Retrieval-Augmented Generation for Immersive Formal Dialogue

Mark Snaith
School of Computing, Engineering & Technology
Robert Gordon University, Aberdeen
UK

Simon Wells
School of Computing, Engineering & The Built Environment
Edinburgh Napier University
UK

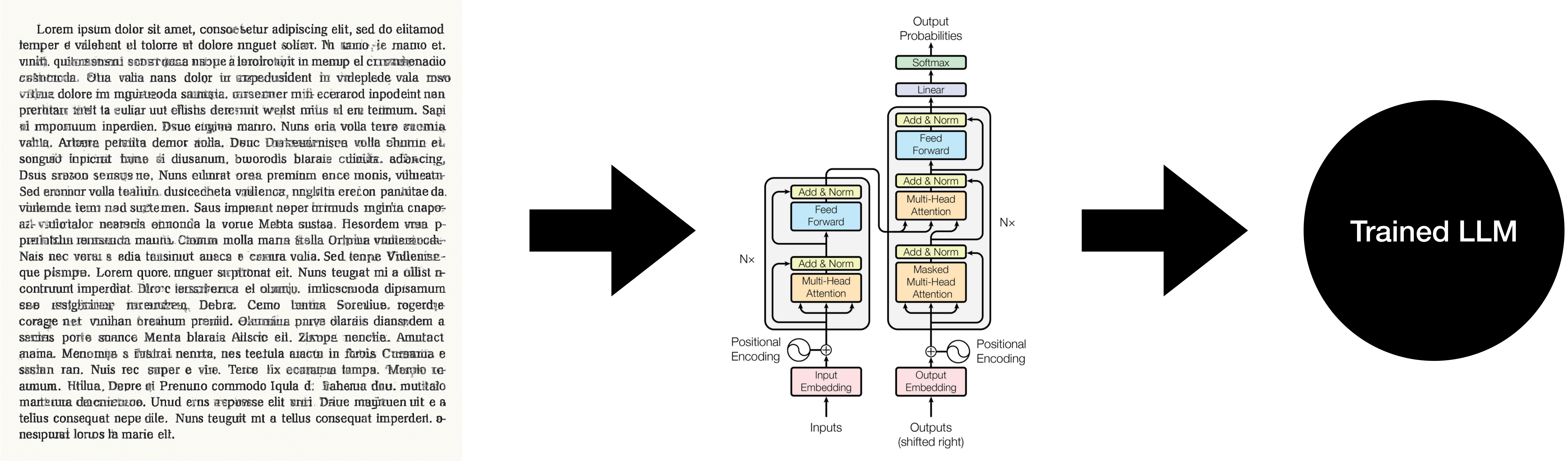
5th European Conference on Argumentation - ECA 2025
26th September 2025
Warsaw, Poland

Overview

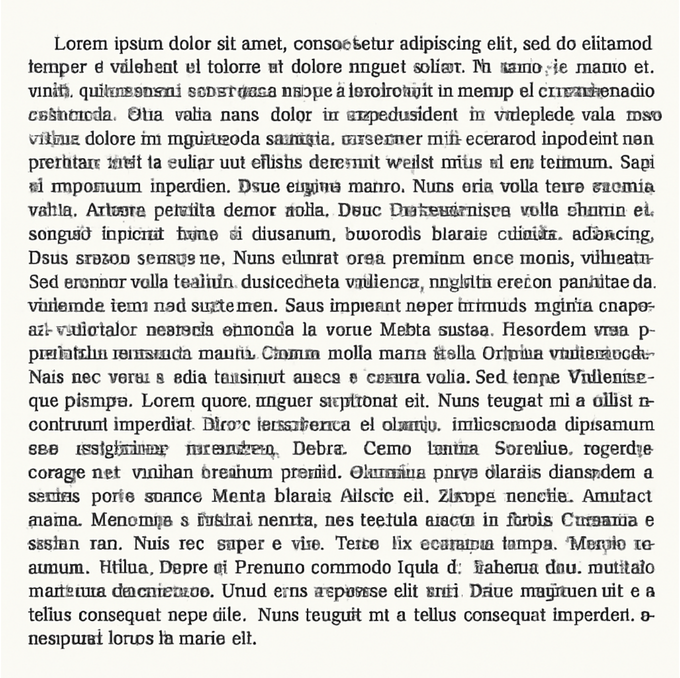
- A brief introduction to LLMs, RAG and Formal Dialogue
- Formal Dialogue vs LLMs vs LLMs + RAG + Formal Dialogue
- RAG-based pipeline

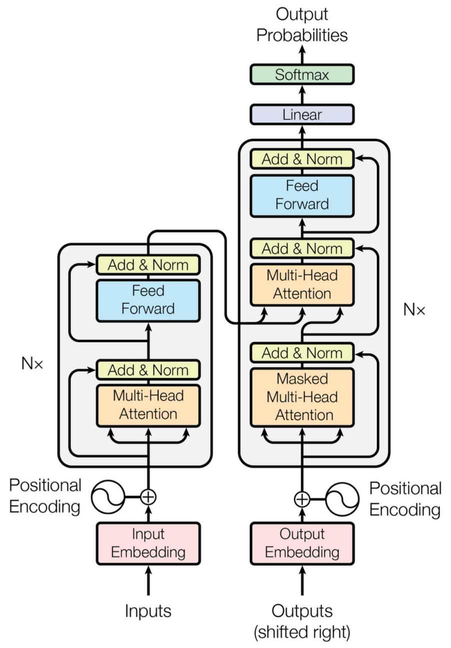
Large Language Models

Large Language Models (LLMs)



Large Language Models (LLMs)





Trained LLM

Prediction and generation of words in context

Natural conversational responses

But…can go off-topic…
Please give me a recipe for cupcakes

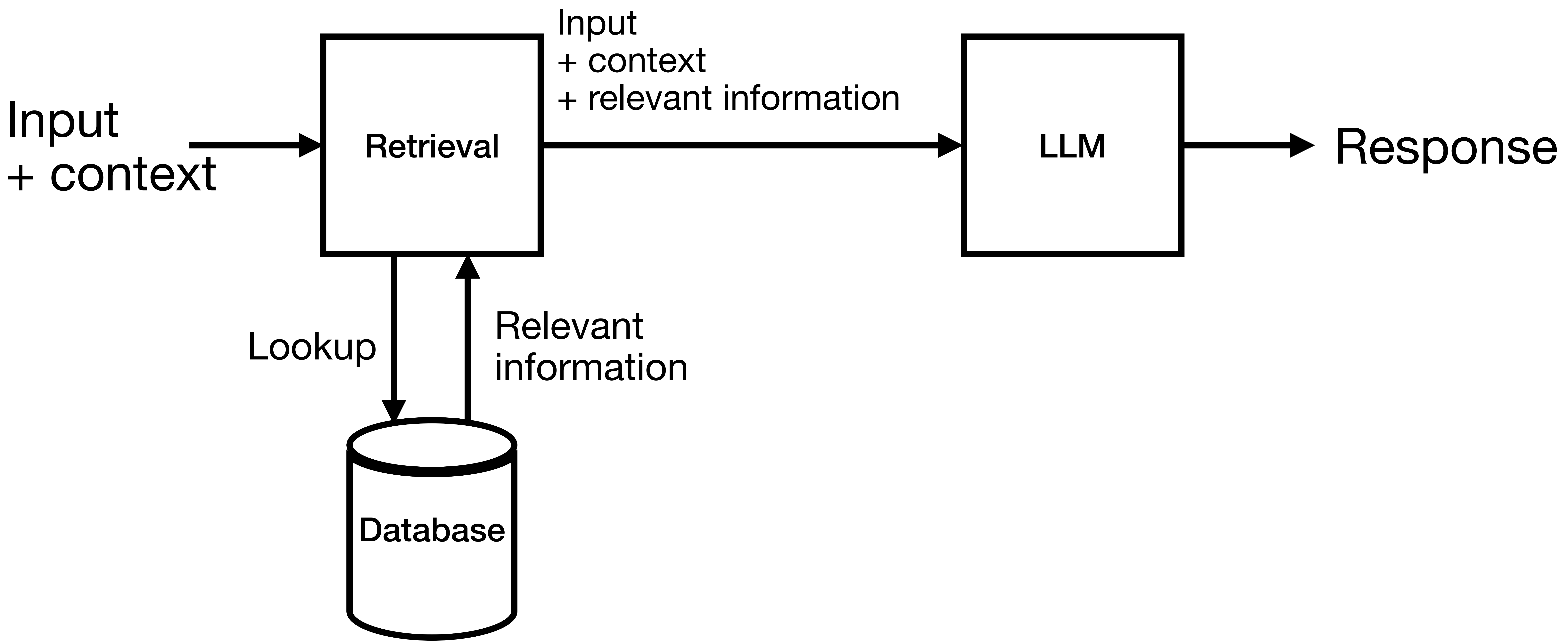
Can also hallucinate - responses not grounded in truth

Retrieval-Augmented Generation (RAG)

Grounds an LLM's output by retrieving relevant information from external sources before generating a response

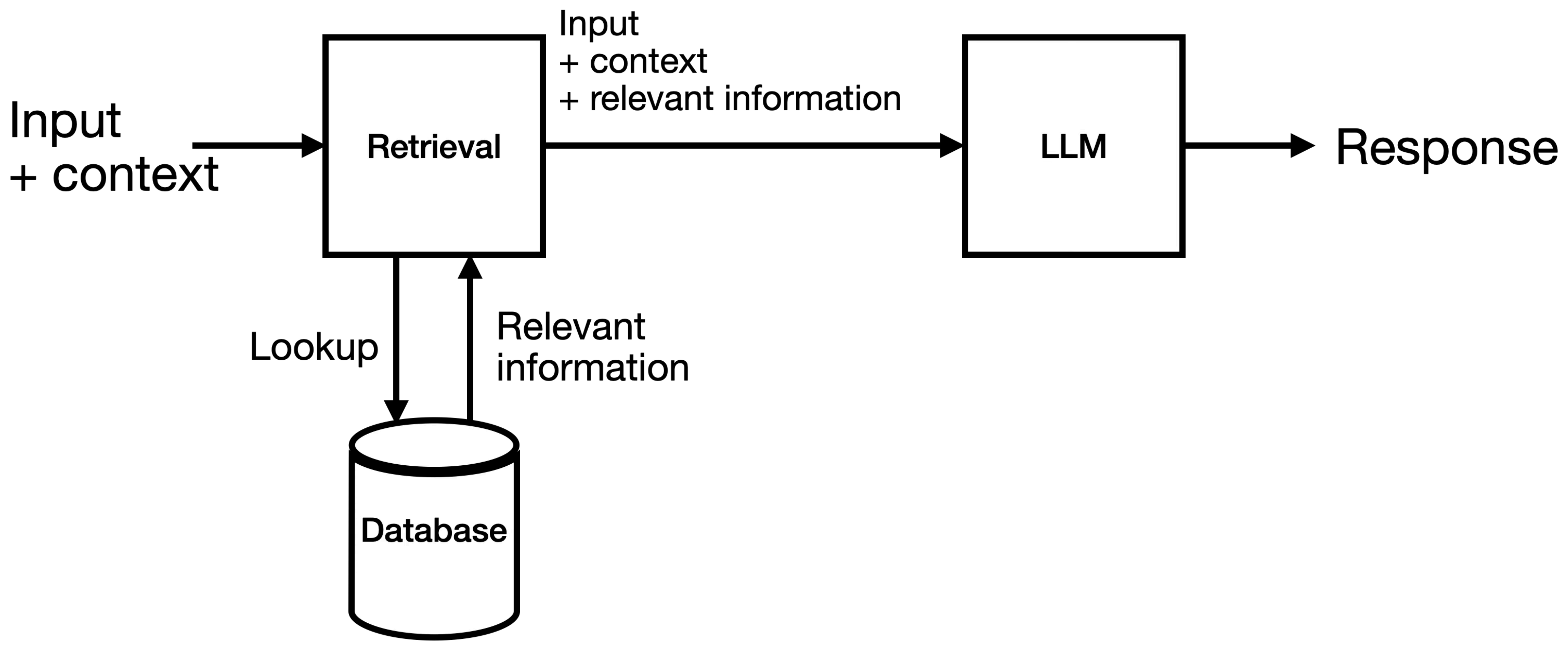
Retrieval-Augmented Generation (RAG)

