

Conversational Recommender Systems

Group 7

Lim Wei Jie

Chia Dehan

Colin Jiang Kelin

Ng Juan Yong



Table of contents

01 Problem Statement

03 Algorithms

HRED, KBRD, UniCRS

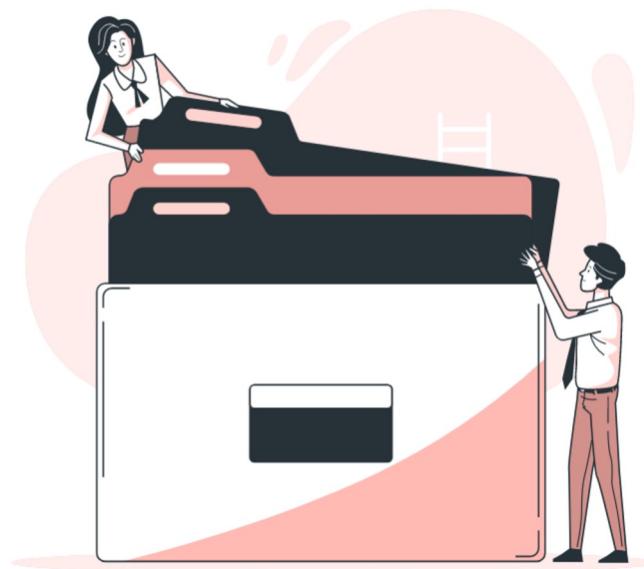
05 Chatbot Demo

02 Data Sets

ReDial, MovieLens 1M

04 Evaluations

Results, Metrics



Introduction



“Recommender systems have become indispensable in our daily lives, guiding our choices and helping us navigate the digital landscape.”

-Chatgpt

Problem Statement



In conventional recommender systems, personalized recommendations are highly based on the previous actions of users, including searching, clicking and purchasing.

These actions can be regarded as users' feedbacks that reflect users' interest.

However, due to its implicitness, such feedback can only reflect a part of users' interest, causing inaccuracy in recommendation.

Problem Statement



Another information source about user preferences is the dialog between users and services. In such dialog, users often provide more information about their preferences.

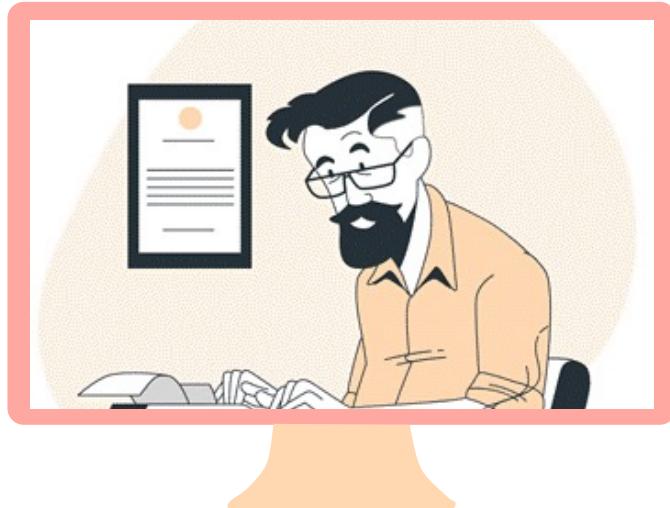
Compared with the implicit feedback, **the feedback from the dialog is more explicit and more related to users' preferences.**

Therefore, a recommender dialog system possesses high commercial potential.

Conversational Recommender System

A recommender dialog system can be regarded as a combination of a recommender system and a dialog system, which can respond to users' utterances with informative natural language expressions and provide high-quality recommendation **based on the content of users' utterances**.

In this setting, information from the recommender system can provide vital information to **maintain multi-turn dialog**, while information from the dialog system that contains **implication of users' preferences can enhance the quality of recommendation**. Besides, the incorporation of external knowledge can strengthen the connections between systems and enhance their performances.



02



Data Sets

ReDial (Recommender Dialogue)

Table 3: Sample conversation from validation set.

SEEKER:	hi ! i 'm looking for an action filled movie similar to jurassic park
HUMAN:	ok , have you seen jurassic world ?
HRED:	hi ! i love that one ! have you seen the conjuring or the conjuring 2 ?
OURS:	hi what kind of movies do you like ?
SEEKER:	yes i loved that movie as well . are there any good movies without dinosaurs ?
HUMAN:	have you seen king kong ?
HRED:	no , i have n't seen it yet . have you seen the avengers ? it 's about a couple of friends who are their own .
OURS:	i have n't seen that one . i will have to check it out . have you seen jurassic park
SEEKER:	yes ! that was another great giant monster movie :)
HUMAN:	what about jaws & ; jaws 2 ? oldies but goodies scary action ? suspense gives me the rumbles
HRED:	i have n't seen that one yet but i heard it was good . i have n't seen it yet . i have n't seen it yet . i have n't seen it yet .
OURS:	i have n't seen that one yet . i will have to check it out . what about walking with dinosaurs ?
SEEKER:	i absolutely loved the first one . one of my favorite movies ever :) honestly i can't say i remember much about the second one . jaws 3-d was quite funny and jaws : the revenge was just weird . i do love suspense too ...
HUMAN:	i like suspense but sometimes i can barely handle it ! it gives me anxiety ! lol
HRED:	no , i have n't seen it yet . i will have to check it out . thanks for the chat .
OURS:	i have n't seen that one . i will have to check it out . i hope i helped .
SEEKER:	that 's understandable . as long as it 's not too gory i do n't generally have a problem with those things .
HUMAN:	well , it was great chatting with you ! have a great one !
HRED:	i 'm glad i was able to help you today
OURS:	have you seen immortals ? it 's a war movie but it 's pretty good
SEEKER:	you too ! thanks !

Background:

- A dataset of dialogues where users recommend movies to each other.
- Consists of over 10,000 conversations, across 1112 workers, 6637 movies, 48873 ratings.
- Conversations are centered around the theme of providing movie recommendations.
- Collected by a team of researchers working at Polytechnique Montréal, MILA – Quebec AI Institute, Microsoft Research Montréal, HEC Montreal, and Element AI.

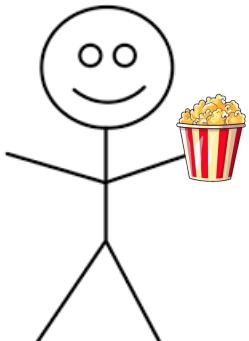
Use Case:

- Can be used to train and evaluate models for conversational recommendations.
- Can also be used to study the different factors that influence how users recommend movies to each other.

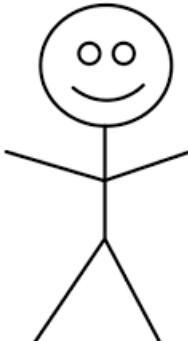
How was ReDial collected?

- Explain what kind of movies he likes
- Ask for suggestions from the recommender

- Understand the seeker's needs
- Recommend appropriate movies



Movie Seeker



Movie Recommender

To improve data quality:

- Use formal language
- Conversations contain a minimum of 10 messages
- Require 4 movies to be mentioned in every conversation.

How was ReDial validated?

Additional labels are collected in the shape of **Movie Dialogue Forms.**

After the experiment, workers are asked 3 questions for each movie:

1. Get "Suggested" label: *Binary*

Is the movie mentioned by the seeker?

Or

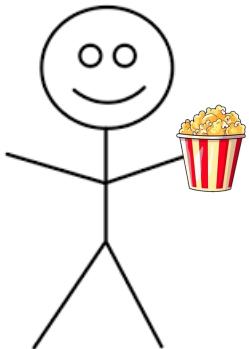
Is the movie suggested by the recommender?

2. Get "Seen" label: *Seen it, Hasn't seen it or Didn't say*

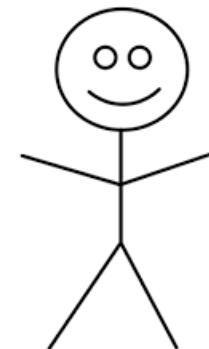
Has the seeker seen the movie?

3. Get "liked" label: *Liked, Didn't like, Didn't say*

Does the seeker like the movie or suggestion?



Movie Seeker



Movie Recommender

1

Both seeker and recommender must answer these questions separately to ensure that they have been conscientious. This ensures high data quality.

How was ReDial validated?

Additional labels are collected in the shape of Movie Dialogue Forms.

After the experiment, workers are asked 3 questions for each movie:

1. Get "Suggested" label: *Binary*

Is the movie mentioned by the seeker?

Or

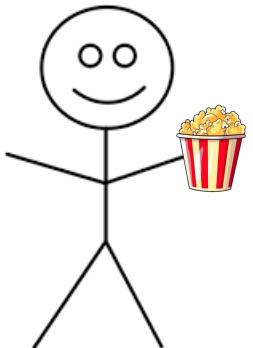
Is the movie suggested by the recommender?

2. Get "Seen" label: *Seen it, Hasn't seen it or Didn't say*

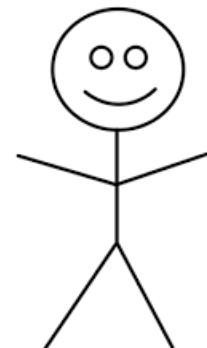
Has the seeker seen the movie?

3. Get "liked" label: *Liked, Didn't like, Didn't say*

Does the seeker like the movie or suggestion?



Movie Seeker



Movie Recommender

2

These labels collected also allow for sub tasks, such as sentiment analysis, to be trained as well, to better evaluate the overall dialogue.

ReDial

Movie Mentions	Respondent Questions	Messages	Conversation Id	Respondent Worker Id	Initiator Worker Id	Initiator Questions
<pre>[{'moviedb': '203371', 'movieName': 'Final Fantasy: The Spirits Within (2001)'} {'moviedb': '84779', 'movieName': 'The Triplets of Belleville (2003)'} ...]</pre>	<pre>[{'moviedb': '203371', 'suggested': 1, 'seen': 0, 'liked': 1} {'moviedb': '84779', 'suggested': 0, 'seen': 1, 'liked': 1} ...]</pre>	<p>Conversational text data</p> <pre>[{'timeOffset': 0, 'text': "Hi there, how are you? I'm looking for movie recommendations", 'senderWorkerId': 0, 'messageId': 1021} {'timeOffset': 15, 'text': "I am doing okay. What kind of movies do you like?", 'senderWorkerId': 1, 'messageId': 1022} {'timeOffset': 66, 'text': "I like animations like @84779 and @191602", 'senderWorkerId': 1, 'messageId': 1023} ...]</pre>	391	1	0	<p>Moviedb : suggested / seen / liked</p> <p>Worker IDs allowed for tracking of user's preferences</p> <p>These binary 1/0 liked allowed for the creation of user-item-rating triplets</p>

MovieLens

Feature	Class	Shape	Dtype	Description
FeaturesDict				
movie_genres	Sequence(ClassLabel)	(None,)	int64	
movie_id	Tensor		string	
movie_title	Tensor		string	
timestamp	Tensor		int64	
user_id	Tensor		string	
user_rating	Tensor		float32	

Background:

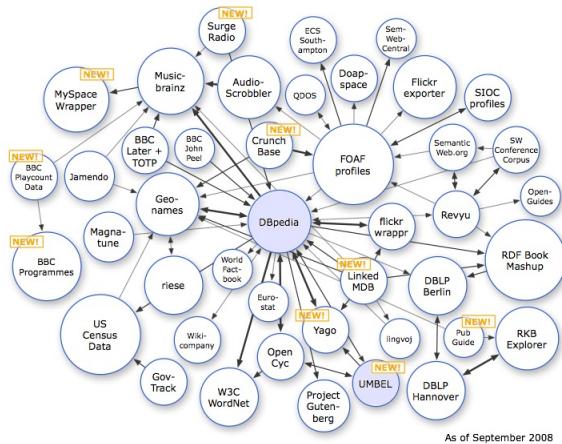
- This dataset describes 5-star rating and free-text tagging activity from MovieLens, a movie recommendation service.
- It contains 27753444 ratings and 1108997 tag applications across 58098 movies.
- These data were created by 283228 users between January 09, 1995 and September 26, 2018.
- This dataset was generated on September 26, 2018.

