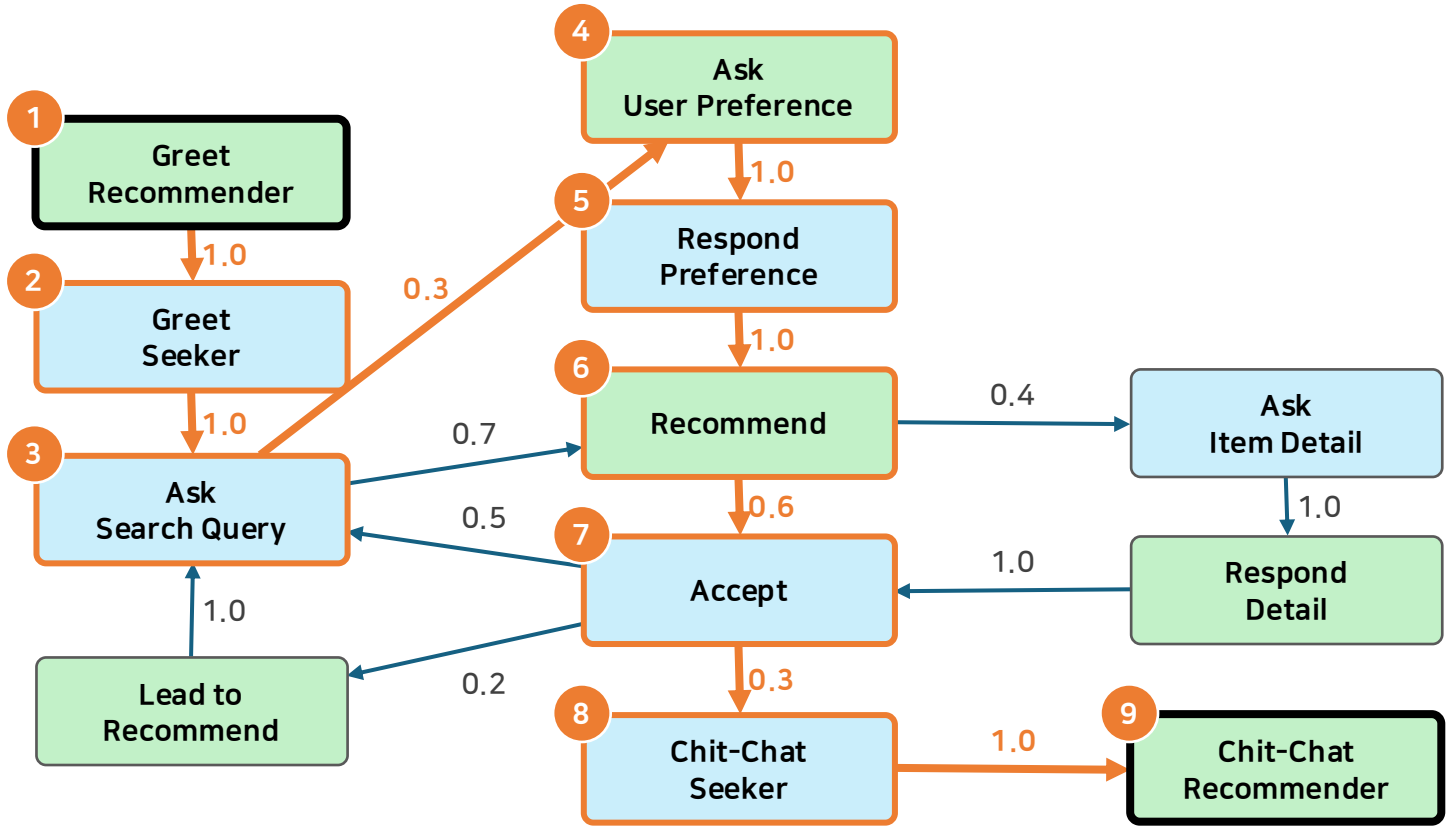
Target Contents



Inception

02 Dialogue Turn Planning

- A **probabilistic state machine** samples a sequence of subgoals for each dialogue
  - Consecutive subgoals of the same speaker are **merged into one turn**