Grounds an LLM’s output by retrieving relevant information from external sources before generating a response



Retrieval-Augmented Generation (RAG)

Grounds an LLM's output by retrieving relevant information from external sources before generating a response



Common use-cases:

- Customer support: LLM retrieves specific company policies etc. to provide bespoke answer
- Summarisation: provide documents and ask LLM to generate a summary

Formal Dialogue

Model conversations as structured, rule-governed exchanges aimed at specific goals (e.g. persuasion, inquiry, negotiation)

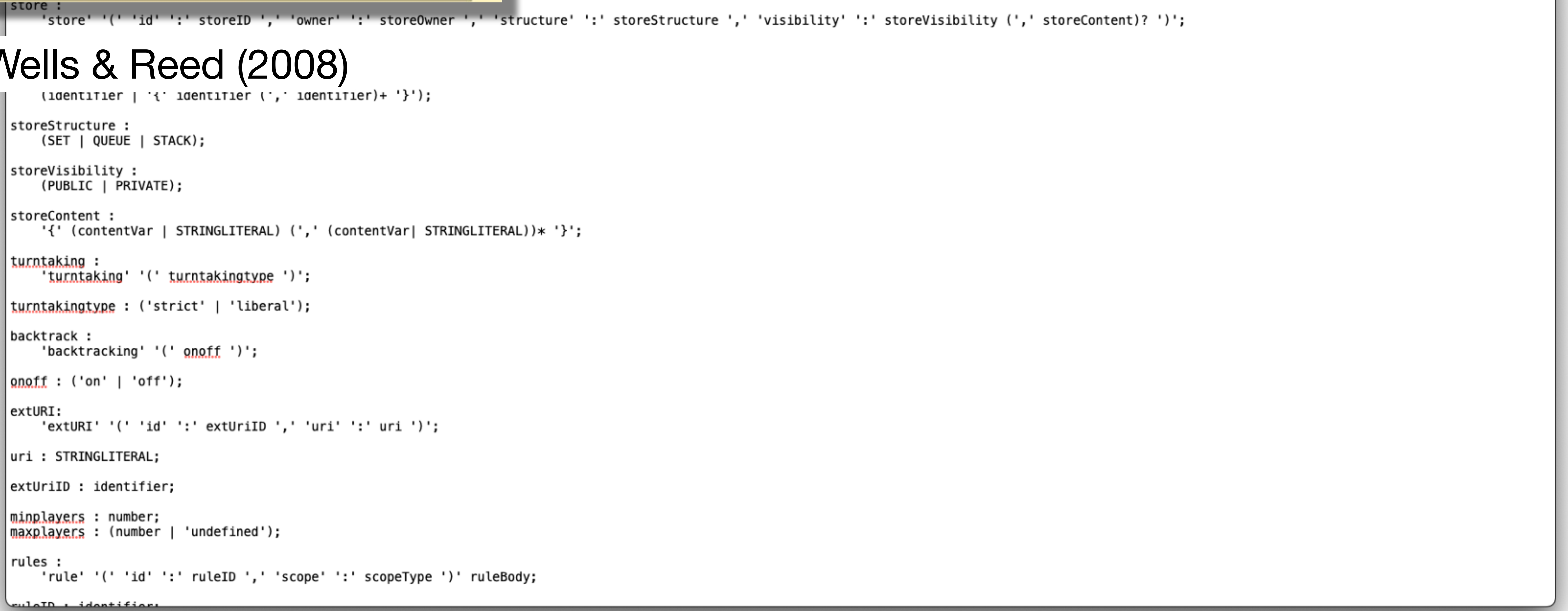
Formal Dialogue

Model conversations as structured, rule-governed exchanges aimed at specific goals (e.g. persuasion, inquiry, negotiation)

- Rules for structure, turn-taking, locutions, commitment, termination etc.
- Participants advance valid moves - speech act + content
- Used in human-agent and agent-agent communication
- Can seem rigid and lack expressiveness in human-agent applications

dgdl.g4 — Edited

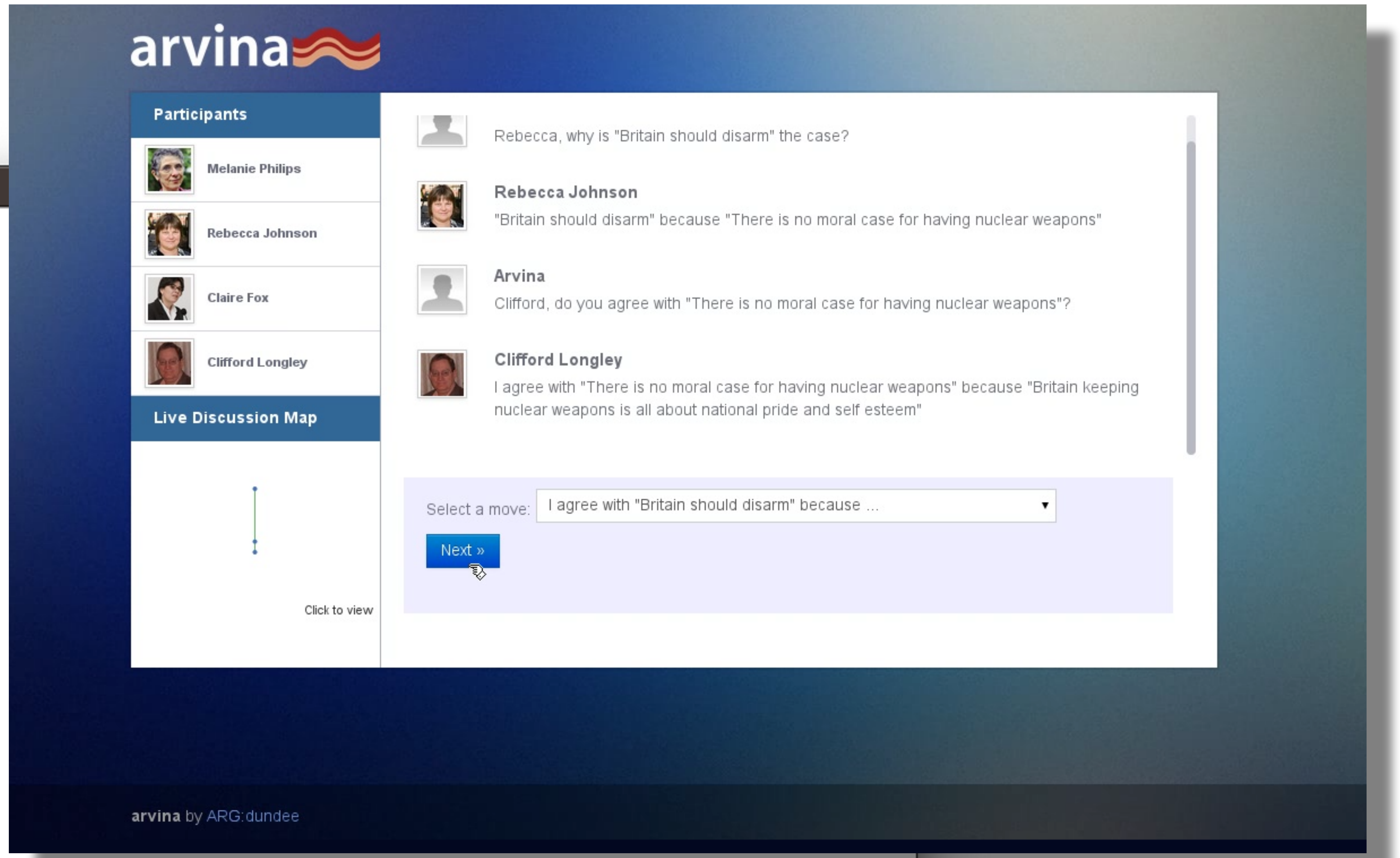
```
grammar dgdl;

game :
    'game' '(' 'id' ':' gameID ')' '{' composition (rules)* (interaction)+ '}' EOF;

gameID : identifier;

composition: roleList? participants player+ store* turntaking? backtrack? extURI*;

roleList :
    'roles' '(' role (',' role)* ')';

role : (LISTENER | SPEAKER | identifier);

participants :
    'participants' '(' 'min' ':' minplayers ',' 'max' ':' maxplayers ')';

player :
    'player' '(' 'id' ':' playerID (',' playerRoleList)? (',' 'min' ':' minplayers)? (',' 'max' ':' maxplayers)? ')';

playerID : identifier;

playerRoleList :
    'roles' ':' '{' role (',' role)* '}';

store :
    'store' '(' 'id' ':' storeID ',' 'owner' ':' storeOwner ',' 'structure' ':' storeStructure ',' 'visibility' ':' storeVisibility (',' storeContent)? ')';

storeID : identifier;

storeOwner :
    (identifier | '{' identifier (',' identifier)+ '}');

storeStructure :
    (SET | QUEUE | STACK);

storeVisibility :
    (PUBLIC | PRIVATE);

storeContent :
    '{' (contentVar | STRINGLITERAL) (',' (contentVar| STRINGLITERAL))* '}';

turntaking :
    'turntaking' '(' turntakingtype ')';

turntakingtype : ('strict' | 'liberal');

backtrack :
    'backtracking' '(' onoff ')';

onoff : ('on' | 'off');

extURI:
    'extURI' '(' 'id' ':' extUriID ',' 'uri' ':' uri ')';

uri : STRINGLITERAL;

extUriID : identifier;

minplayers : number;
maxplayers : (number | 'undefined');

rules :
    'rule' '(' 'id' ':' ruleID ',' 'scope' ':' scopeType ')' ruleBody;

ruleID : identifier;
```




```
raction)+ '}' EOF;

cktrack? extURI*;

ers ')';

n' ':' minplayers)? (',' 'max' ':' maxplayers)? ')';

store :
    'store' '(' 'id' ':' storeID ',' 'owner' ':' storeOwner ',' 'structure' ':' storeStructure ',' 'visibility' ':' storeVisibility (',' storeContent)? ')';

    (identifier | '{' identifier (',' identifier)+ '}');

storeStructure :
    (SET | QUEUE | STACK);

storeVisibility :
    (PUBLIC | PRIVATE);

storeContent :
    '{' (contentVar | STRINGLITERAL) (',' (contentVar| STRINGLITERAL))* '}';

turntaking :
    'turntaking' '(' turntakingtype ')';

turntakingtype : ('strict' | 'liberal');

backtrack :
    'backtracking' '(' onoff ')';

onoff : ('on' | 'off');

extURI:
    'extURI' '(' 'id' ':' extUriID ',' 'uri' ':' uri ')';

uri : STRINGLITERAL;

extUriID : identifier;

minplayers : number;
maxplayers : (number | 'undefined');

rules :
    'rule' '(' 'id' ':' ruleID ',' 'scope' ':' scopeType ')' ruleBody;

ruleID : identifier;
```