DBpedia

Open knowledge graph that provides structured information derived from Wikipedia

Incorporates a manually created, shallow, cross-domain ontology, built from the most frequently used Wikipedia infoboxes

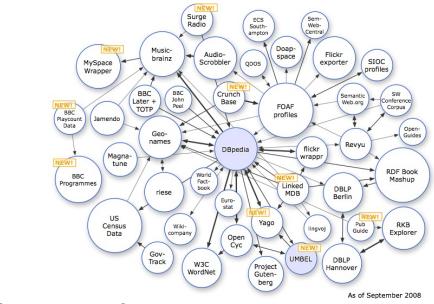
-The ontology helps in mapping structured data from Wikipedia into a consistent, usable format.



Dbpedia is especially useful as it converts unstructured Wikipedia data into structured data, making relationships between different objects and concepts readily available as a knowledge graph.

TLDR:
Dbpedia is **huge graph of Entities** that are **connected to each other via Edges** that are derived from data in Wikipedia

DBpedia



Class	Instances
Resource (overall)	4,233,000
Place	735,000
Person	1,450,000
Work	411,000
Species	251,000
Organisation	241,000

```

<http://dbpedia.org/resource/Monsters_vs._Aliens> <http://dbpedia.org/ontology/director> <http://dbpedia.org/resource/Conrad_Vernon> .
<http://dbpedia.org/resource/Monsters_vs._Aliens> <http://dbpedia.org/ontology/director> <http://dbpedia.org/resource/Rob_Letterman> .
<http://dbpedia.org/resource/Monsters_vs._Aliens> <http://dbpedia.org/ontology/starring> <http://dbpedia.org/resource/Reese_Witherspoon> .
<http://dbpedia.org/resource/Monsters_vs._Aliens> <http://dbpedia.org/ontology/starring> <http://dbpedia.org/resource/Seth_Rogen> .
<http://dbpedia.org/resource/Monsters_vs._Aliens> <http://dbpedia.org/ontology/starring> <http://dbpedia.org/resource/Hugh_Laurie> .
<http://dbpedia.org/resource/Monsters_vs._Aliens> <http://dbpedia.org/ontology/starring> <http://dbpedia.org/resource/Will_Arnett> .
<http://dbpedia.org/resource/Monsters_vs._Aliens> <http://dbpedia.org/ontology/starring> <http://dbpedia.org/resource/Kiefer_Sutherland> .
<http://dbpedia.org/resource/Monsters_vs._Aliens> <http://dbpedia.org/ontology/starring> <http://dbpedia.org/resource/Rain_Wilson> .
<http://dbpedia.org/resource/Monsters_vs._Aliens> <http://dbpedia.org/ontology/starring> <http://dbpedia.org/resource/Paul_Rudd> .
<http://dbpedia.org/resource/Monsters_vs._Aliens> <http://dbpedia.org/ontology/starring> <http://dbpedia.org/resource/Stephen_Colbert> .
<http://dbpedia.org/resource/Monsters_vs._Aliens> <http://dbpedia.org/ontology/musicComposer> <http://dbpedia.org/resource/Henry_Jackson> .
<http://dbpedia.org/resource/Monsters_vs._Aliens> <http://dbpedia.org/ontology/distributor> <http://dbpedia.org/resource/Paramount_Pictures> .
<http://dbpedia.org/resource/Monsters_vs._Aliens> <http://dbpedia.org/ontology/writer> <http://dbpedia.org/resource/Maya_Forges> .
<http://dbpedia.org/resource/Monsters_vs._Aliens> <http://dbpedia.org/ontology/writer> <http://dbpedia.org/resource/Wallace_Woldarsky> .
<http://dbpedia.org/resource/Monsters_vs._Aliens> <http://dbpedia.org/ontology/writer> <http://dbpedia.org/resource/Rob_Letterman> .
<http://dbpedia.org/resource/Monsters_vs._Aliens> <http://dbpedia.org/ontology/writer> <http://dbpedia.org/resource/Jonathan_Albe_and_Glenn_Berger> .
```

- Open knowledge graph that provides structured information derived from Wikipedia
- Incorporates a manually created, shallow, cross-domain ontology, built from the most frequently used Wikipedia infoboxes
- The ontology helps in mapping structured data from Wikipedia into a consistent, usable format.
- Dbpedia is especially useful as it converts unstructured Wikipedia data into structured data, make relationships between different objects and concepts readily available as a knowledge graph.

Dbpedia - examples

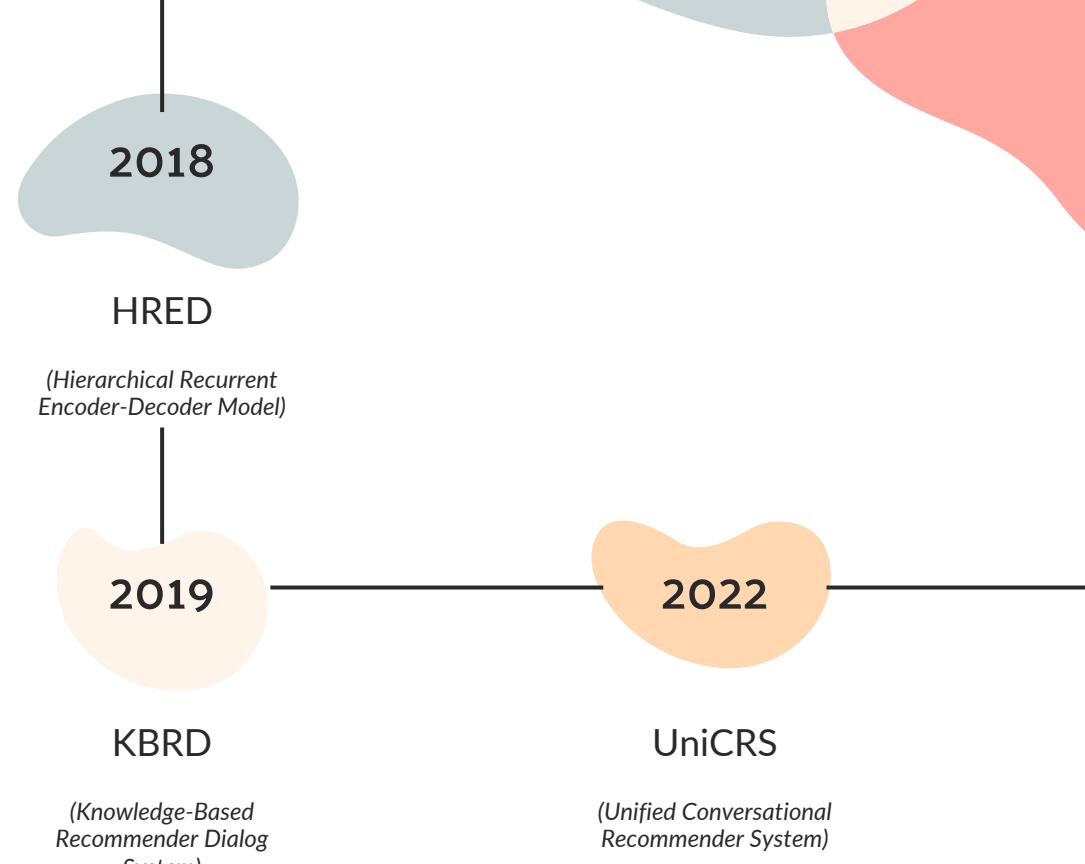
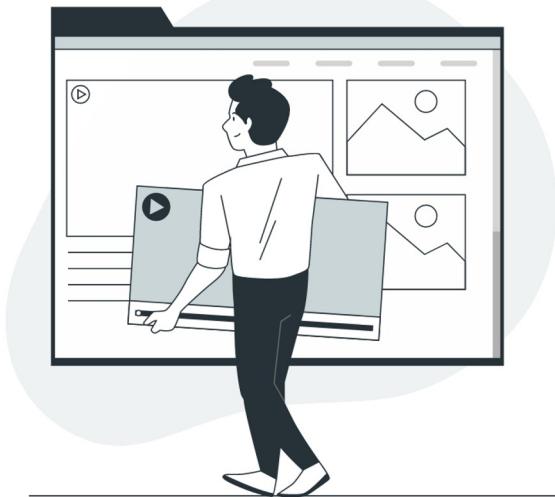
- **Entity Description:** DBpedia describes entities like "Barack Obama" in a structured way, including attributes such as birthdate, birthplace, and occupation. You can easily extract this structured information e.g., to compare different politicians' birthplaces.
- **Category Information:** DBpedia can be used to find entities of a certain category, e.g., all cities in Germany with more than 100,000 inhabitants. This is possible because DBpedia extracts the category information from Wikipedia and transforms it into a machine-readable format.
- **Link Analysis:** The links between entities in Wikipedia are transformed into semantic relationships in DBpedia, which enables the analysis of the network of entities. For example, you can analyze which philosophers influenced each other.
- **Multilingual Support:** DBpedia supports multilingual queries. For example, if you want to find out what "Paris" is called in different languages, DBpedia can provide the answer as it contains the language links from Wikipedia.
- **Integration with Other Datasets:** Because DBpedia uses standard Semantic Web formats, it can be easily linked with other datasets on the web. For instance, you can combine DBpedia with Geonames (a geographical database) to analyze the correlation between the population of a city (from DBpedia) and its altitude (from Geonames).