Sampled turn plan

T1	**Recommender:** Greet	(1)
T2	**Seeker:** Greet + Ask search query	(2–3)
T3	**Recommender:** Ask user preference	(4)
T4	**Seeker:** Respond preference	(5)
T5	**Recommender:** Recommend	(6)
T6	**Seeker:** Accept + Chit-chat	(7–8)
T7	**Recommender:** Chit-chat	(9)

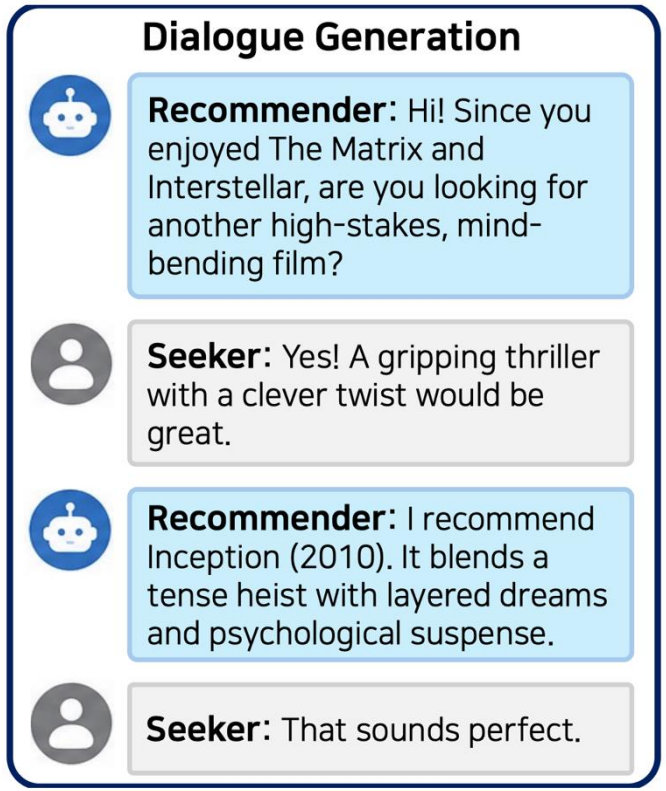
02 What Each Turn in the Plan Specifies

- Each turn specifies the **active speaker**, **subgoals**, **content items**, and **turn-specific constraints**
  - The plan fixes the recommendation flow, while the agents decide the wording

Turn	Speaker	Subgoal(s)	Content items	Constraint
T1	Recommender	Greet	User context: Ttoltoli: The Secret of the Baby Dinosaur	
T2	Seeker	Greet, Ask search query	Search query: “movie”	Do not reveal the target title
T3	Recommender	Ask user preference		
T4	Seeker	Respond preference		
T5	Recommender	Recommend	Target: The Super Mario Bros. Movie (2023)	Include the target and supporting information
T6	Seeker	Accept, Chit-chat		
T7	Recommender	Chit-chat		