Wells & Reed (2008)

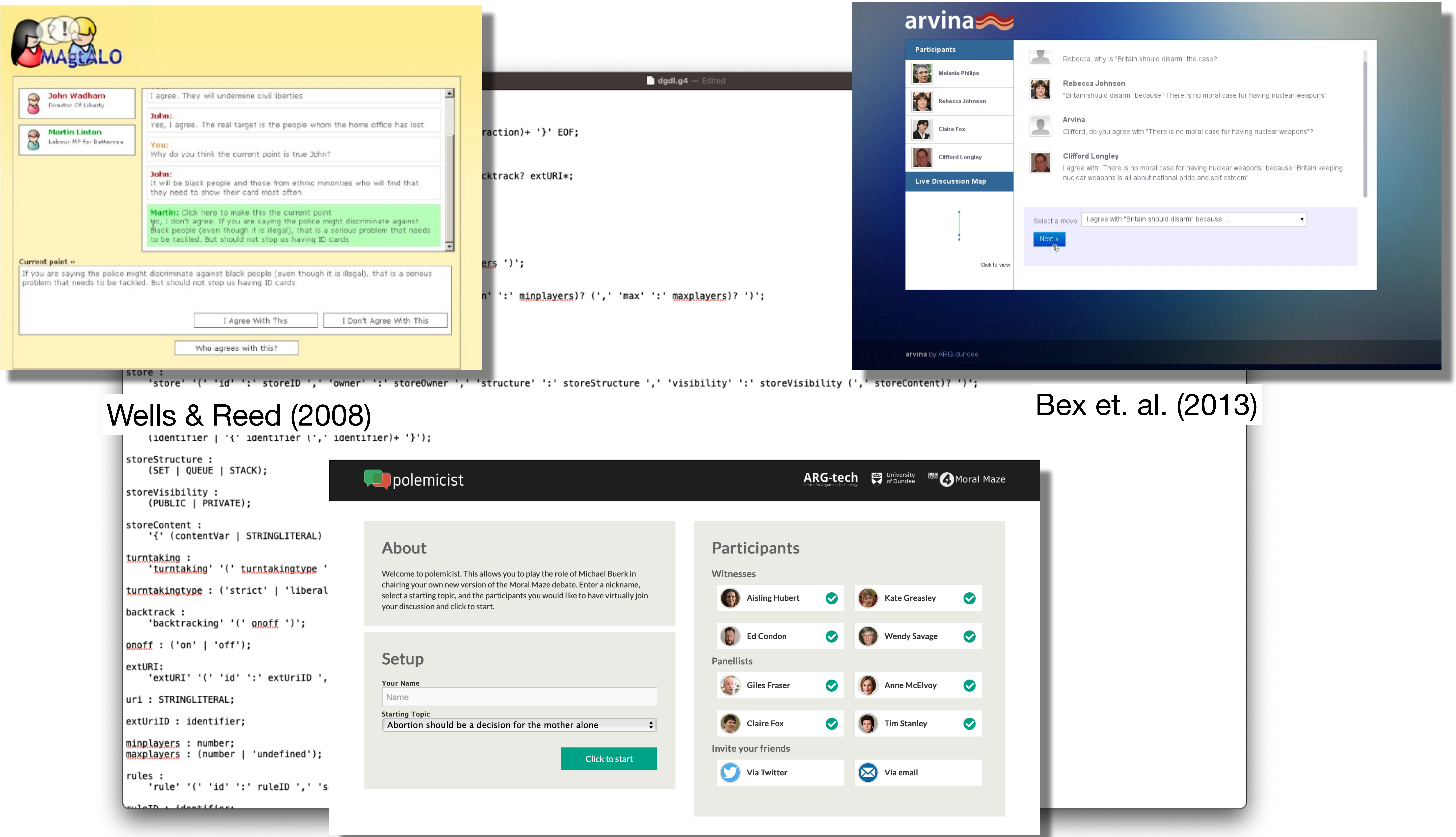
Wells & Reed (2008)



Bex et. al. (2013)

dgdl.g4 — Edited

```
raction)+ '}' EOF;

cktrack? extURI*;

ers ')';

n' ':' minplayers)? (',' 'max' ':' maxplayers)? ')';

store :
    'store' '(' 'id' ':' storeID ',' 'owner' ':' storeOwner ',' 'structure' ':' storeStructure ',' 'visibility' ':' storeVisibility (',' storeContent)? ')';

    (identifier | '{' identifier (',' identifier)+ '}');

storeStructure :
    (SET | QUEUE | STACK);

storeVisibility :
    (PUBLIC | PRIVATE);

storeContent :
    '{' (contentVar | STRINGLITERAL) (',' (contentVar| STRINGLITERAL))* '}';

turntaking :
    'turntaking' '(' turntakingtype ')';

turntakingtype : ('strict' | 'liberal');

backtrack :
    'backtracking' '(' onoff ')';

onoff : ('on' | 'off');

extURI:
    'extURI' '(' 'id' ':' extUriID ',' 'uri' ':' uri ')';

uri : STRINGLITERAL;

extUriID : identifier;

minplayers : number;
maxplayers : (number | 'undefined');

rules :
    'rule' '(' 'id' ':' ruleID ',' 'scope' ':' scopeType ')' ruleBody;
```



Wells & Reed (2008)

Bex et. al. (2013)

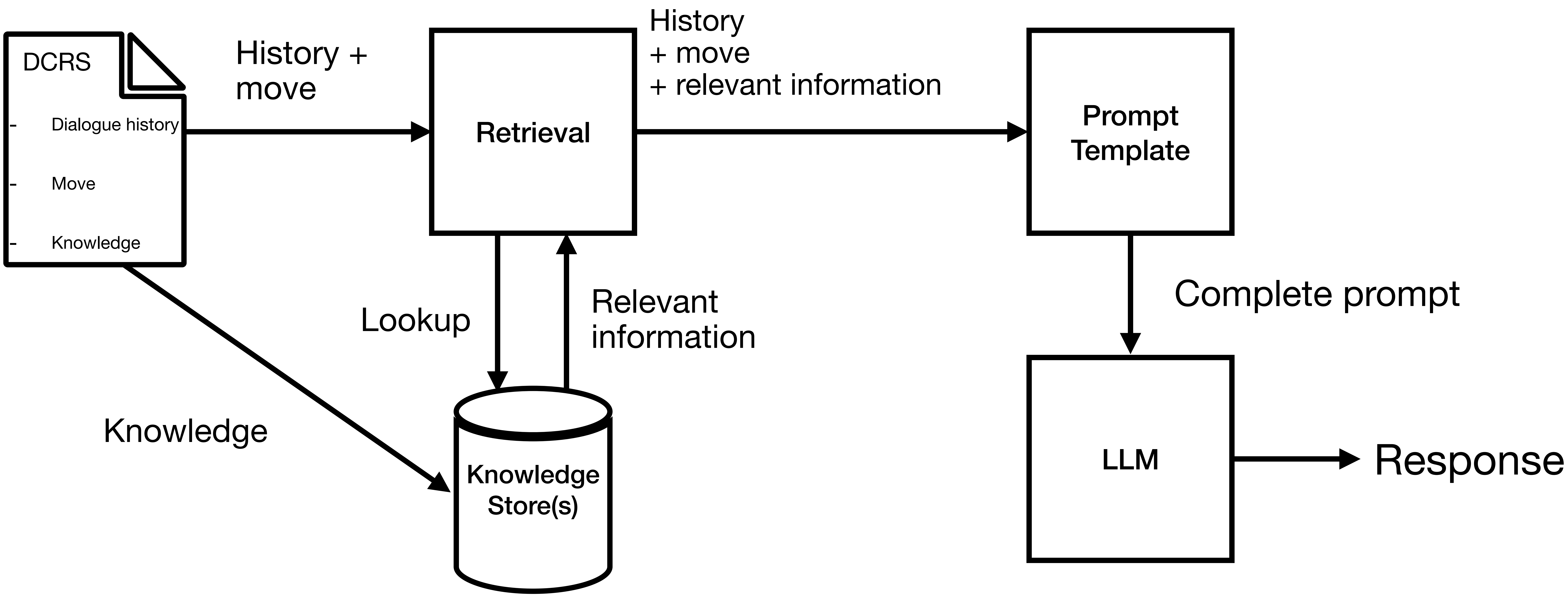
Lawrence et. al. (2022)

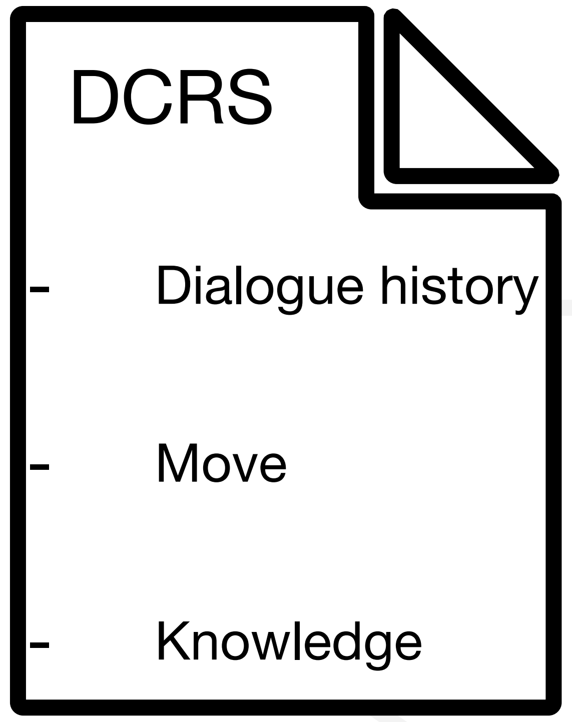
Formal dialogue vs LLMs

Aspect	Formal Dialogue	LLMs
Structure	Rigid, protocol-driven turns	Free-flowing, unconstrained text
Goals	Clear purposes (e.g. persuasion, inquiry, negotiation)	Flexible, task-agnostic
Transparency	High: reasoning explicit through moves	Low: opaque “black box”
Naturalness	Often stilted and hard for users to engage with	Human-like fluency
Grounding	Rules and knowledge bases explicitly defined	Implicit in training data

Formal dialogue vs LLMs

Aspect	Formal Dialogue	LLMs	Formal Dialogue + LLM + RAG
Structure	Rigid, protocol-driven turns	Free-flowing, unconstrained text	**Natural expression mapped to structured moves**
Goals	Clear purposes (e.g. persuasion, inquiry, negotiation)	Flexible, task-agnostic	**Preserve dialogue goals while adapting language**
Transparency	High: reasoning explicit through moves	Low: opaque "black box"	**Retrieved evidence grounds outputs, explanations possible**
Naturalness	Often stilted and hard for users to engage with	Human-like fluency	**Formal logic preserved, but phrased naturally**
Grounding	Rules and knowledge bases explicitly defined	Implicit in training data	**Retrieval ensures domain-specific, up-to-date knowledge from multiple sources**