03



Algorithms

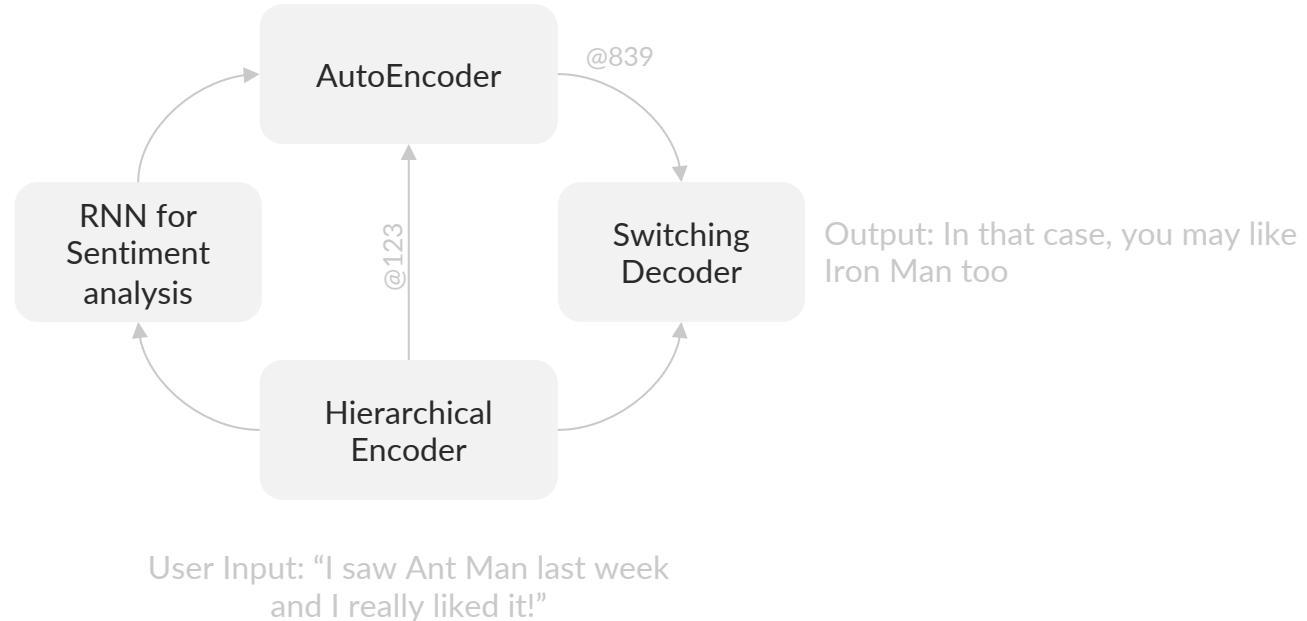
Brief timeline of research



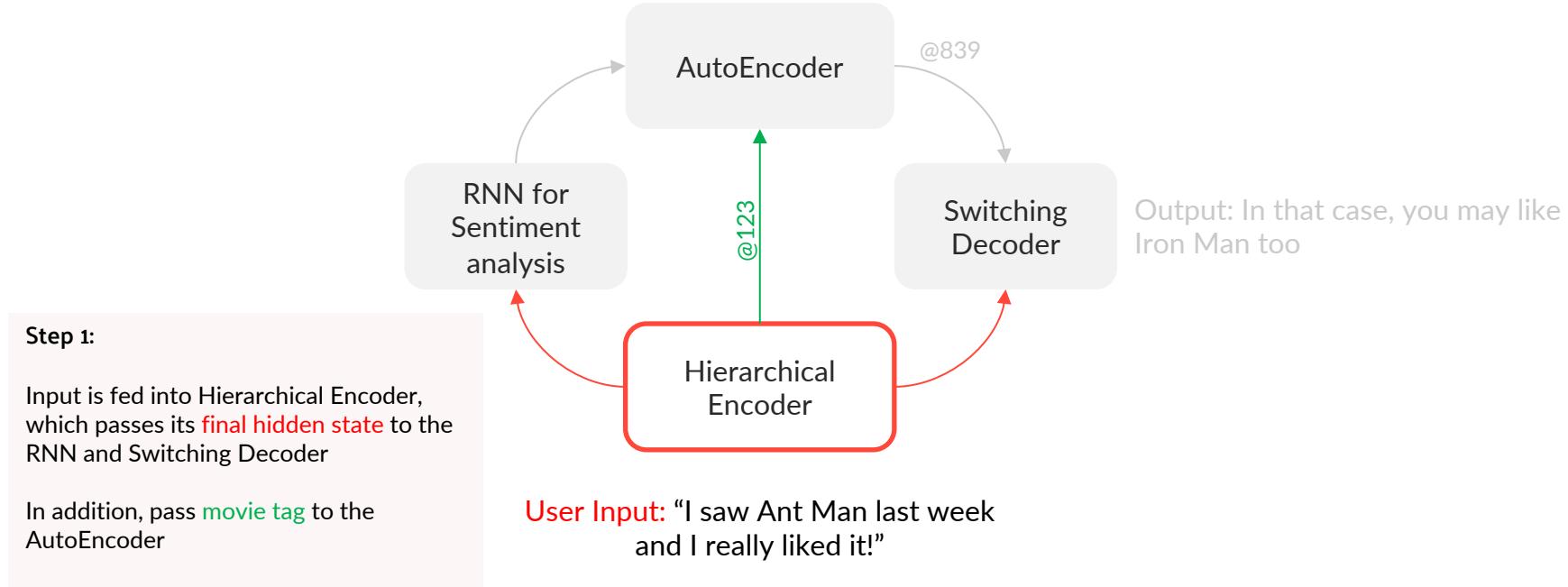
HRED (2018)

- Dynamically instantiated RNN for movie sentiment analysis
- Autoencoder recommender
- Decoder with movie recommendation switching mechanism

HRED Architecture



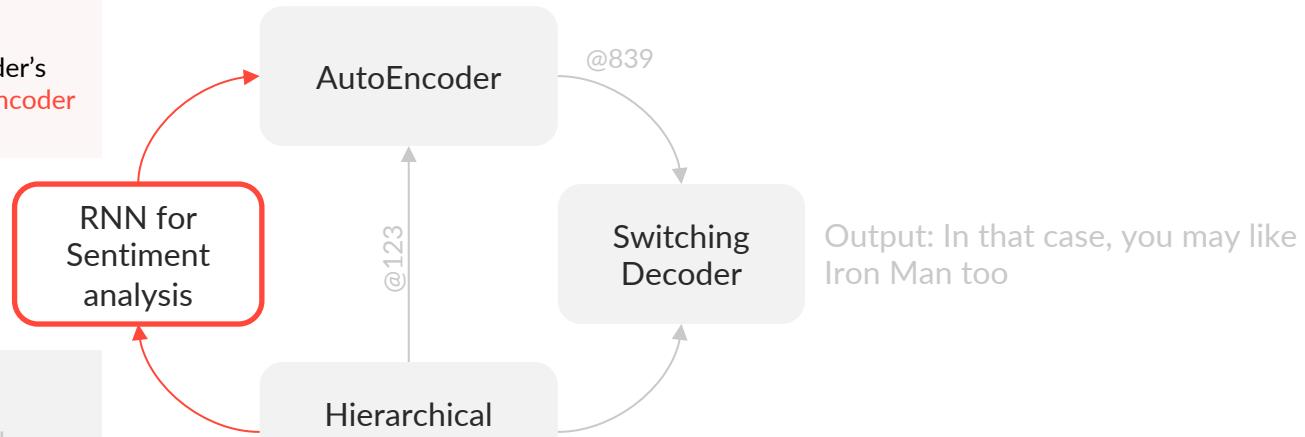
HRED Architecture



HRED Architecture

Step 2:

Sentiment is generated from encoder's hidden state, and passed to AutoEncoder



Step 1:

Input is fed into Hierarchical Encoder, which passes its final hidden state to the RNN and Decoder

In addition, pass movie tag to the AutoEncoder

User Input: "I saw Ant Man last week
and I really liked it!"

Output: In that case, you may like
Iron Man too

HRED Architecture

Step 2:

Sentiment is generated from encoder's hidden state, and passed to AutoEncoder

Step 1:

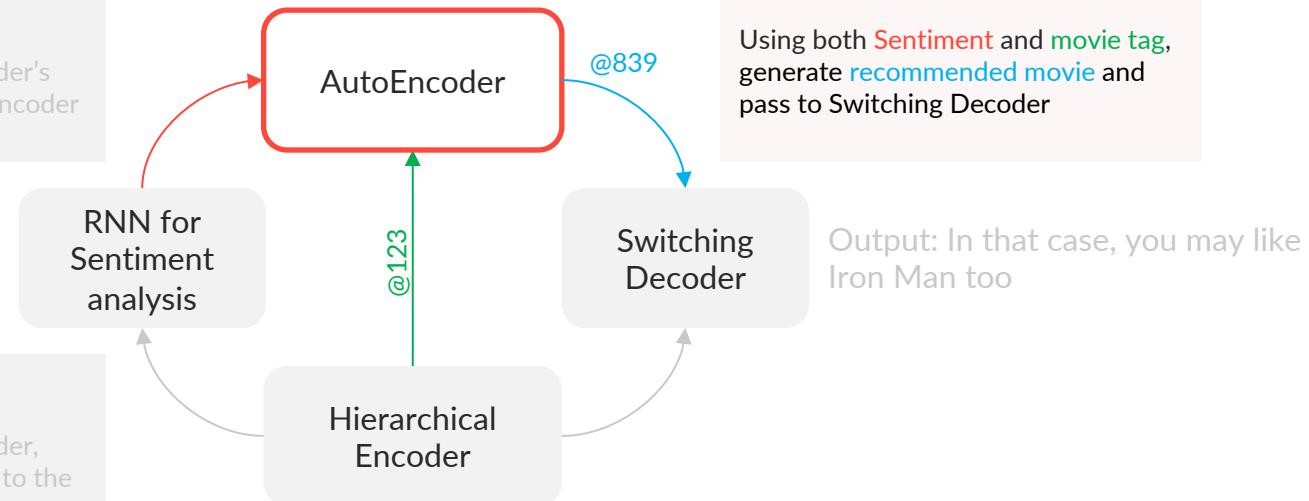
Input is fed into Hierarchical Encoder, which passes its final hidden state to the RNN and Decoder

In addition, pass movie tag to the AutoEncoder

Step 3:

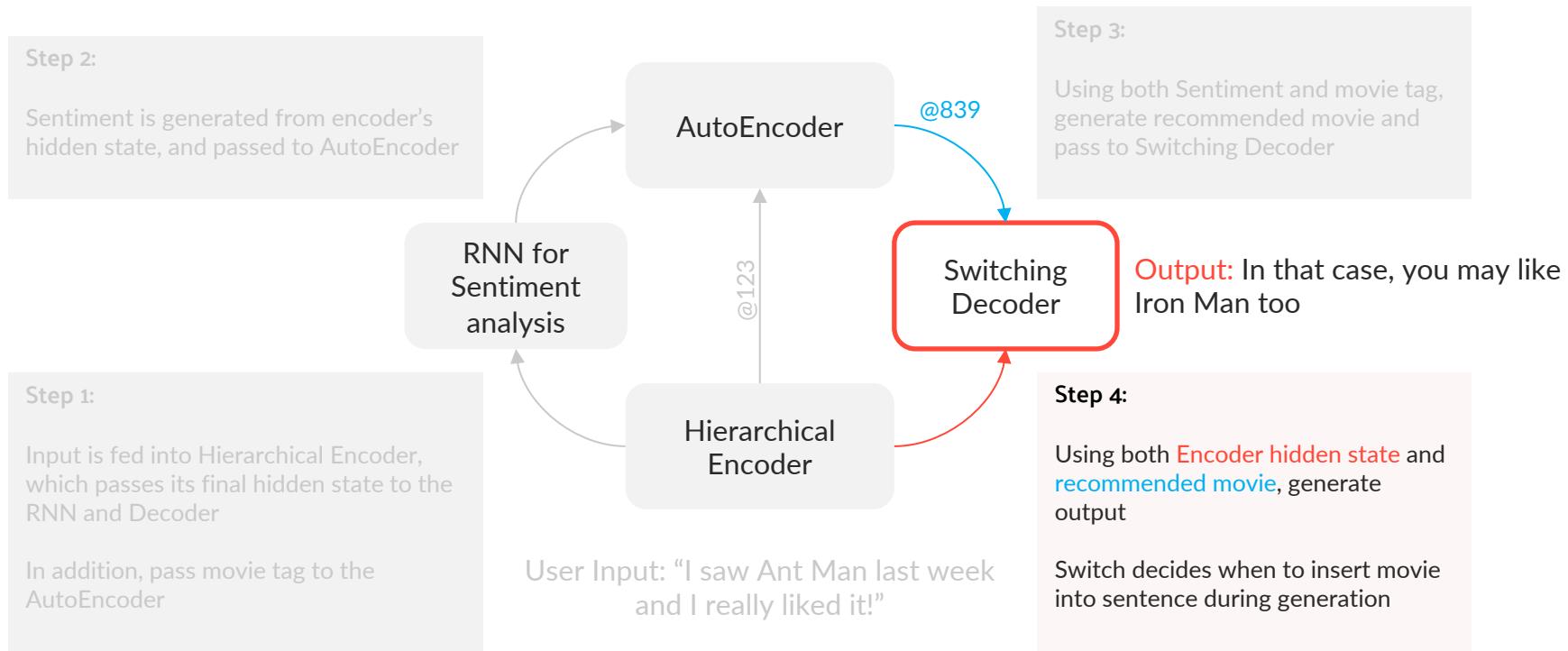
Using both **Sentiment** and **movie tag**, generate **recommended movie** and pass to Switching Decoder

Output: In that case, you may like Iron Man too



User Input: "I saw Ant Man last week and I really liked it!"