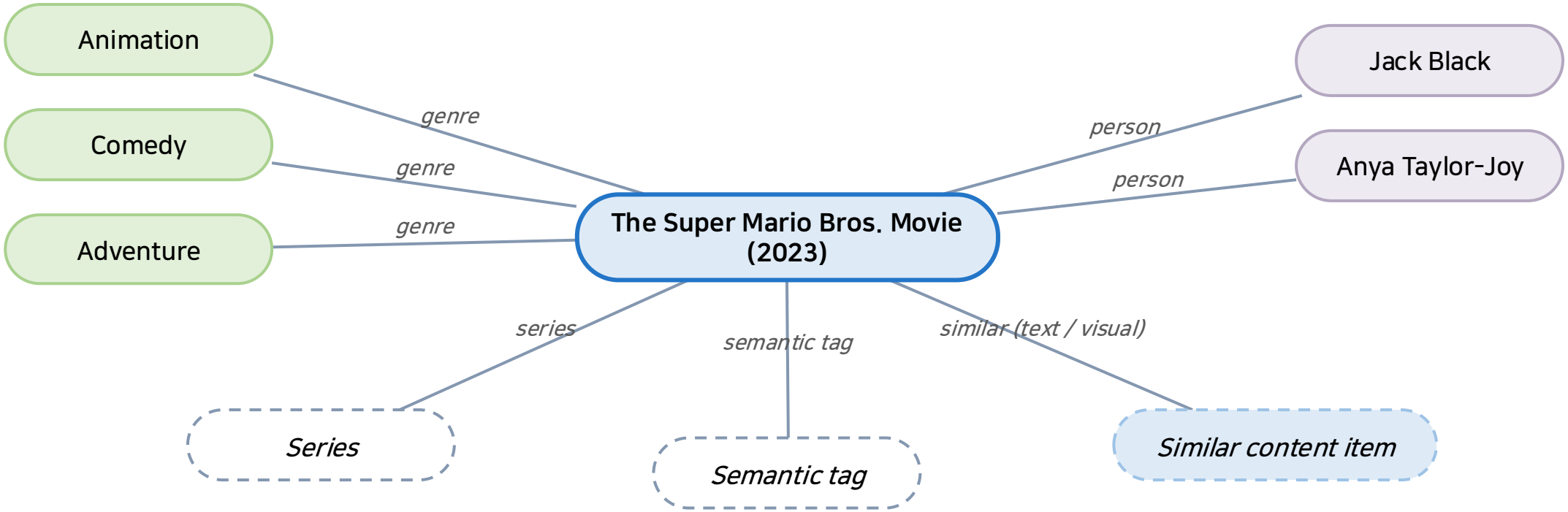
02 Automatic Dialogue Generation

- **Seeker** and **recommender** LLM agents generate the dialogue one turn at a time
  - Each agent receives its role, dialogue history, current subgoal, content items, and turn constraints
  - Search-request turns **hide the target title**; recommendation turns **must include the target**