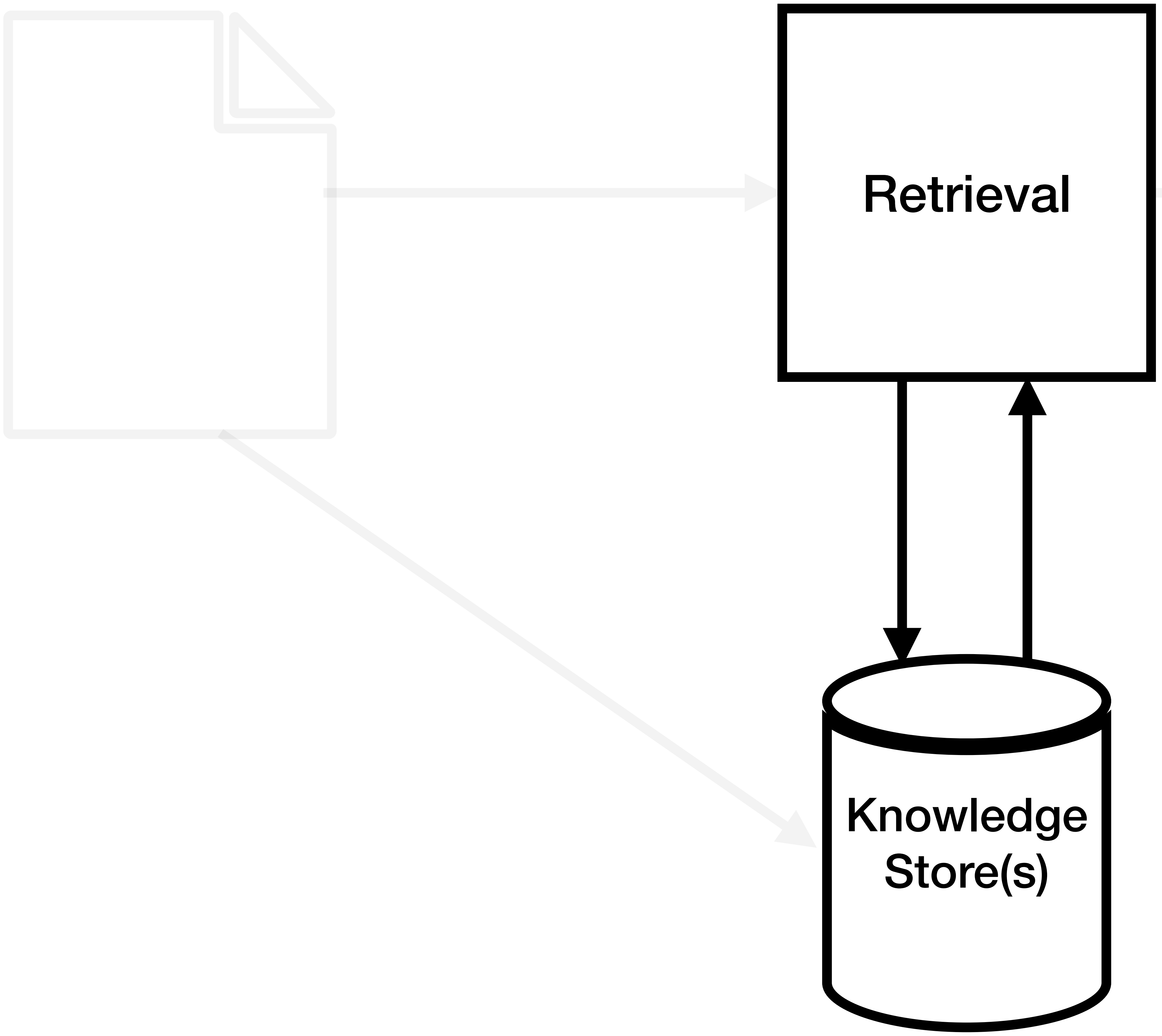

DCRS
- Dialogue history
- Move
- Knowledge
History + move
Retrieval
History
+ move
+ relevant information
Prompt Template
Lookup
Relevant information
Knowledge
Knowledge Store(s)
Complete prompt
LLM
Response



Dialogue Context Request Structure

Contains:

- Knowledge (from the agent)
- Dialogue history (for context)
- Move scaffold (to help the LLM structure its response)



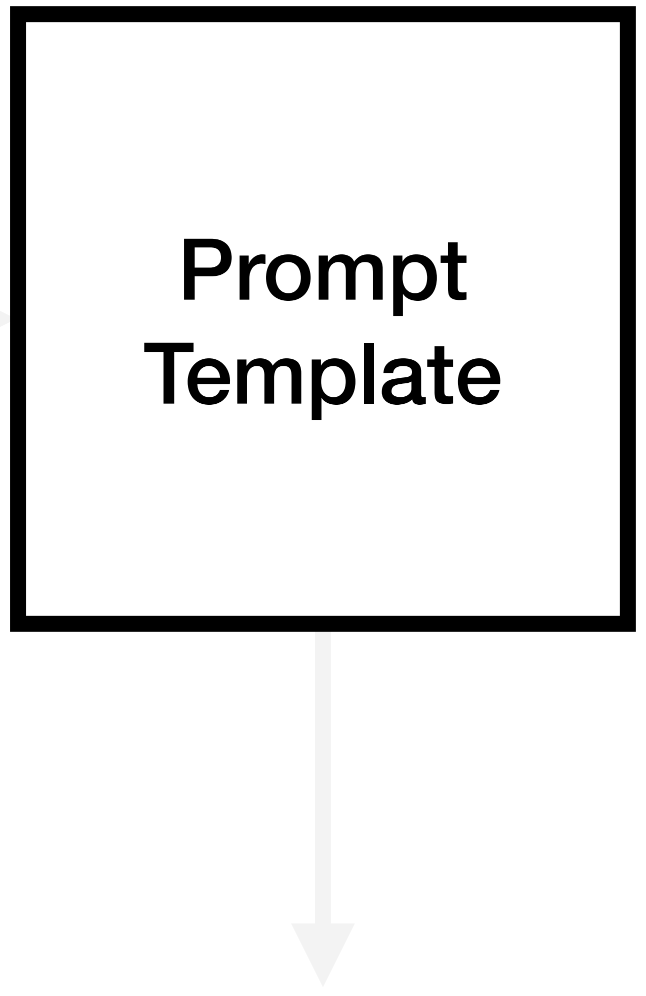
Vector Database(s) - semantic search

Returns *relevant* knowledge based on a query assembled from:

- Dialogue history
- Move scaffold

Four key elements:

- Task description (instructions)
- Dialogue history
- Move scaffold
- Retrieved knowledge
    - And how it should be interpreted



Prompt Template

You are assisting in a dialogue by generating the next utterance. Use the provided scaffold, dialogue history, and structured knowledge sources below. *Prioritise the provided knowledge sources in generating the next utterance, but if a suitable utterance can't be generated you may use your own knowledge, in which case identify the source as "generated".*

Dialogue History
{{history}}

Utterance Scaffold
{{Utterance}}

You may adapt the scaffold to better match the tone, style, or conversational flow of the dialogue history. The goal is to maintain coherence and sound natural within the ongoing dialogue.

Knowledge Sources

The knowledge provided below includes different structures. Each structure has a name, a formal representation, and a description of how to interpret it.

{{knowledge}}

Prompt Template

You are assisting in a dialogue by generating the next utterance. Use the provided scaffold, dialogue history, and structured knowledge sources below. ***Prioritise the provided knowledge sources in generating the next utterance, but if a suitable utterance can't be generated you may use your own knowledge, in which case identify the source as "generated".***

Dialogue History
{{history}}

Utterance Scaffold
{{Utterance}}

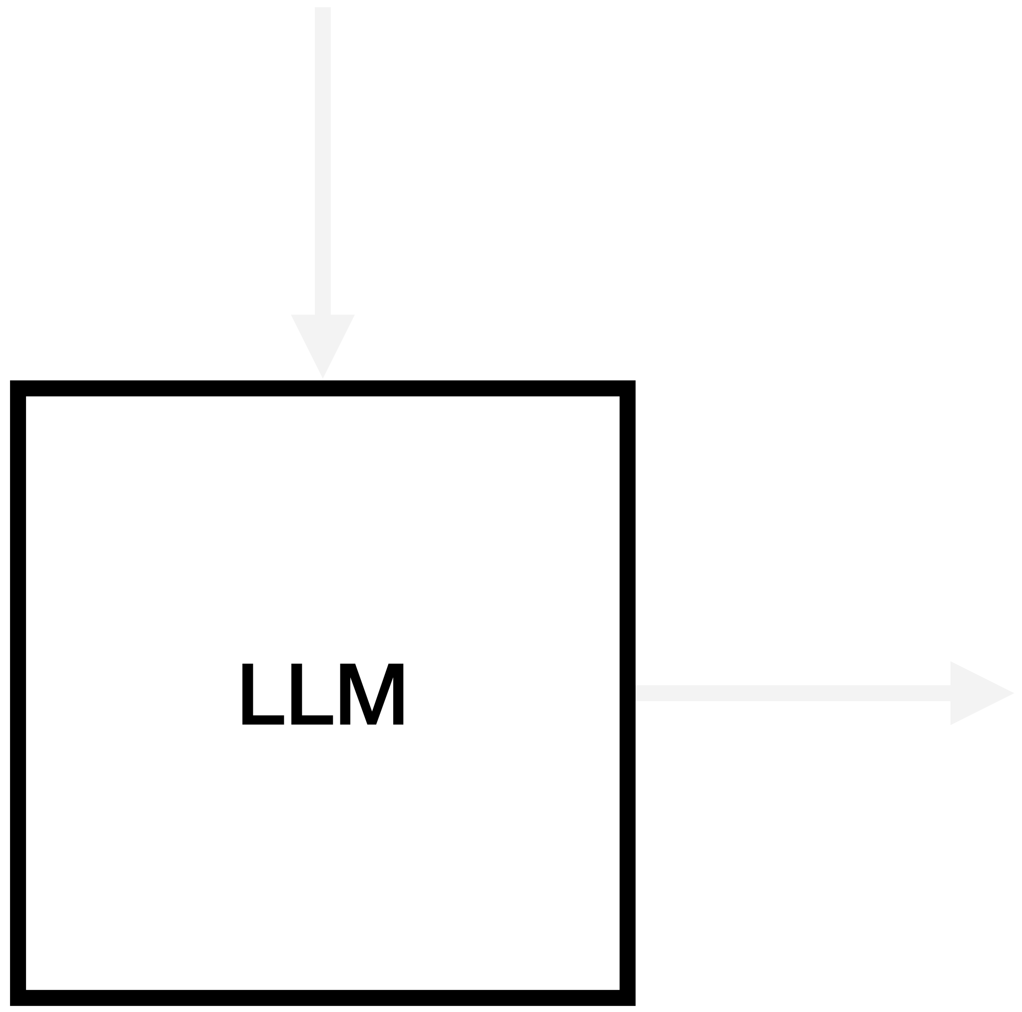
You may adapt the scaffold to better match the tone, style, or conversational flow of the dialogue history. The goal is to maintain coherence and sound natural within the ongoing dialogue.