HRED Architecture



HRED - Issues

```
CONVERSATION 0
SEEKER: <s> hi i am looking for a movie like toy story 2 </s>
GENERATED T=1: <s> hi there </s>

CONVERSATION 1
SEEKER: <s> i really like harry potter and the <unk> of fire </s>
GENERATED T=1: <s> hello </s>

CONVERSATION 2
SEEKER: <s> i love all kinds of movies like the haunting </s>
GENERATED T=1: <s> hi there ! </s>

CONVERSATION 3
SEEKER: <s> world war z is my favorite </s>
GENERATED T=1: <s> hello </s>

CONVERSATION 4
SEEKER: <s> i loved step up revolution </s>
GENERATED T=1: <s> hello </s>

CONVERSATION 5
SEEKER: <s> hello </s>
GENERATED T=1: <s> hello </s>
SEEKER: <s> i loved step up revolution </s>
GENERATED T=1: <s> what kind of movies do you like ? </s>
```

Requires movie token

Unable to generate recommendations for inputs like “Any Sci-fi Movies”

Separate Models

Recommender and Language model works independently, leading to incoherence in outputs

Long Training Time

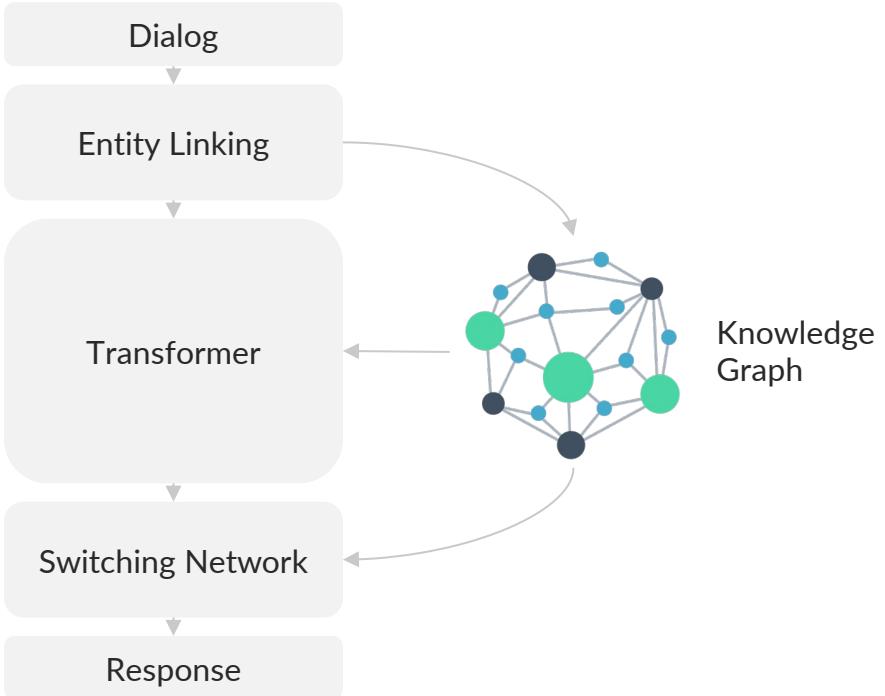
Took 1 day for 1 Epoch. Unable to train for long enough period to observe substantial results

Need better cooperation between Dialog system and Language model! -> Use KBRD, UniCRS

KBRD (2019)

- It's a conversational recommendation system, engaging users in dialogue to refine and personalize recommendations. It integrates the NLP and the recommender portion into a hybrid model.
- Utilizes a knowledge graph to leverage context, user preferences, and product features for providing recommendations.
- Employs collaborative filtering for capturing user-item interactions and rule-based methods for dialogue management.
- Capable of handling a wide range of natural language user queries, using these interactions to improve and personalize recommendations.
- Note that KBRD is unique in its dialogue-driven approach and its use of knowledge graphs to drive interactions and recommendations, which distinguishes it from more traditional recommendation systems.

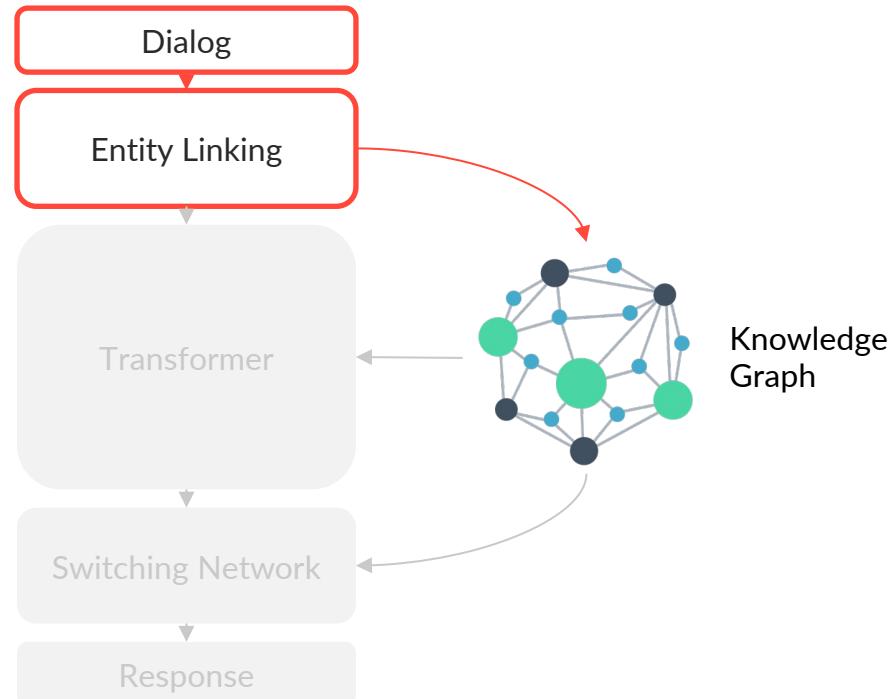
KBRD Architecture



KBRD Architecture

Step 1:

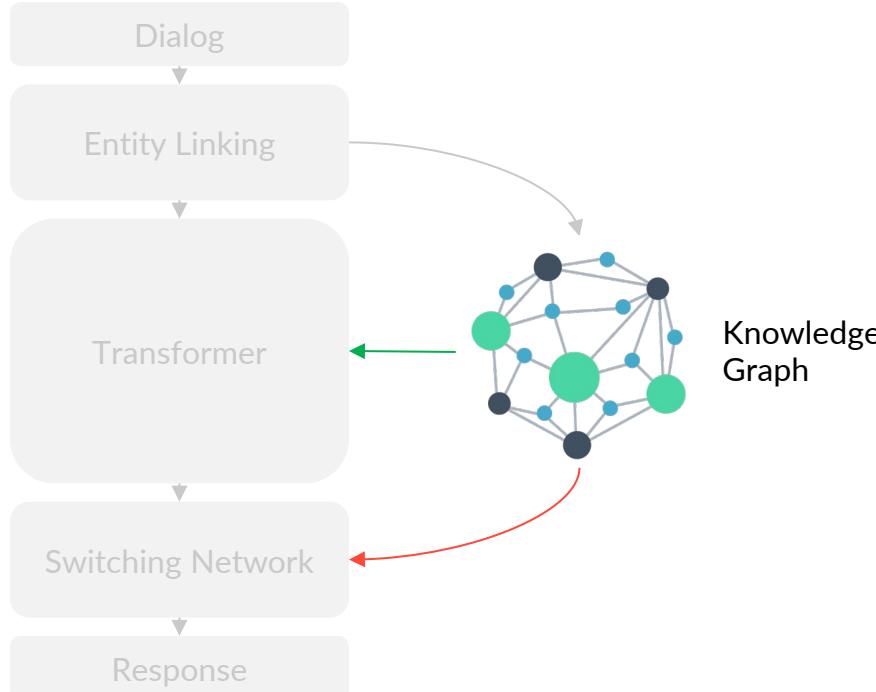
Movie Items and Informative Entities in Dialog input is linked to external knowledge graph (Dbpedia) using Entity Linking proposed by (Daiber et al., 2013)



KBRD Architecture

Step 1:

Movie Items and Informative Entities in Dialog input is linked to external knowledge graph (Dbpedia) using Entity Linking proposed by (Daiber et al., 2013)



Step 2:

Relational Graph Convolutional Networks (R-GCNs) is applied to Knowledge Graph to obtain recommendations

Final hidden state is passed to transformer that acts as a Vocabulary Bias

Final movie recommendation is passed to Switching Network

KBRD Architecture

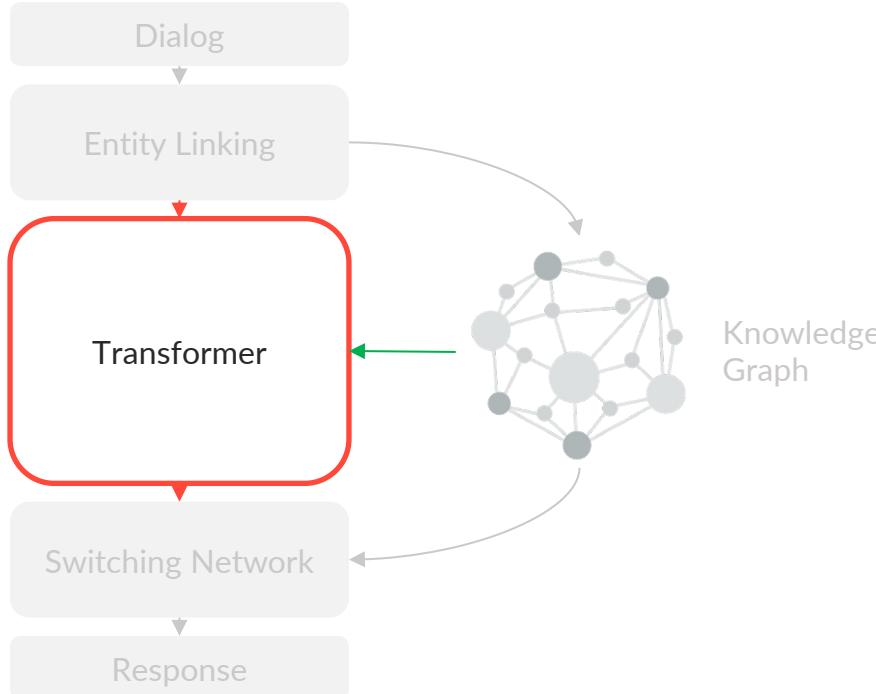
Step 1:

Movie Items and Informative Entities in Dialog input is linked to external knowledge graph (Dbpedia) using Entity Linking proposed by (Daiber et al., 2013)

Step 3:

Transformer use **input dialog** to generate output sequence

Final output layer is concatenated with **Vocabulary Bias** that influences the words that will be generated



Step 2:

Relational Graph Convolutional Networks (R-GCNs) is applied to Knowledge Graph to obtain recommendations