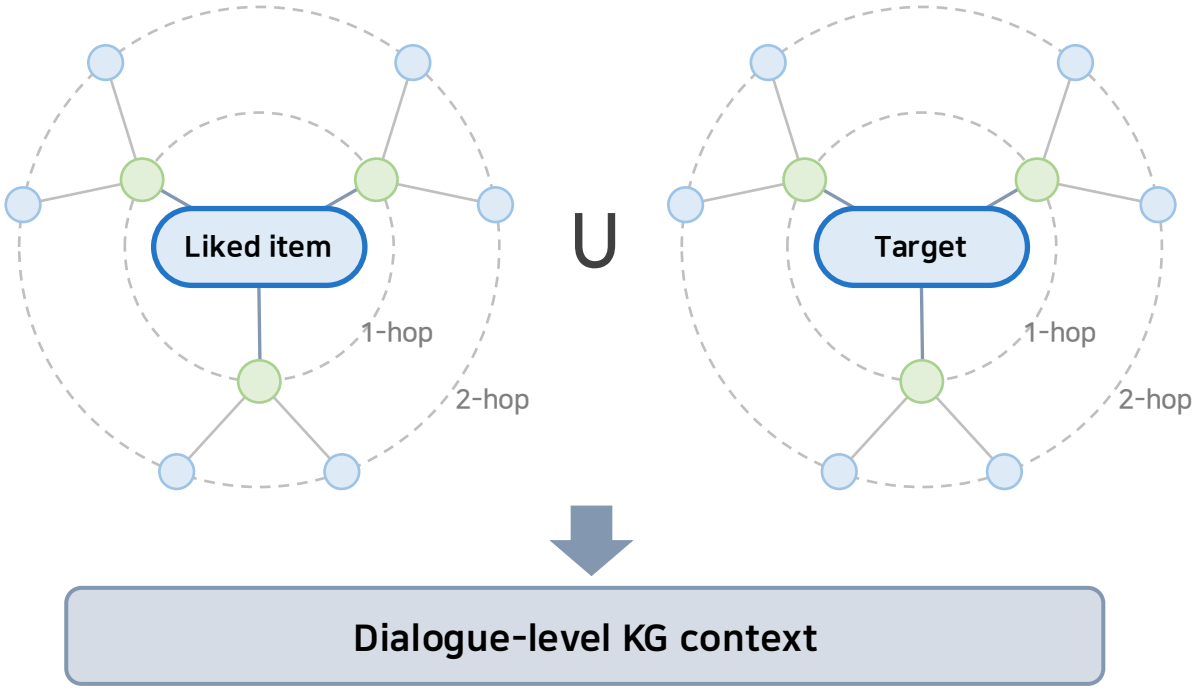
02 Content Knowledge Graph

- Built from **structured content metadata** of movies and TV shows
  - Links items to **genres, people, series, and semantic tags**, and to **related items with similar textual or visual features**
  - **792,395** entities, **5,618,485** triplets, and **14** relations



02 KG Grounding and Filtering

- For every item in the user context, the target, or the plan, extract a **bounded two-hop subgraph**
  - The number of neighbors kept at each hop is limited to keep prompts short
- Their union forms the **dialogue-level KG context** given to the recommender



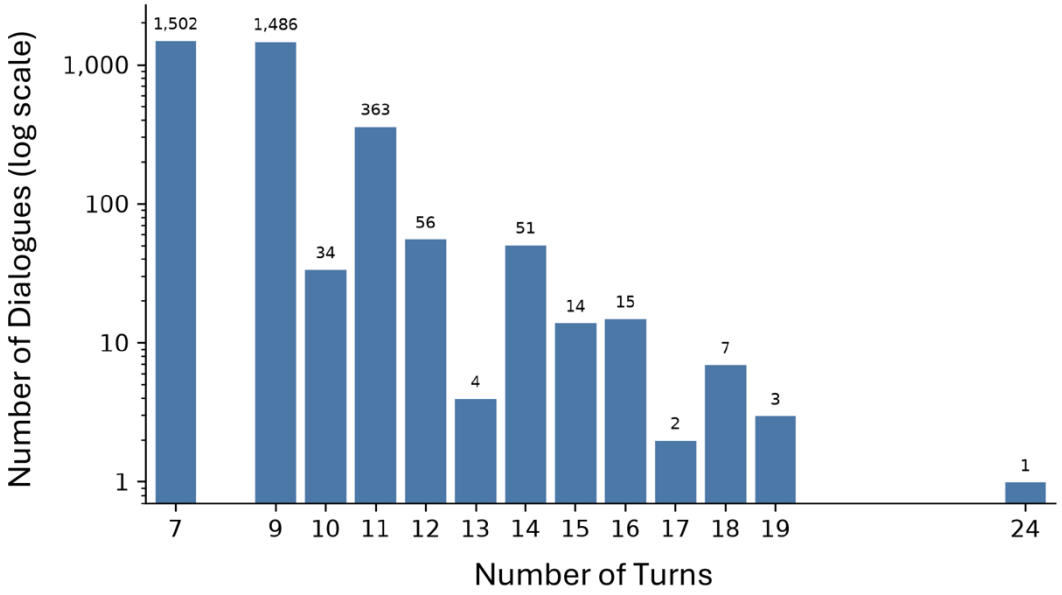
Automatic filters after generation

- Target title revealed before the recommendation turn
- Search query directly matches the target title