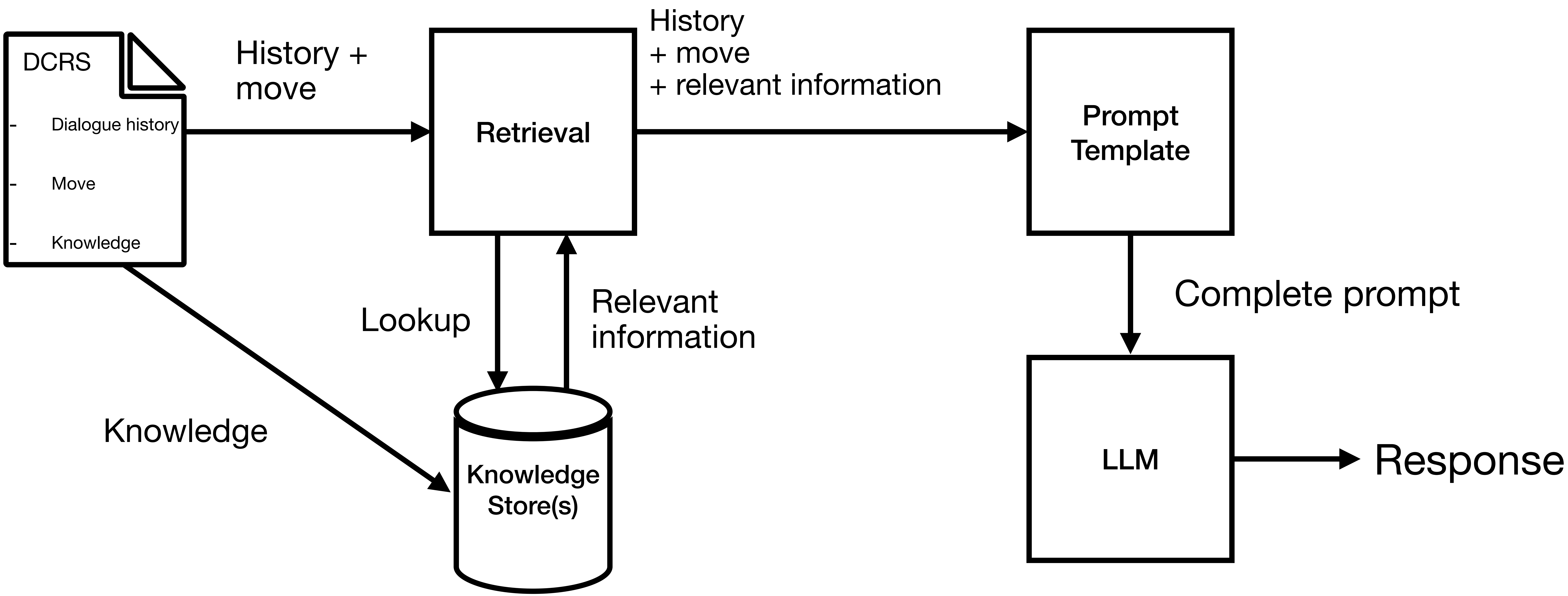
Knowledge Sources

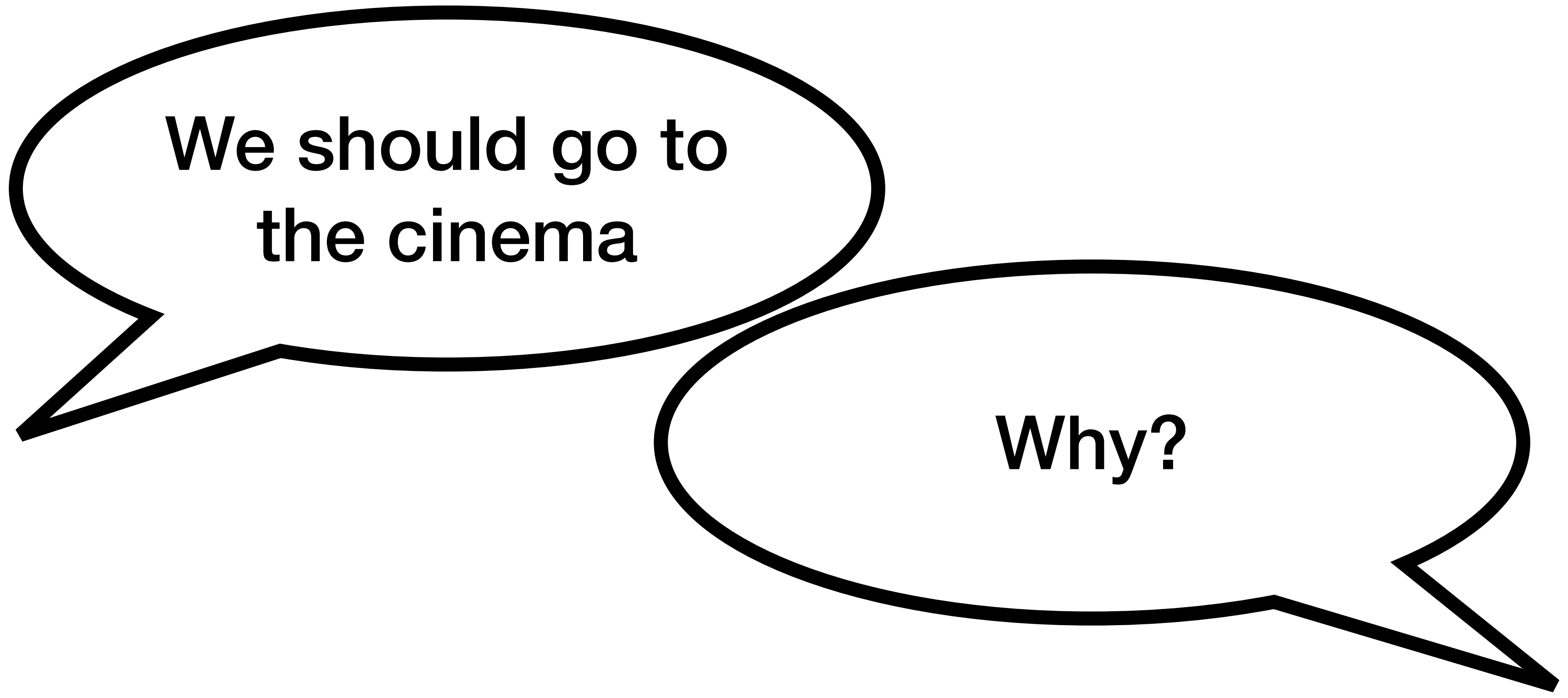
The knowledge provided below includes different structures. Each structure has a name, a formal representation, and a description of how to interpret it.

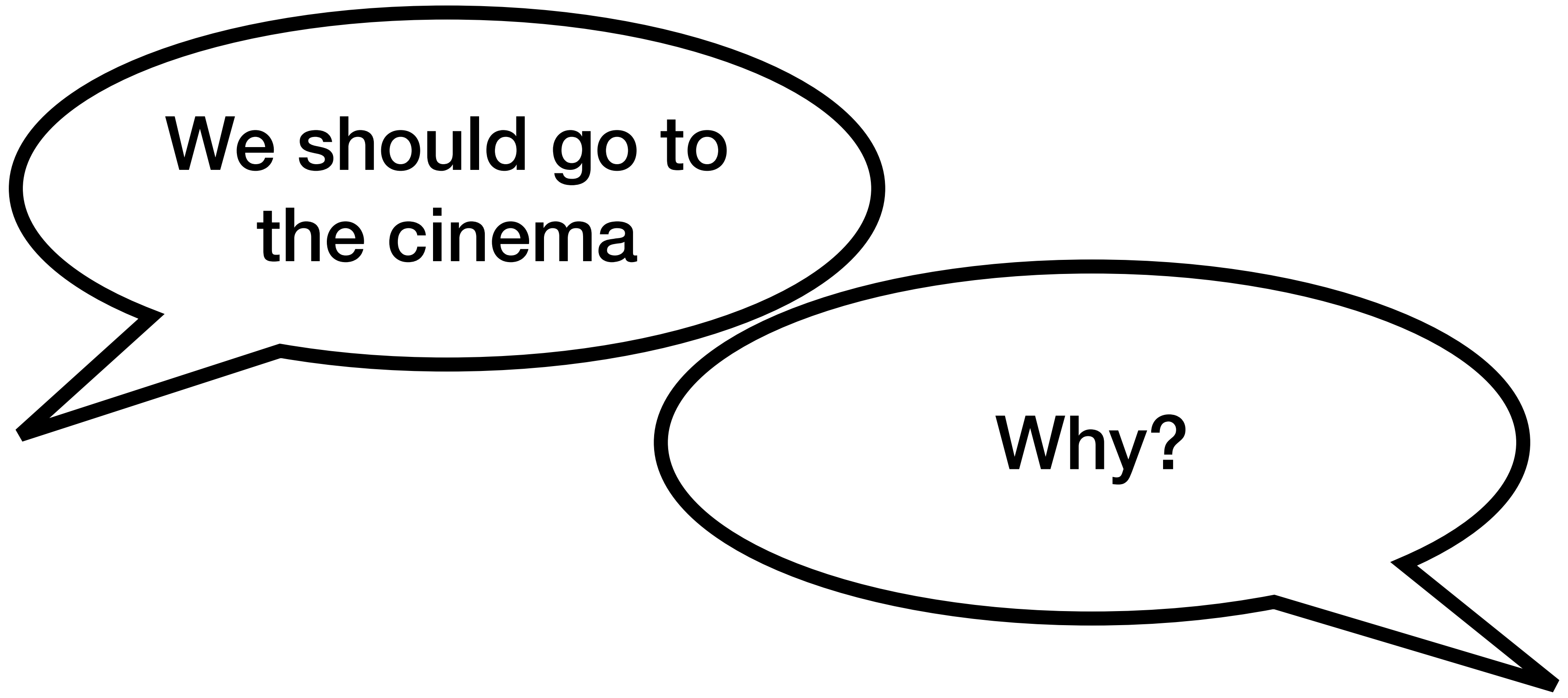
{{knowledge}}

- Swappable LLMs - cloud or local
- Response: structured JSON object
- Returns multiple utterances, tagged by source