Final hidden state is passed to transformer that acts as a Vocabulary Bias

Final movie recommendation is passed to Switching Network

KBRD Architecture

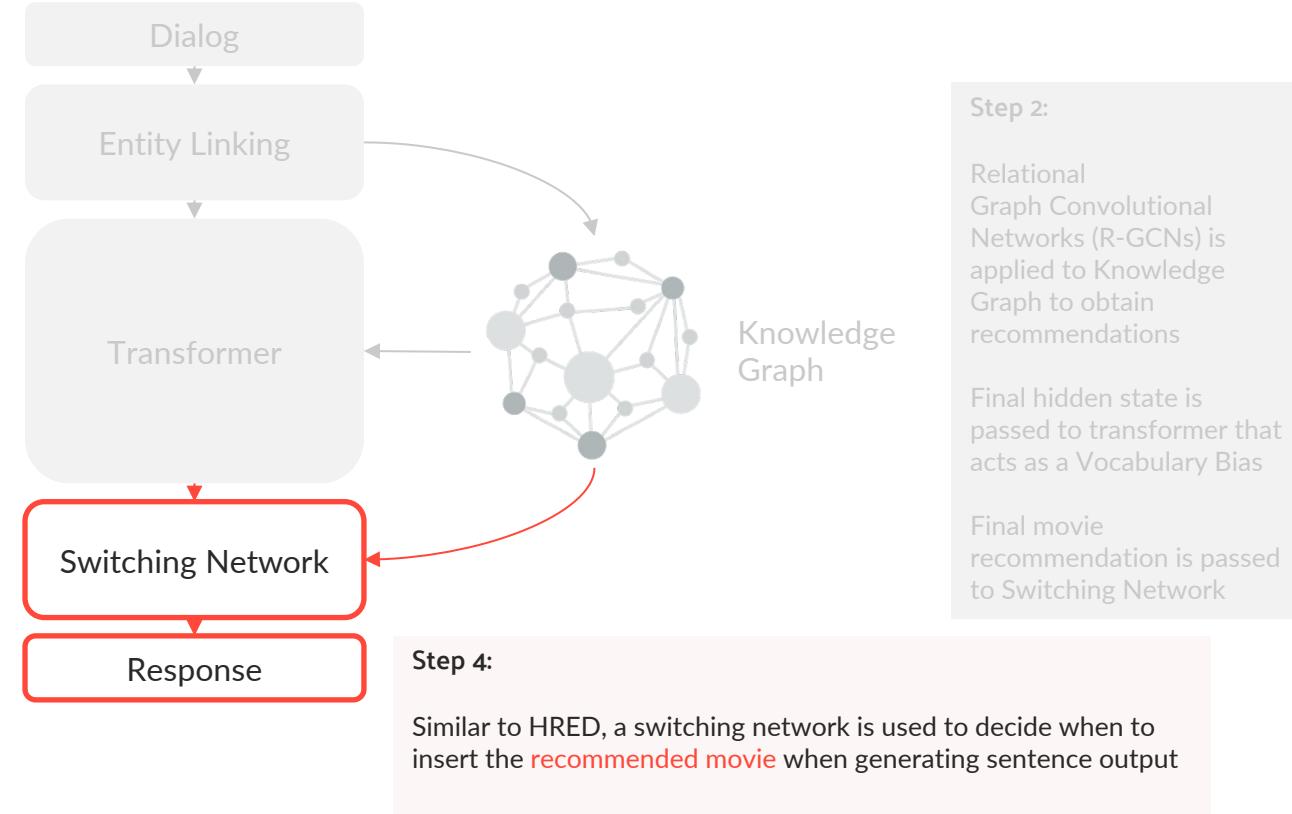
Step 1:

Movie Items and Informative Entities in Dialog input is linked to external knowledge graph (Dbpedia) using Entity Linking proposed by (Daiber et al., 2013)

Step 3:

Transformer use input dialog to generate output sequence

Final output layer is concatenated with Vocabulary Bias that influences the words that will be generated



Step 2:

Relational Graph Convolutional Networks (R-GCNs) is applied to Knowledge Graph to obtain recommendations

Final hidden state is passed to transformer that acts as a Vocabulary Bias

Final movie recommendation is passed to Switching Network

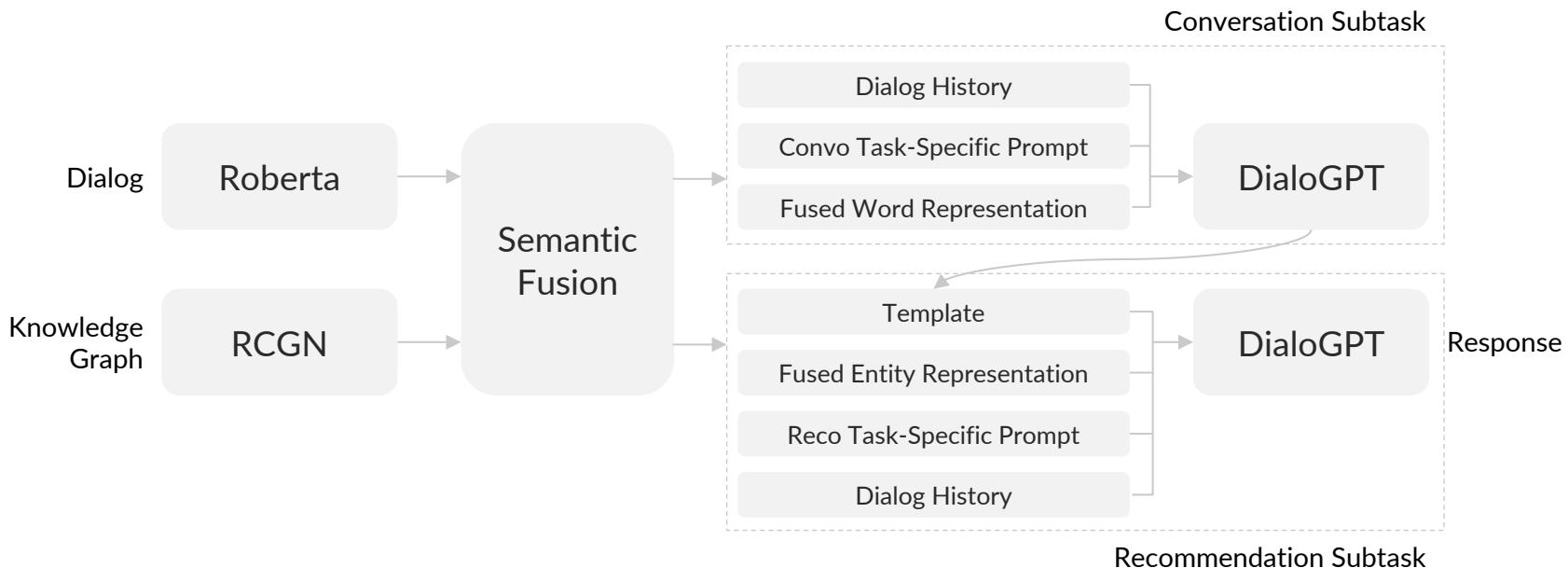
Step 4:

Similar to HRED, a switching network is used to decide when to insert the **recommended movie** when generating sentence output

UniCRS (2022)

- Unifies the recommendation and conversation subtasks into the prompt learning paradigm.
- Utilizes knowledge-enhanced prompts based on a fixed pre-trained language model (PLM) to fulfill both subtasks in a unified approach.
- Incorporates the generated response template as an important part of the prompt, to enhance the information interaction between the two subtasks.

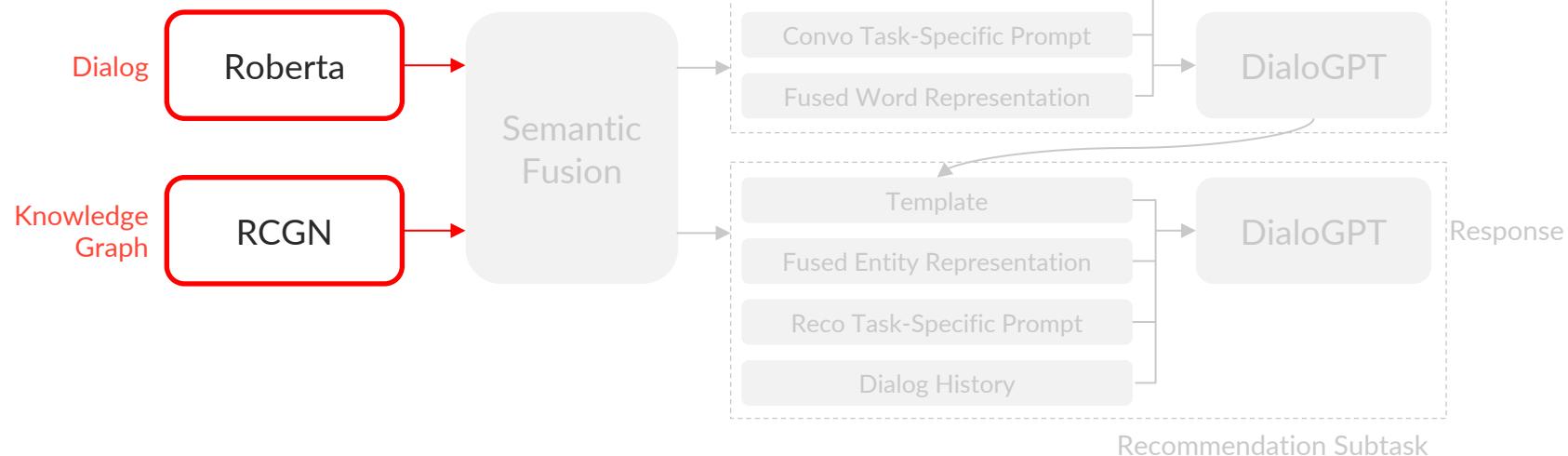
UniCRS



UniCRS

Step 1:

Roberta converts Dialog into **Word embeddings**. RCGN converts linked entities into **Entity Embeddings**.



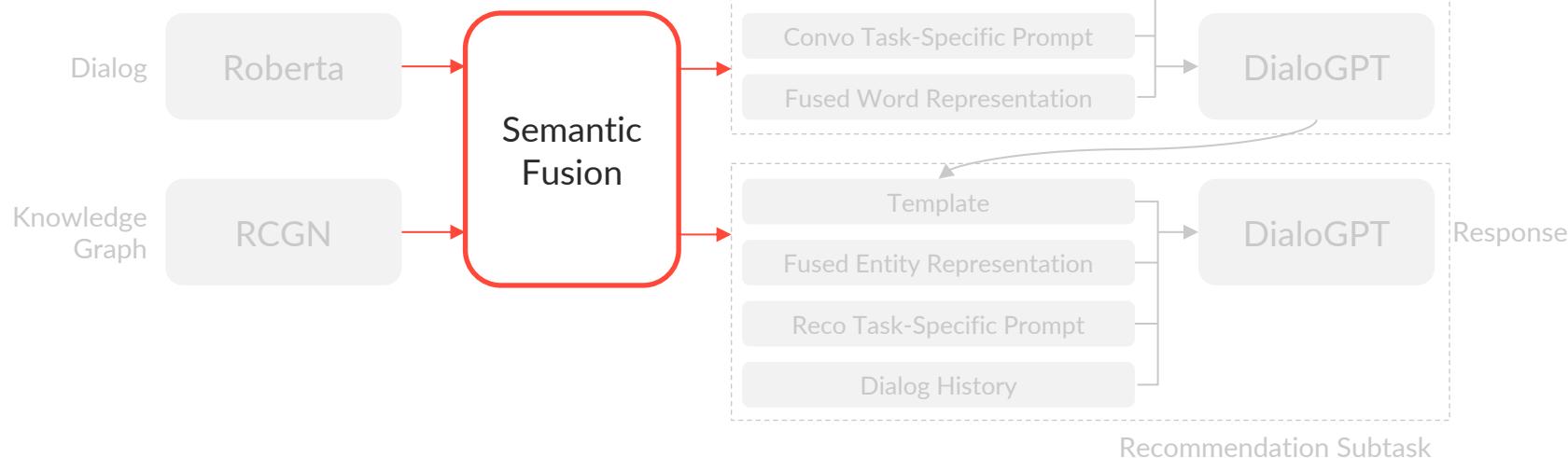
UniCRS

Step 1:

Roberta converts Dialog into Word embeddings. RCGN converts linked entities into Entity Embeddings.