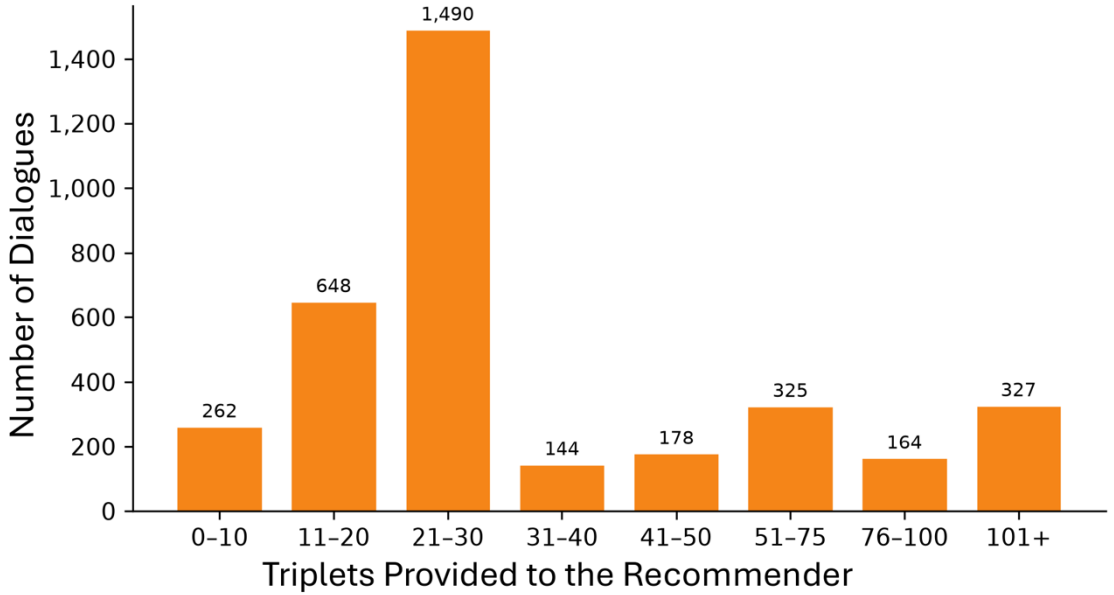
10,000 candidates → **3,538 retained**

03 Generated Dataset

- **3,538 dialogues** retained from 10,000 candidates after filtering
  - 1,731 unique targets, 8.58 turns on average, and **164 unique turn plans**
  - Generated with gpt-5.4-mini from anonymized TV search and viewing records



Turn-plan length



KG triplets per dialogue

03 Example of a Generated Dialogue

Search query: movie | **User context**: Ttoltoli: The Secret of the Baby Dinosaur | **Target**: The Super Mario Bros. Movie (2023)

GREET_RECOMMENDER	Hello! You previously enjoyed children's animation such as **Ttoltoli: The Secret of the Baby Dinosaur**. Would you like to watch something similar today, or should I suggest something with a different atmosphere?
GREET_SEEKER → ASK_SEARCH_QUERY	Hello! Yes, please choose something for me again today. I am looking for a light and entertaining movie that is not too serious and is suitable for watching with my family. Could you recommend one?
ASK_USER_PREFERENCE	For something that you can comfortably watch with your family, animation or comedy would be a good fit. Would you prefer something completely bright and cheerful, or a story with a little more adventure?
RESPOND_PREFERENCE	I would prefer something bright and cheerful!
RECOMMEND	In that case, I recommend The Super Mario Bros. Movie (2023). Its combination of comedy, animation, and adventure makes it bright and entertaining. With voice performances by Jack Black and Anya Taylor-Joy, it is also a fun choice for the whole family.
ACCEPT → CHITCHAT_SEEKER	Great, I will watch that one! Its bright and cheerful atmosphere sounds perfect for a relaxing movie night with my family. If we enjoy it, I will ask for something similar next time.
CHITCHAT_RECOMMENDER	I hope you have a great time watching it with your family! If you enjoy it, I can recommend another bright and lighthearted movie next time.