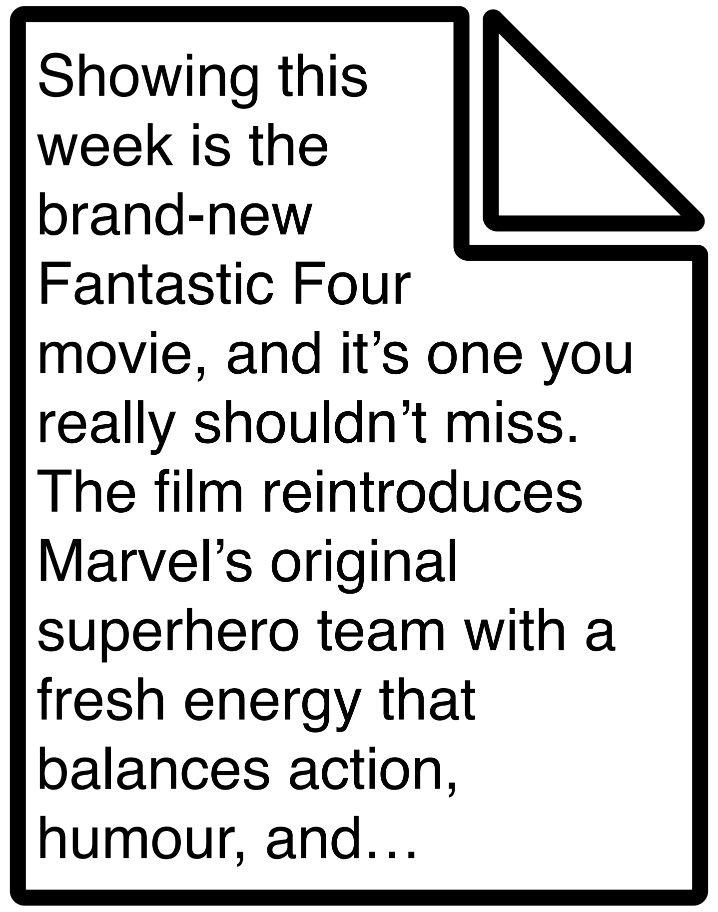


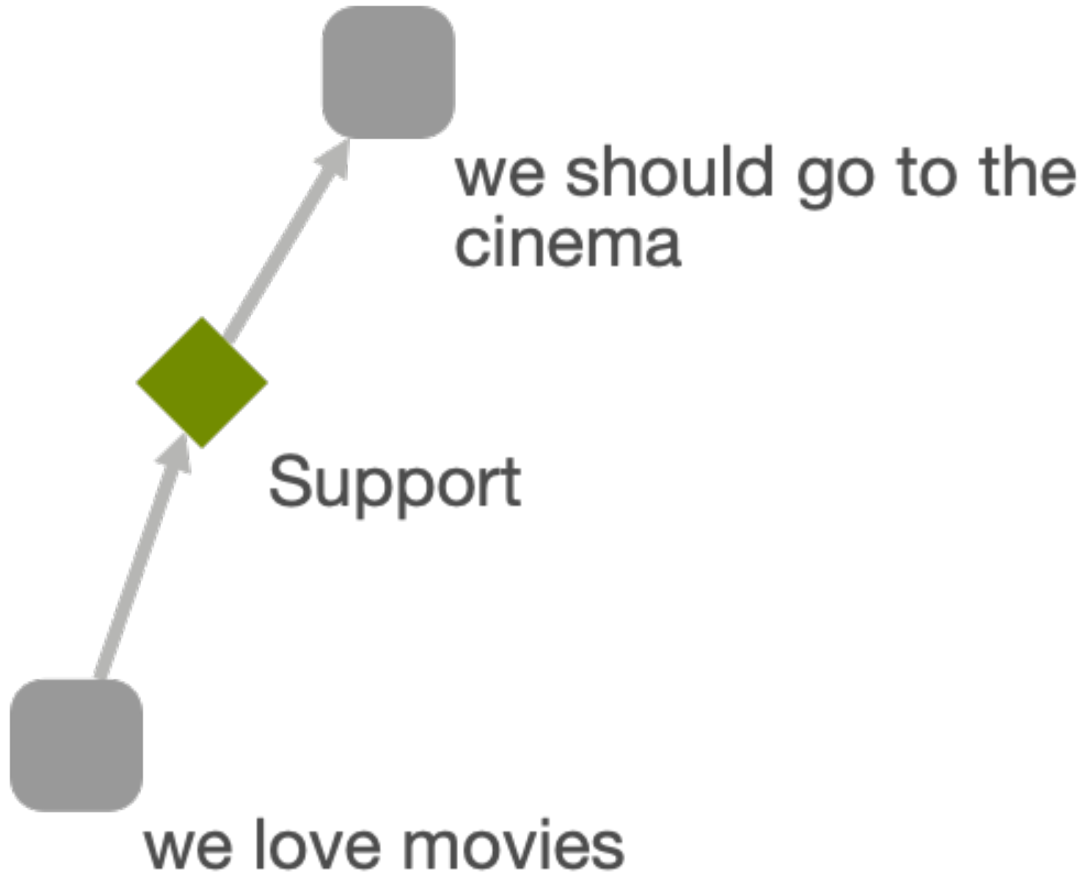

DCRS
- Dialogue history
- Move
- Knowledge
History + move
Retrieval
History
+ move
+ relevant information
Prompt Template
Lookup
Relevant information
Knowledge
Knowledge Store(s)
Complete prompt
LLM
Response

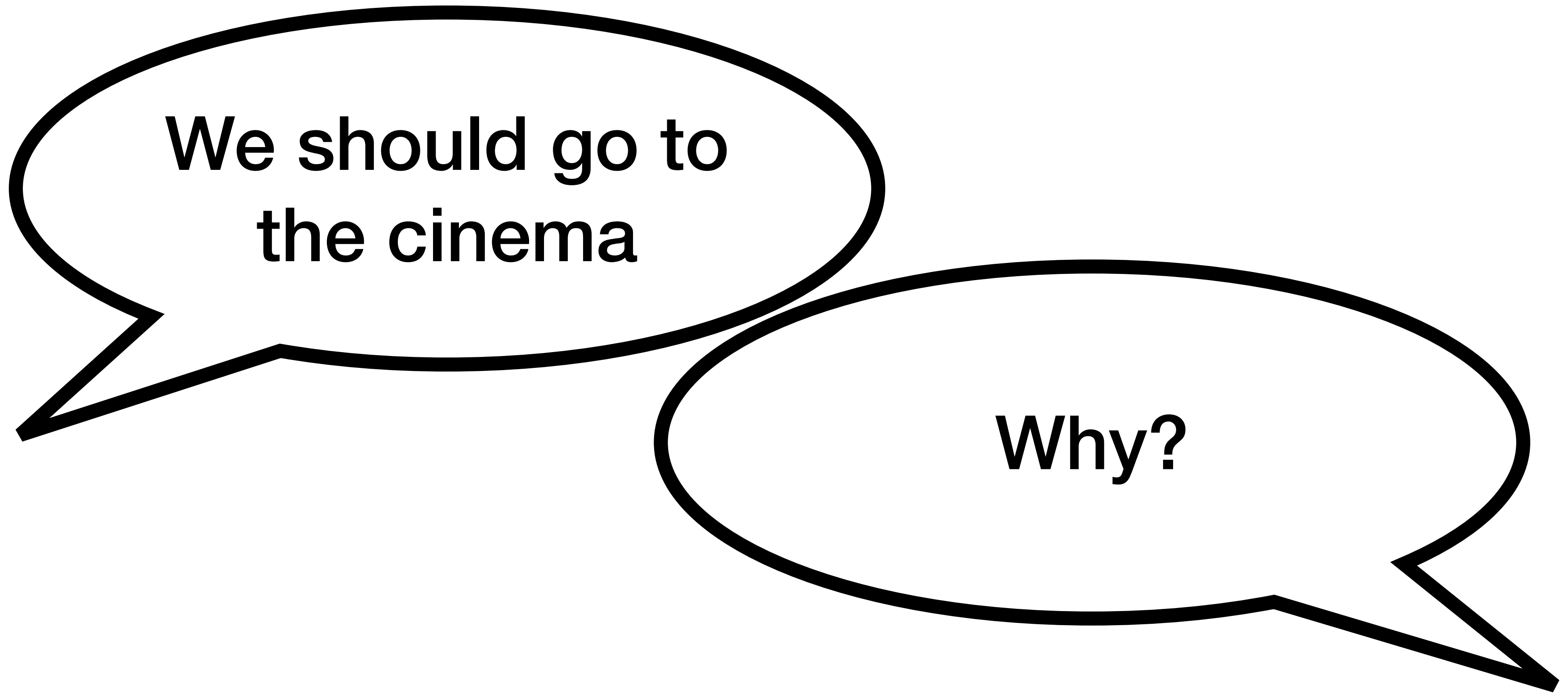

We should go to the cinema
Why?


We should go to the cinema
Why?