Step 2:

Semantic Fusion is used to combine both **Embeddings** to obtain a **Fused Representation**



UniCRS

Step 1:

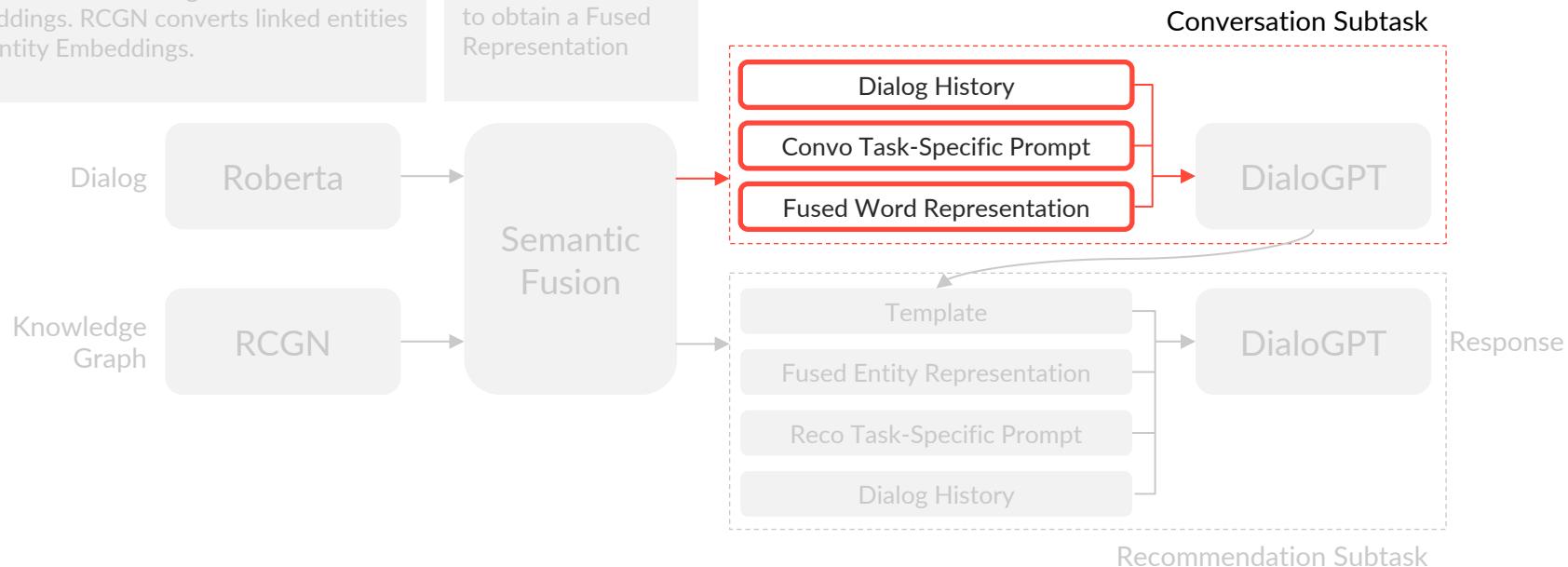
Roberta converts Dialog into Word embeddings. RCGN converts linked entities into Entity Embeddings.

Step 2:

Semantic Fusion is used to combine both Embeddings to obtain a Fused Representation

Step 3:

Dialog History, Conversation Task-Specific Prompt, and Fused Word Representation are concatenated as input into DialoGPT.



UniCRS

Step 1:

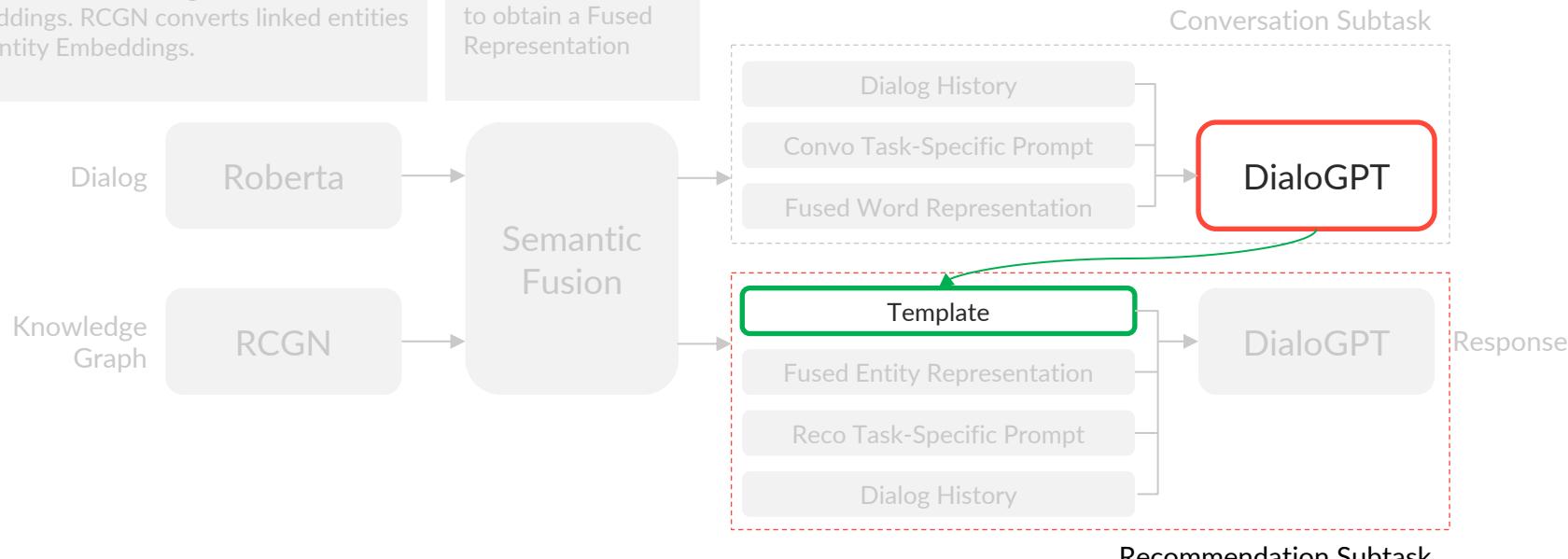
Roberta converts Dialog into Word embeddings. RCGN converts linked entities into Entity Embeddings.

Step 2:

Semantic Fusion is used to combine both Embeddings to obtain a Fused Representation

Step 3:

Dialog History, Conversation Task-Specific Prompt, and Fused Word Representation are concatenated as input into DialoGPT.



Step 4:

A template is generated and passed to Recommender Subtask

Template:

[Insert] is good for you. It is a classic thriller movie

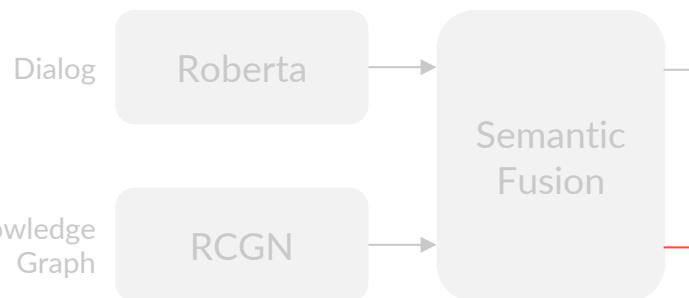
UniCRS

Step 1:

Roberta converts Dialog into Word embeddings. RCGN converts linked entities into Entity Embeddings.

Step 2:

Semantic Fusion is used to combine both Embeddings to obtain a Fused Representation



Step 3:

Dialog History, Conversation Task-Specific Prompt, and Fused Word Representation are concatenated as input into DialoGPT.

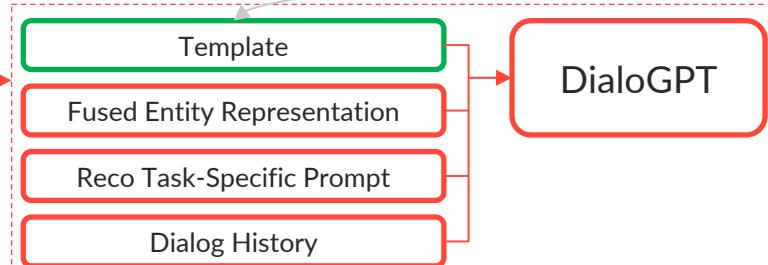
Conversation Subtask



Step 5:

Using the **Fused Entity Representation**, **Recommendation Task-Specific Prompt**, and **Dialog History**, DialoGPT generates a recommendation and constructs a **Response** together with the **Template**.

IT(2017) is good for you. It is a classic thriller movie



Recommendation Subtask

Step 4:

A template is generated and passed to Recommender Subtask

Template:

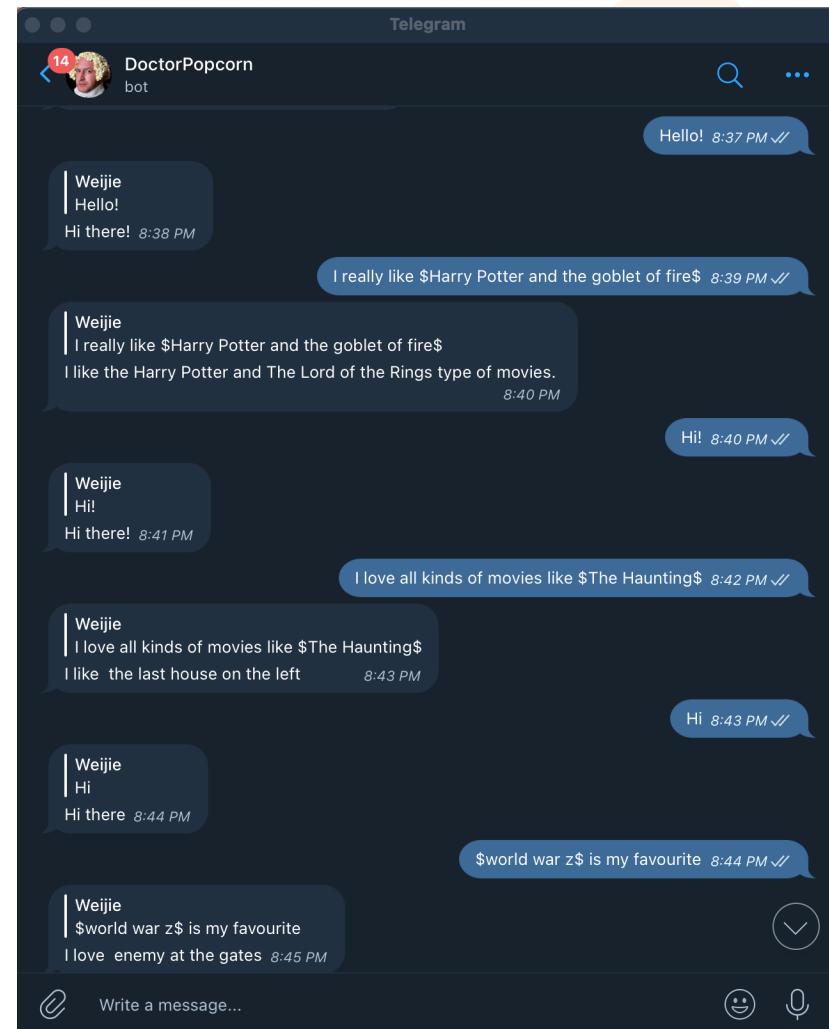
[Insert] is good for you. It is a classic thriller movie

03

Evaluations



Output - UniCRS



Do the movie recommendations make sense?

User: I really like \$Harry Potter and the goblet of fire\$

Bot: I like the Harry Potter and The Lord of the Rings type of movies.



Harry Potter and the Goblet of Fire (2005)
12 17/11/2005 (SG) • Adventure, Fantasy, Family • 2h 37m

78⁺ User Score

Dark And Difficult Times Lie Ahead.