Original Korean utterances translated into English

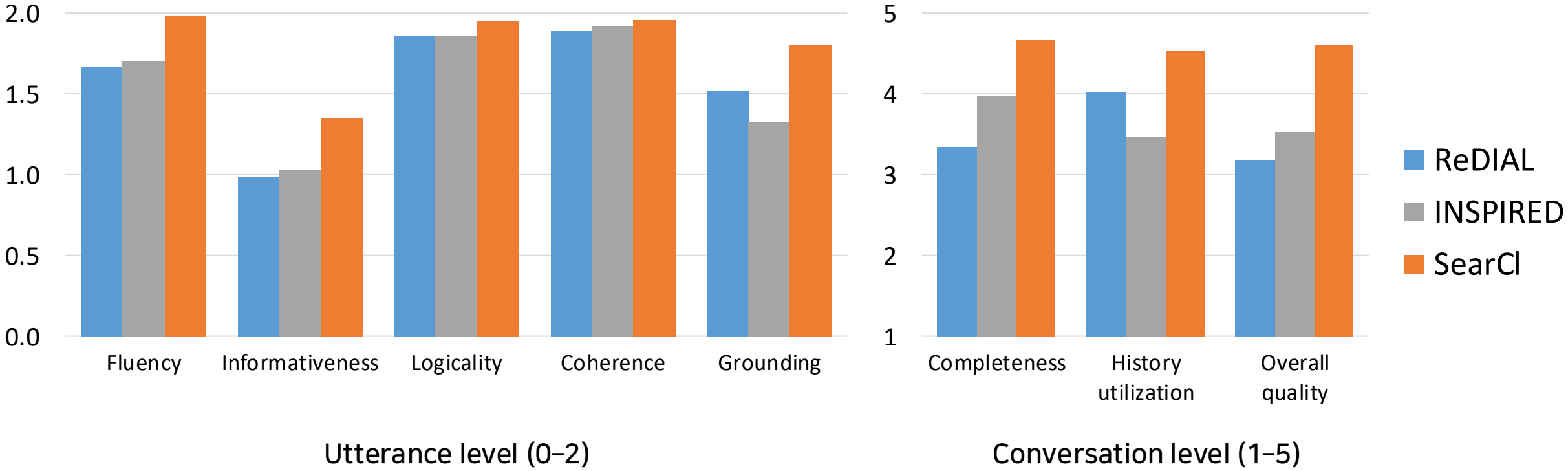
03 Evaluation Metrics

- LLM-as-a-judge (gpt-5.6-luna) scores SearCI and 100-dialogue samples from ReDIAL and INSPIRED
  - The first four metrics follow LLM-REDIAL; metrics in **blue** are introduced in this work

Level	Metric	What the judge evaluates
Utterance (0–2)	Fluency	Grammatical correctness and naturalness
	Informativeness	Whether the utterance gives useful and specific information
	Logicality	Internal consistency
	Coherence	How well the utterance follows the preceding dialogue
	Constraint adherence	Designated speaker and subgoal, target included, and no target leakage
	Grounding	Recommendations supported by user context, metadata, or KG evidence
Conversation (1–5)	**Completeness**	Expected recommendation stages in an appropriate order
	History utilization	How effectively the interaction history is used
	Overall quality	Overall naturalness and consistency as a dialogue

03 Dialogue Quality Evaluation

- LLM-as-a-judge comparison with 100-dialogue samples from ReDIAL and INSPIRED
  - SearCI achieves **higher scores on all comparable metrics**
  - Largest gain in **Overall quality** (4.62 vs. 3.18 and 3.53)



03 Ablation Study: Effect of KG Context

- Paired dialogues per query session, differing **only in KG context**
  - Same user context, target, turn plan, and generation model
- KG context **significantly improves overall quality** (two-sided paired t-test)

Condition	Overall quality
w/o KG context	4.556
w/ KG context	**4.620**

$$p = 6.3 \times 10^{-8}$$

03 Conclusion

- Propose SearCI, which generates CRS dialogues grounded in observed search-click behavior
- A configurable probabilistic state machine enables flexible dialogue design
- A content KG grounds recommendation explanations in auditable evidence
- Generated dialogues achieve higher quality scores than ReDIAL and INSPIRED