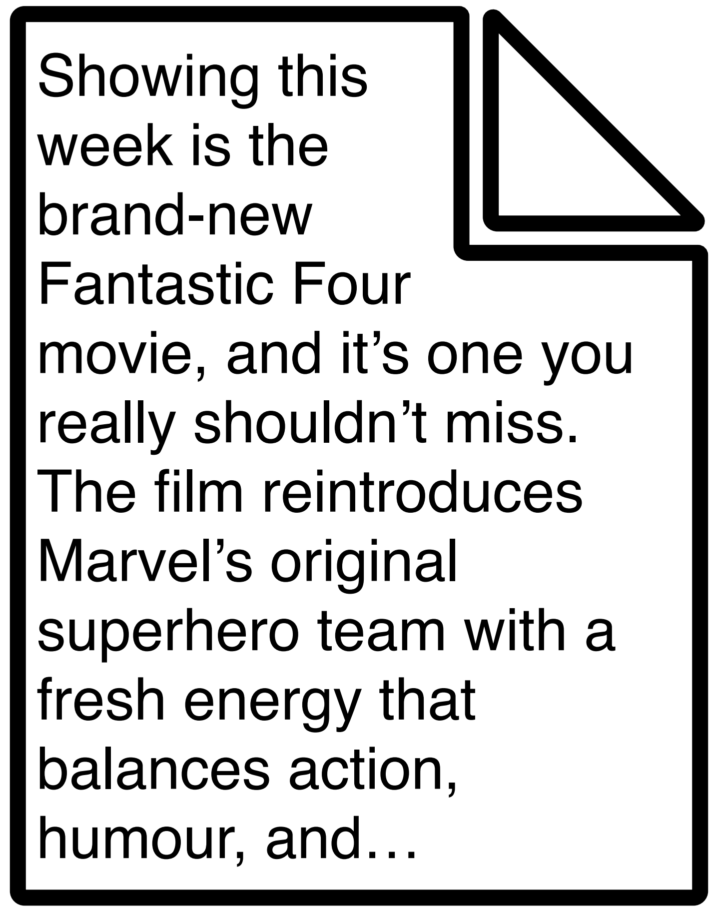
Knowledge

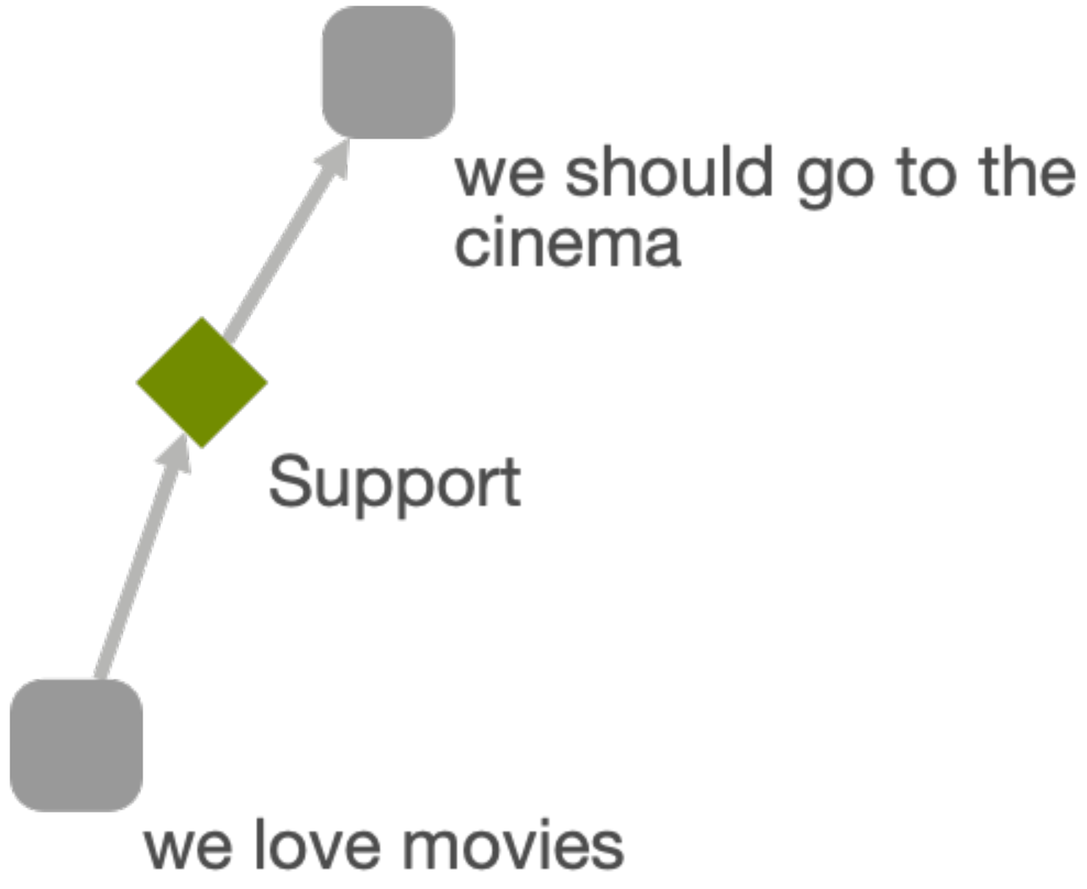

Showing this week is the brand-new Fantastic Four movie, and it’s one you really shouldn’t miss. The film reintroduces Marvel’s original superhero team with a fresh energy that balances action, humour, and…


we should go to the cinema
Support
we love movies



Knowledge

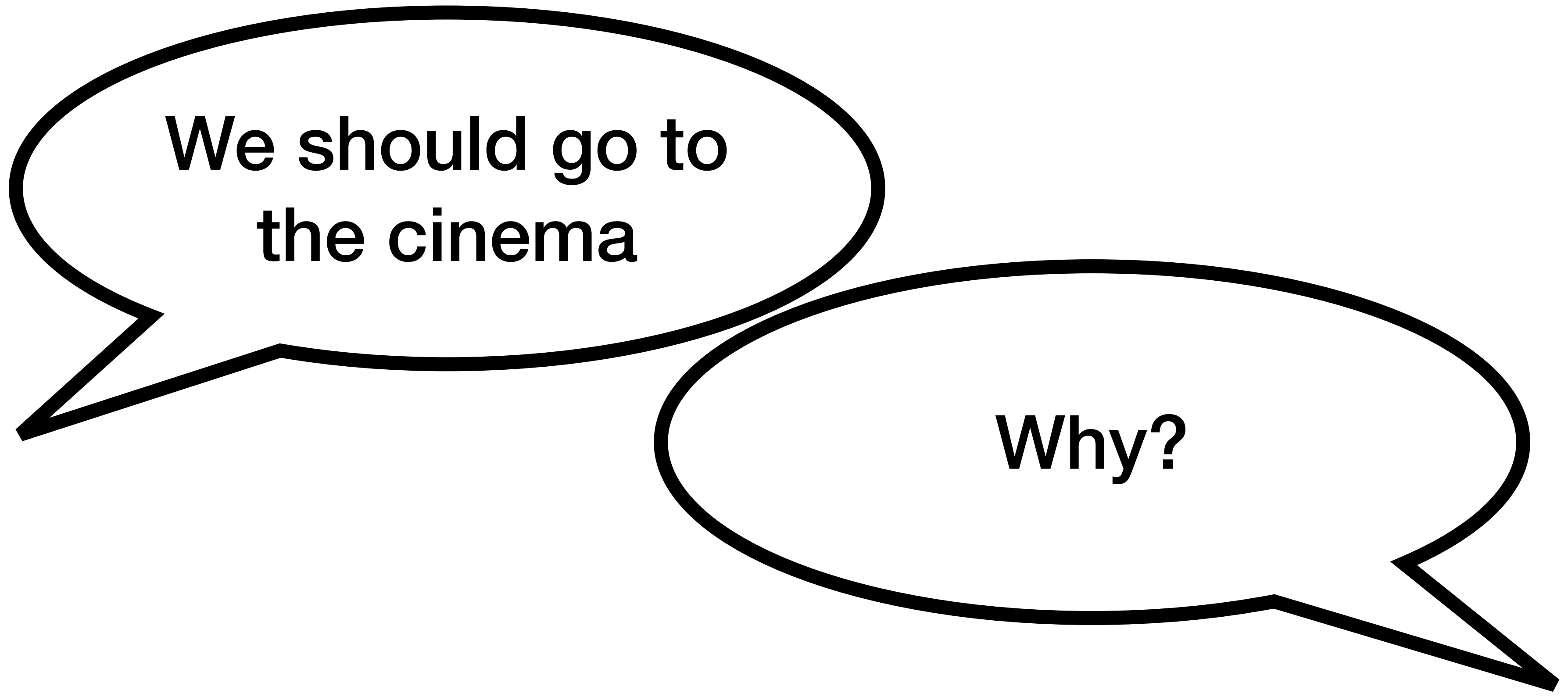




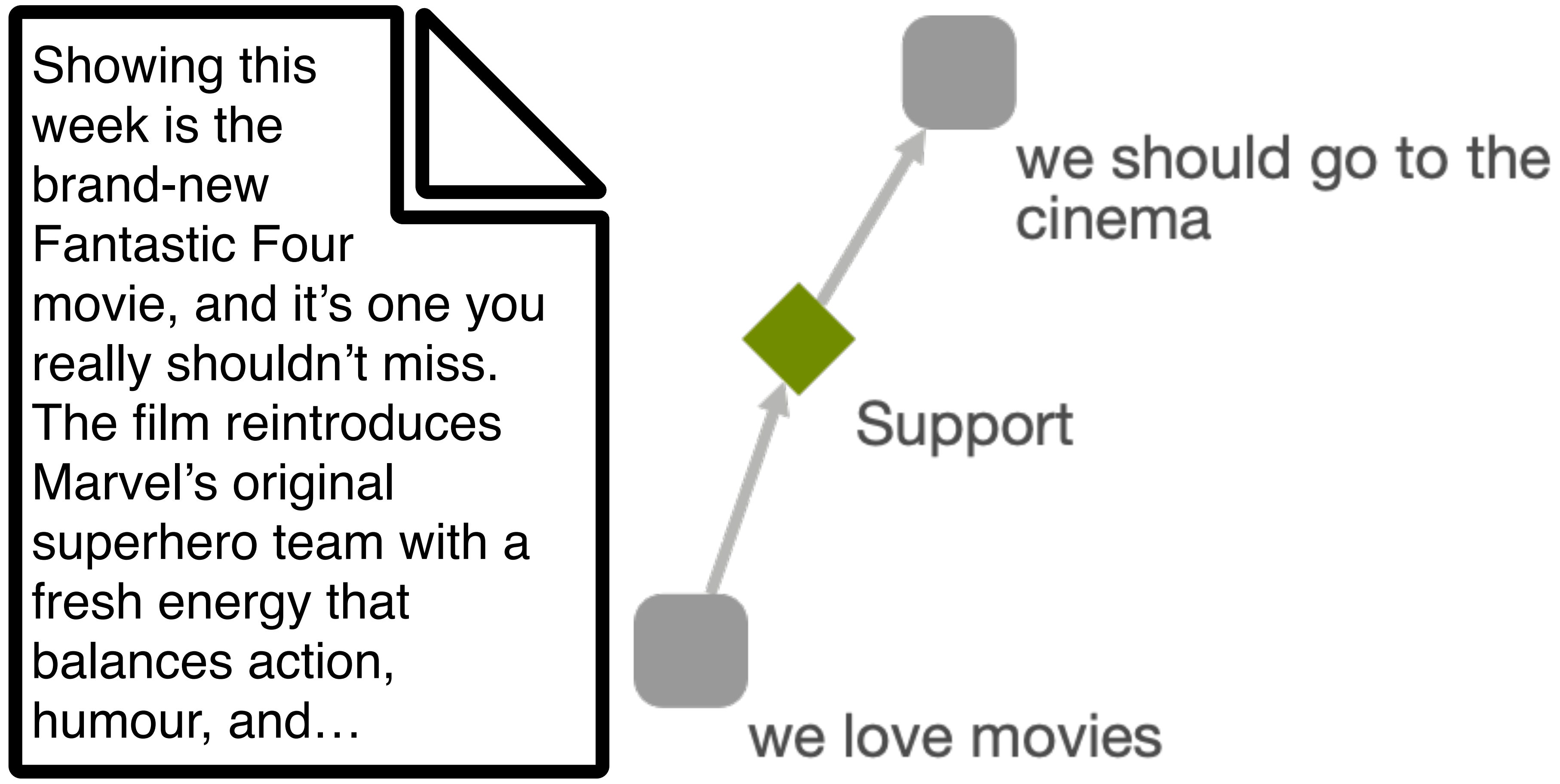
Dialogue history

Alice: We should go to the cinema [assert]

Bob: Why? [challenge]