Overview

When Harry Potter's name emerges from the Goblet of Fire, he becomes a competitor in a grueling battle for glory among three wizarding schools—the Triwizard Tournament. But since Harry never submitted his name for the Tournament, who did? Now Harry must confront a deadly dragon, fierce water demons and an enchanted maze only to find himself in the cruel grasp of He Who Must Not Be Named.

Mike Newell
Director

J.K. Rowling
Novel

Steve Kloves
Screenplay

Now Streaming Watch Now



The Lord of the Rings: The Fellowship of the Ring (2001)
PG-13 20/12/2001 (SG) • Adventure, Fantasy, Action • 2h 59m

84⁺ User Score

One ring to rule them all

Overview

Young hobbit Frodo Baggins, after inheriting a mysterious ring from his uncle Bilbo, must leave his home in order to keep it from falling into the hands of its evil creator. Along the way, a fellowship is formed to protect the ringbearer and make sure that the ring arrives at its final destination: Mt. Doom, the only place where it can be destroyed.

Peter Jackson
Director, Screenplay

J.R.R. Tolkien
Novel

Philippa Boyens
Screenplay

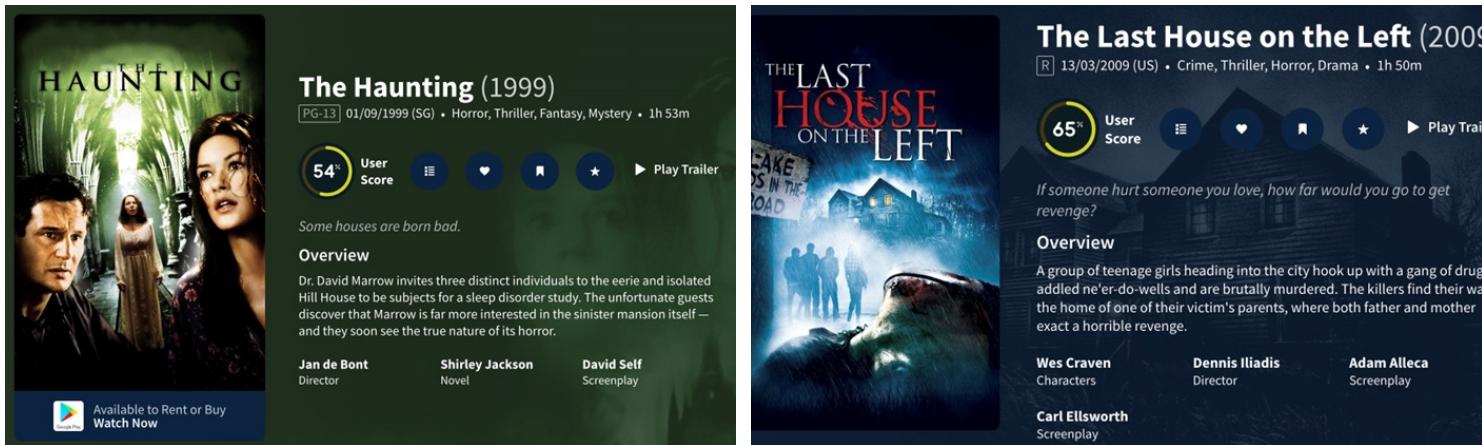
Fran Walsh
Screenplay

Amazon Prime Video Now Streaming Watch Now

Do the movie recommendations make sense?

User: I love all kinds of movies like \$The Haunting\$

Bot: I like the last house on the left



Do the movie recommendations make sense?

User: \$world war z\$ is my favourite

Bot: I love enemy at the gates



Need for a more systematic form of evaluation! -> Metrics, Rule-based Approach

Recommender Evaluation

- Higher recall is better
- Recall@1 is more significant than recall@50

$$Recall = \frac{\# \text{ of relevant recommendations}}{\# \text{ of all possible relevant items}}$$

	Recall@1	Recall@10	Recall@50
KBRD	0.02656	0.1699	0.3568
UniCRS	0.0238	0.1403	0.3321
WMF	0.0078	0.0836	0.2487
EASE ^R	0.0269	0.1358	0.3073
HRED	0	0	0

Recommender Evaluation

- Other Metrics

NDCG evaluates ranked search results or recommendation lists by considering both relevance and position, providing a comprehensive ranking performance measure.

NDCG assigns higher scores to relevant items at the top of the list, promoting rankings that prioritize highly relevant results, while its normalization allows for easy comparisons across different datasets and evaluation scenarios.

	NDCG@1	NDCG@10	NDCG@50
KBRD	0.02656	0.08743	0.1288
UniCRS	0.0238	0.07272	0.1147

Recommender Evaluation

- Other Metrics

MRR focuses on the highest-ranked relevant item, giving higher scores to lists where relevant items appear closer to the top. It provides a simplified evaluation of ranking quality, making it particularly useful in scenarios where the position of the first relevant item is crucial.

	MRR@1	MRR@10	MRR@50
KBRD	0.02656	0.06246	0.07133
UniCRS	0.0238	0.0523	0.0611

Language Evaluation

Distinct@k is a metric used to assess the diversity of the top-k items in a ranked list. It measures the proportion of unique items among the first k recommendations, indicating the variety or novelty of the suggested results.

Distinct@k encourages systems to provide diverse recommendations by assigning higher scores when a larger number of distinct items are present within the top-k recommendations. It complements other evaluation metrics by capturing the diversity aspect of recommendation algorithms.

	Distinct@1	Distinct@2	Distinct@3	Distinct@4
UniCRS	0.1003	0.2959	0.4829	0.6292
	BLEU@1	BLEU@2	BLEU@3	BLEU@4
UniCRS	0.1915	0.0454	0.0198	0.0116

Human Evaluation

Human evaluation can be biased and subjective. To minimize these factors, we decided to use a binary scoring system.

- Curated a set of binary tasks for Recommender to perform
- If task is performed successfully, we give a score of 1, else 0
- Tasks are separated into different Categories
- There are 5 tasks in each category, which gives each category a total of 5 points

Human Evaluation - Tasks



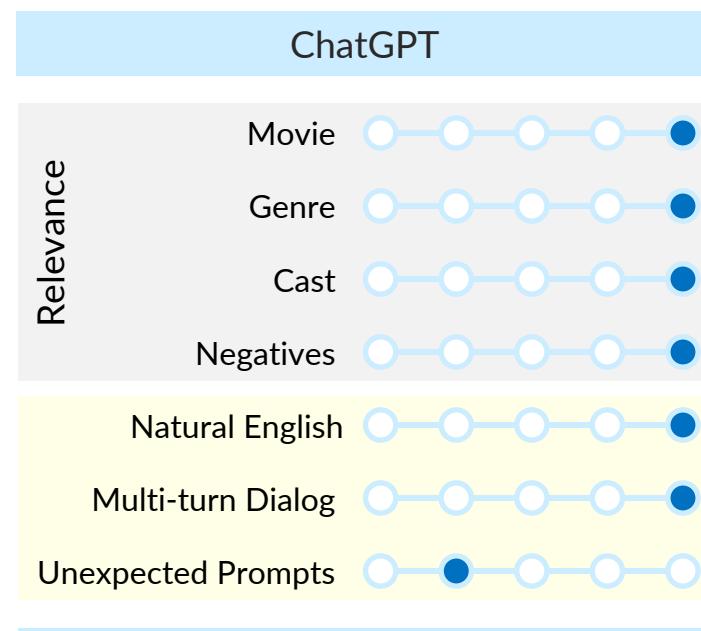
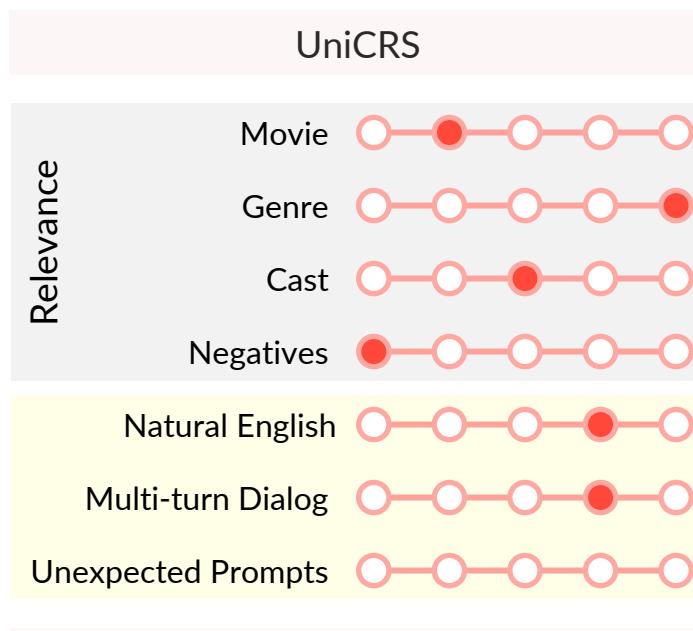
Input	Output	Score
Can you recommend me a movie like Superman ?	Sure, have you seen Batman ?	1
Any good psychological horror like Shutter Island ?	Cast away is a good one.	0
Is there a nice animated film like Kungfu Panda ?	How about The Incredibles ?	1
I really like the Transformers movies!	Then you will definitely like Pacific Rim !	1
I am looking for a movie like Independence Day .	I'd definitely recommend War of the Worlds .	1

Human Evaluation - Categories

Movies	Is recommended item relevant to mentioned movie given positive/negative sentiments?	Natural English	Are replies given as natural language responses?
Genre	Is recommended item relevant to mentioned genre given positive/negative sentiments?	Multi-turn Dialog	Is the recommender able to maintain multi-turn conversations?
Cast	Is recommended item relevant to mentioned cast given positive/negative sentiments?	Unexpected Prompts	Is the recommender able to guide users to correct inputs? (spelling errors, blanks)

*7th category is the negatives

Human Evaluation - Scores



Prompt: Act like a movie recommender system. Recommend me one movie based on my inputs.

Limitations and Future Considerations

Model Performance

More training time, data, and epochs
Leverage better performing LLMs for Dialog

Business Application

More data in other fields besides movies
Lighter weight, deployable solution

Evaluation

More curated tasks
More holistic categories with various levels of difficulties

Any more?

any more?

Any more?

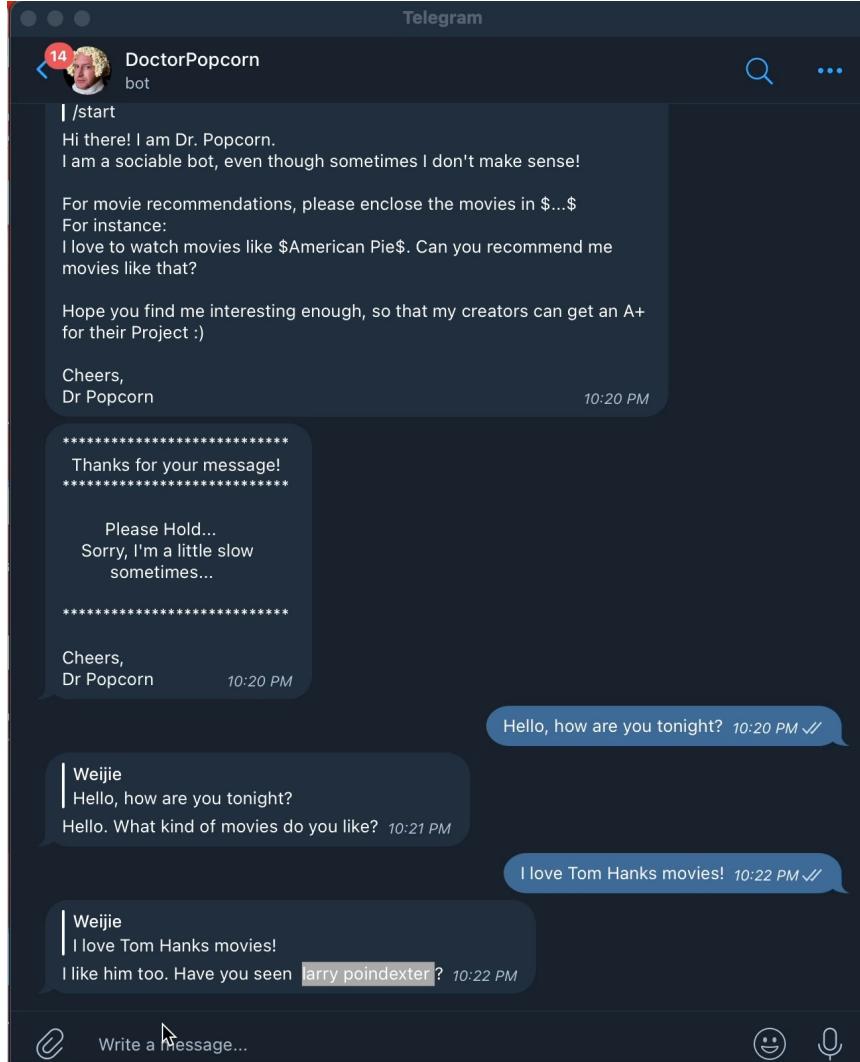
any more?

04

ChatBot Demonstration



Sample Video

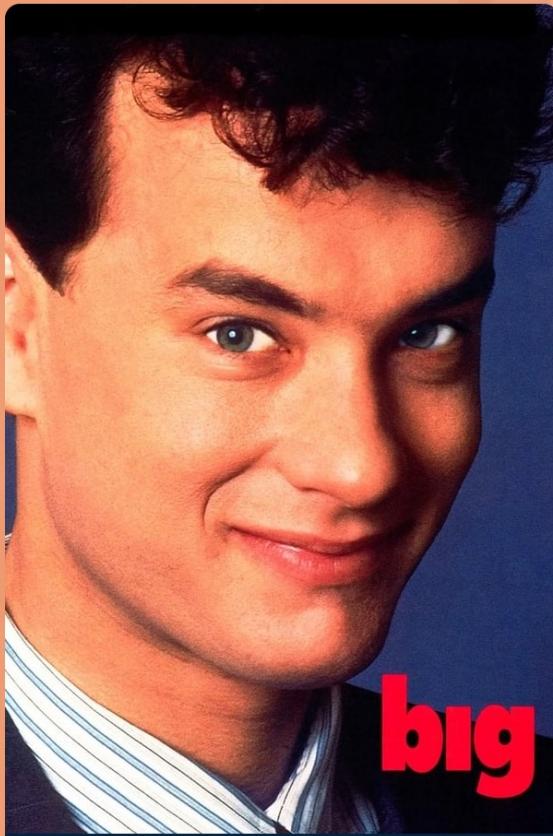


UniCRS



04a

Positives of the Chat Bot



Now Streaming
[Watch Now](#)

Big (1988)

PG

03/06/1988
(US)

Fantasy, Drama, Comedy, Romance, Family

1h
44m



User Score



[Play Trailer](#)

Have you ever had a really big secret?

Overview

When a young boy makes a wish at a carnival machine to be big—he wakes up the following morning to find that it has been granted and his body has grown older overnight. But he is still the same 13-year-old boy inside. Now he must learn how to cope with the unfamiliar world of grown-ups including getting a job and having his first romantic encounter with a woman.

Penny Marshall
Director

Anne Spielberg
Screenplay

Gary Ross
Screenplay

Positives of the ChatBot

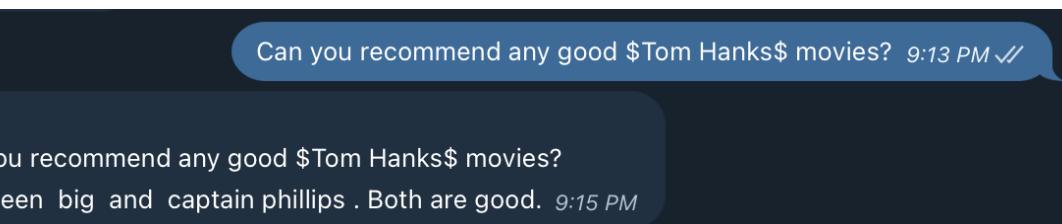
Other instances of good replies:



Understands relationship between genre and movies



Understands relationship amongst movies

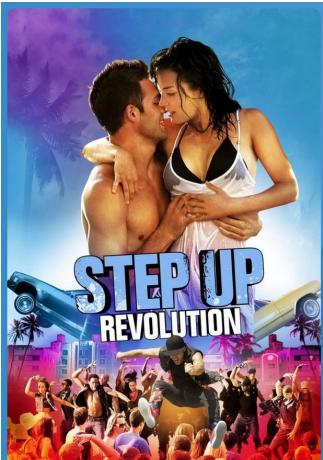


Understands relationship between cast and movies.

Able to recommend more than 1 movie

Positives of the ChatBot

*Sometimes, it understands
contrastive relationship*



Step Up Revolution (2012)

PG-13 30/08/2012 (SG) • Music, Drama, Romance • 1h 39m



User Score



► Play Trailer

One Step Can Change Your World

Overview

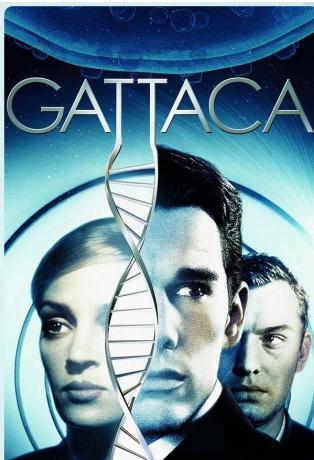
Emily arrives in Miami with aspirations to become a professional dancer. She sparks with Sean, the leader of a dance crew whose neighborhood is threatened by Emily's father's development plans.

Scott Speer
Director

Amanda Brody
Screenplay

Jenny Mayer
Writer

Available to Rent or Buy
Watch Now



Gattaca (1997)

PG-13 19/02/1998 (SG) • Thriller, Science Fiction, Mystery, Romance • 1h 46m



User Score



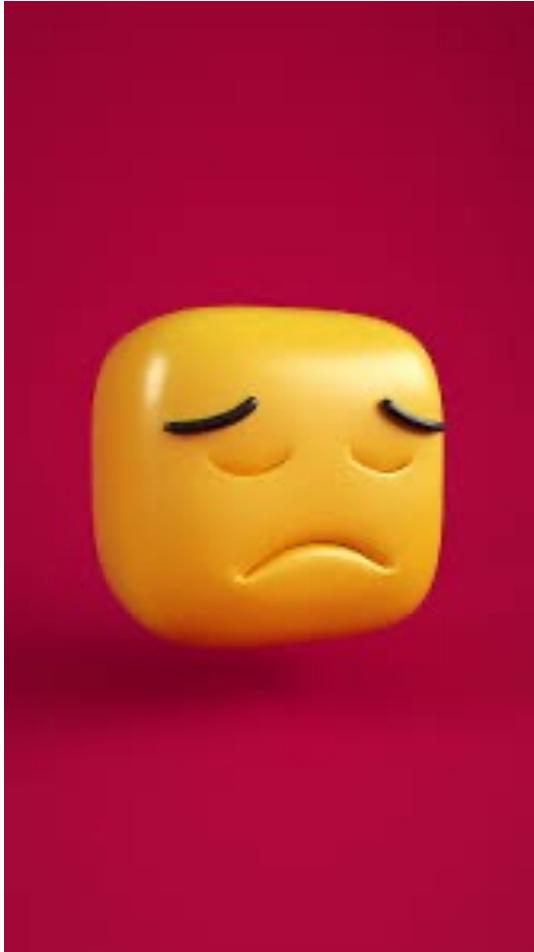
► Play Trailer

There is no gene for the human spirit.

Overview

In a future society in the era of indefinite eugenics, humans are set on a life course depending on their DNA. Young Vincent Freeman is born with a condition that would prevent him from space travel, yet is determined to infiltrate the GATTACA space program.

Andrew Niccol
Director, Writer

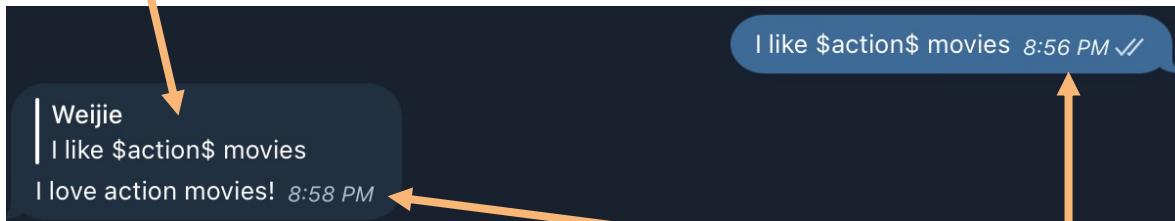


04b

Negatives of the Chat Bot

Negatives of the Chat Bot

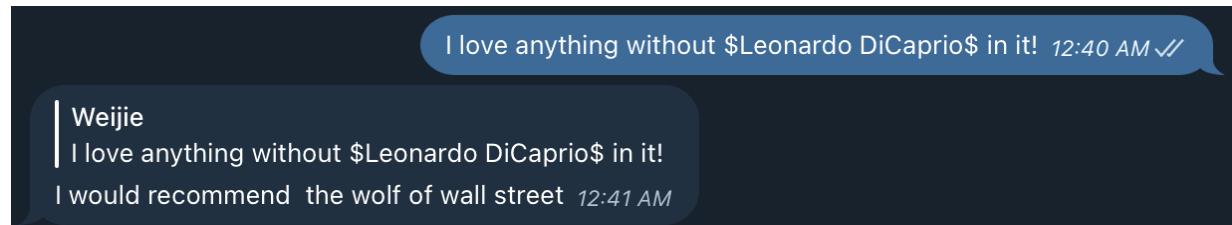
At times, it churns out
unhelpful replies.



At times, it can be **really slow.**

Negatives of the Chat Bot

Cannot differentiate genre from movies



Sometimes, it cannot discern between like and dislike connotations

Negatives of the Chat Bot

I am actually looking for movies for my 12 yr olds sleepover

1:33 AM ✓✓

Weijie

I am actually looking for movies for my 12 yr olds sleepover

I would recommend francesca ciardi

1:33 AM

***Recommends an actress
instead of movies.***

***Unable to answer
complex questions.***

Negatives of the Chat Bot

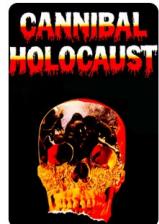


Francesca Ciardi

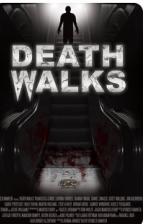
Biography

We don't have a biography for Francesca Ciardi.

Known For



Cannibal Holocaust



Death Walks



Fatal Fix



Cursed Films



Farewell Moscow



Safari

Acting

All ▾ Department ▾

2020 ⚡ Cursed Films
(1 episode)

2016 ⚡ Death Walks
as Lucrezia

1991 ⚡ Safari

1987 ⚡ Farewell Moscow

1980 ⚡ Fatal Fix
as Pusher

The movies the actress acts in are totally not appropriate for 12-year-olds.

Personal Info

Known For
Acting

Known Credits
6

Gender
Female

Birthdate
1954-07-26 (68 years old)

Place of Birth
Rome, Italy

Thanks!

Do you have any questions?



CREDITS: This presentation template was created by Slidesgo, including icons by Flaticon, infographics & images by Freepik and illustrations by Storyset