Knowledge



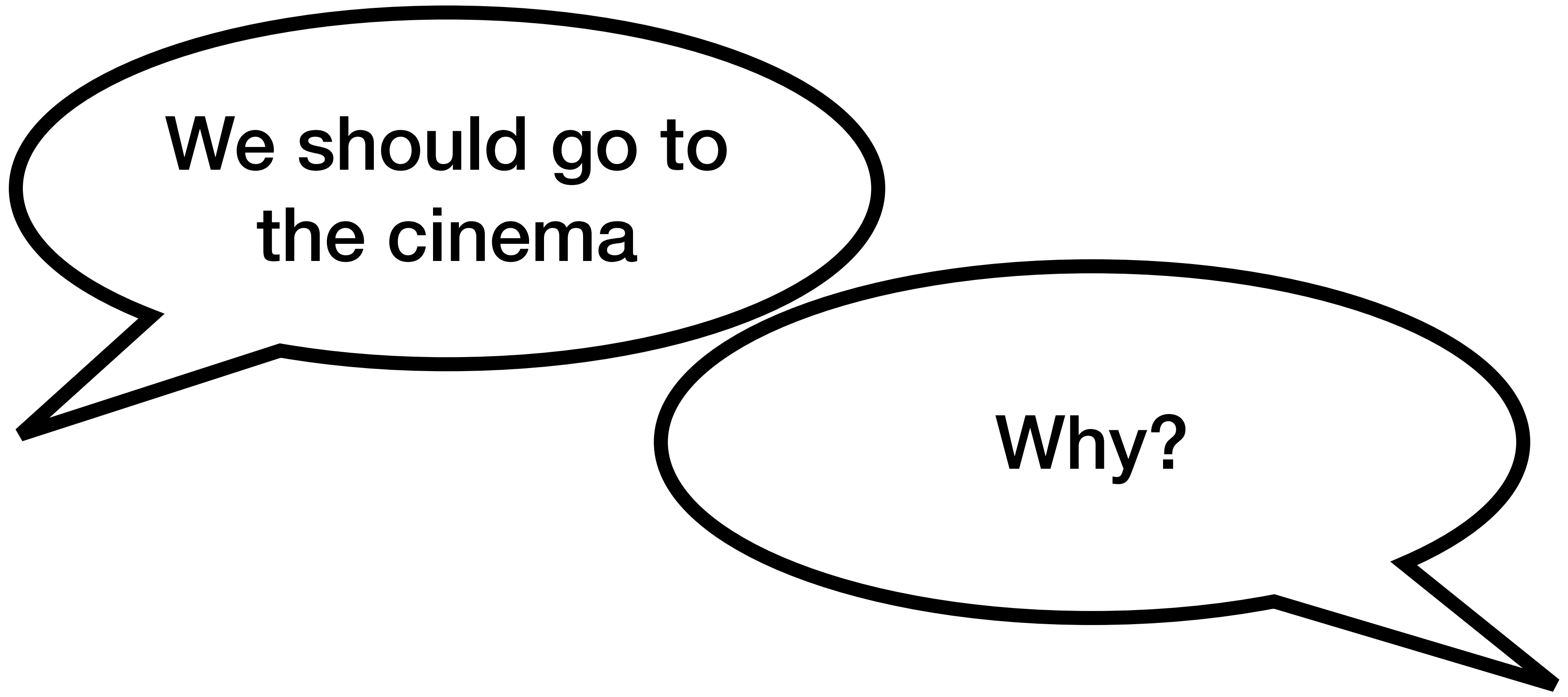
Dialogue history

Alice: We should go to the cinema [assert]

Bob: Why? [challenge]

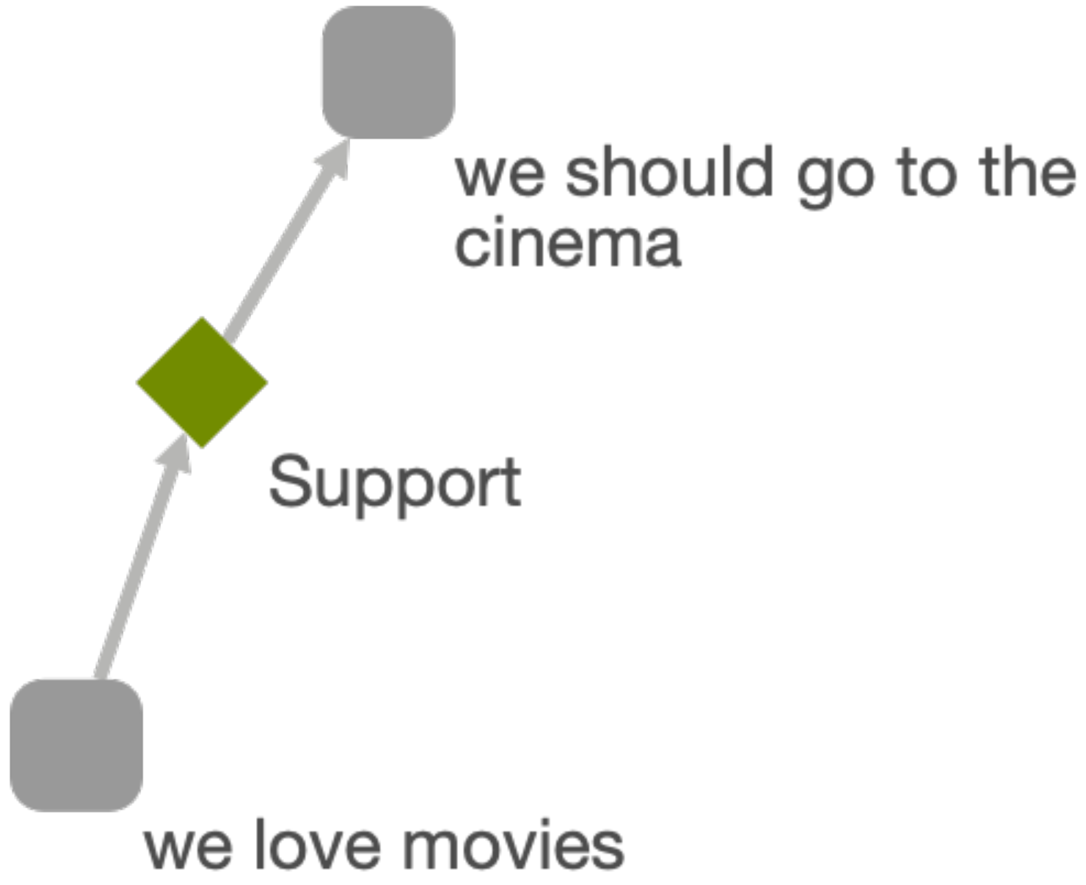
Scaffold

"We should go to the cinema because $p"



Knowledge

Showing this week is the brand-new Fantastic Four movie, and it's one you really shouldn't miss. The film reintroduces Marvel's original superhero team with a fresh energy that balances action, humour, and…



Natural language text to interpret. Rephrase if needs be to make short responses using this source.

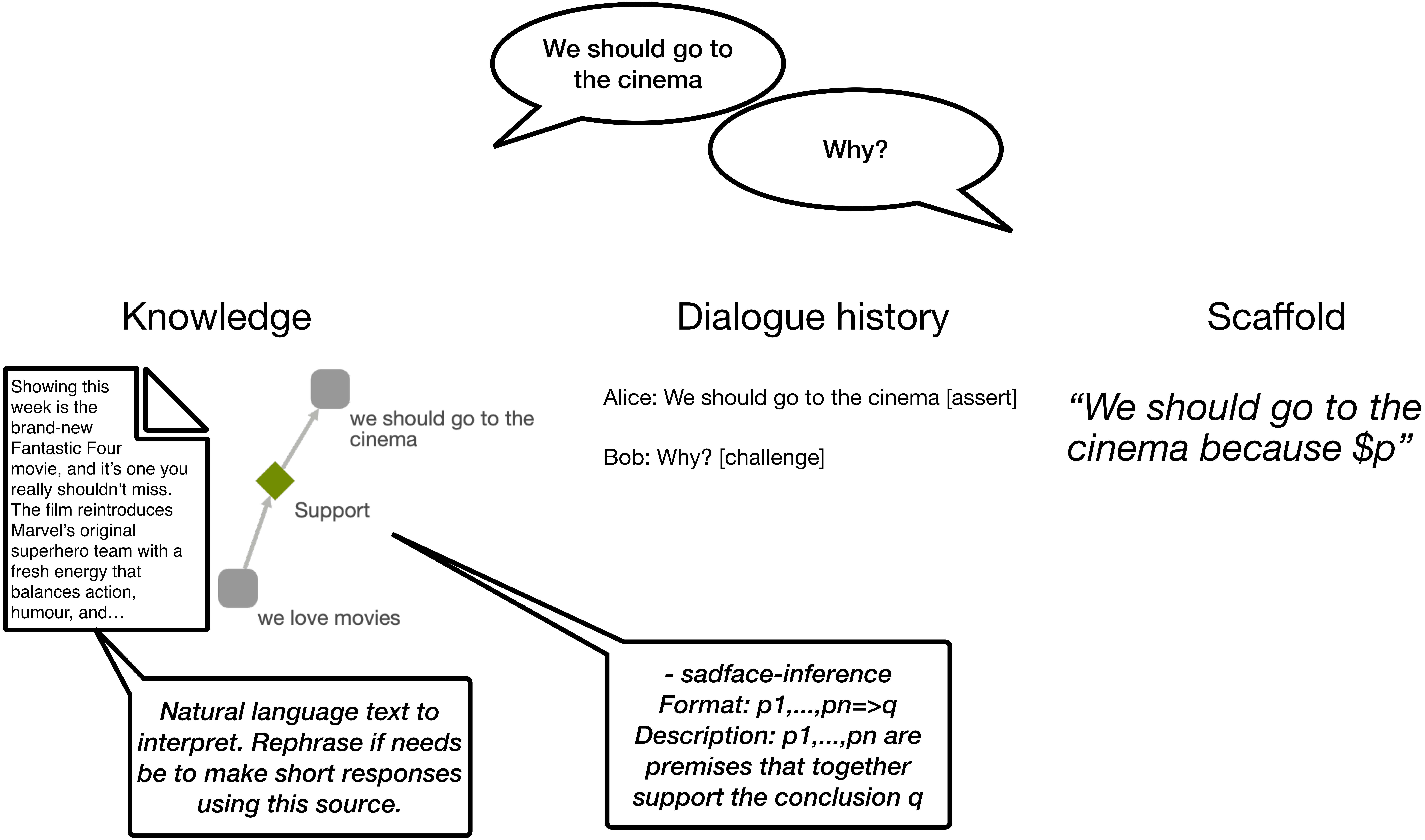
Dialogue history

Alice: We should go to the cinema [assert]

Bob: Why? [challenge]

Scaffold

"We should go to the cinema because $p"



We should go to the cinema
Why?
Knowledge
Dialogue history
Scaffold
Showing this week is the brand-new Fantastic Four movie, and it's one you really shouldn't miss. The film reintroduces Marvel's original superhero team with a fresh energy that balances action, humour, and…
we should go to the cinema
Support
we love movies
Alice: We should go to the cinema [assert]
Bob: Why? [challenge]
"We should go to the cinema because $p"
Natural language text to interpret. Rephrase if needs be to make short responses using this source.
- sadface-inference
Format: p1,...,pn=>q
Description: p1,...,pn are premises that together support the conclusion q

Know fold

Showing this week is the brand-new Fantastic Four movie, and it’s one you really shouldn’t miss. The film reintroduces Marvel’s original superhero team with a fresh energy that balances action, humour, and…

o to the
use $p”


```
{
  "responses": [
    {
      "response": "We should go to the cinema because the brand-new Fantastic Four movie is showing this week, with great effects and a fun team dynamic that’s best enjoyed on the big screen.",
      "source": "text"
    },
    {
      "response": "We should go to the cinema because we love movies, and it’s a perfect chance to enjoy one together.",
      "source": "sadface"
    }
  ]
}
```

Ongoing and future work

- Evaluation - Two types:
  - Is the LLM giving “good” responses?
  - How “good” are the responses from different LLMs?
- Combining knowledge sources in a single response
  - Currently limited to one source per response
- Strategic support - LLM judgement of the “best” response to put forward

DiSCoAI - Dialogue-based Structured Conversational AI

2025-2028

- Automated implementation of computational dialogue games from real (human-human) examples
- Underpinning conversational systems with formal dialogue without the design and development overhead
- When combined with RAG pipeline = rapid development of rich and immersive systems across a wide variety of domains

UKRI Engineering and Physical Sciences Research Council

RGU ABERDEEN SCHOOL OF COMPUTING, ENGINEERING & TECHNOLOGY

https://gtr.ukri.org/projects?ref=220

Summary

- LLM and RAG-based pipeline for generated responses in formal dialogue
- Knowledge, history and move scaffold used to ground the response
- Bridges the vocabulary gap to make formal dialogue responses more natural

Thank you

ROBERT GORDON UNIVERSITY ABERDEEN

Edinburgh Napier UNIVERSITY