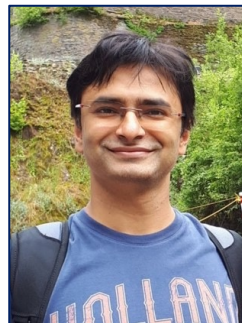
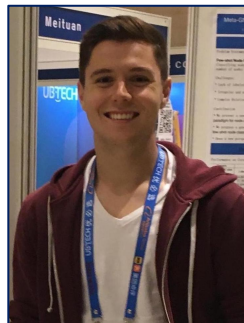


# Conversational Question Answering on Heterogeneous Sources

---

*Philipp Christmann, Rishiraj Saha Roy, Gerhard Weikum*



# Conversational Question Answering

Who played Jaime Lannister in Game of Thrones?

Nikolaj Coster-Waldau

What about the dwarf?

Peter Dinklage

When was he born?

11 June 1969

Release date of first season?

17 April 2012

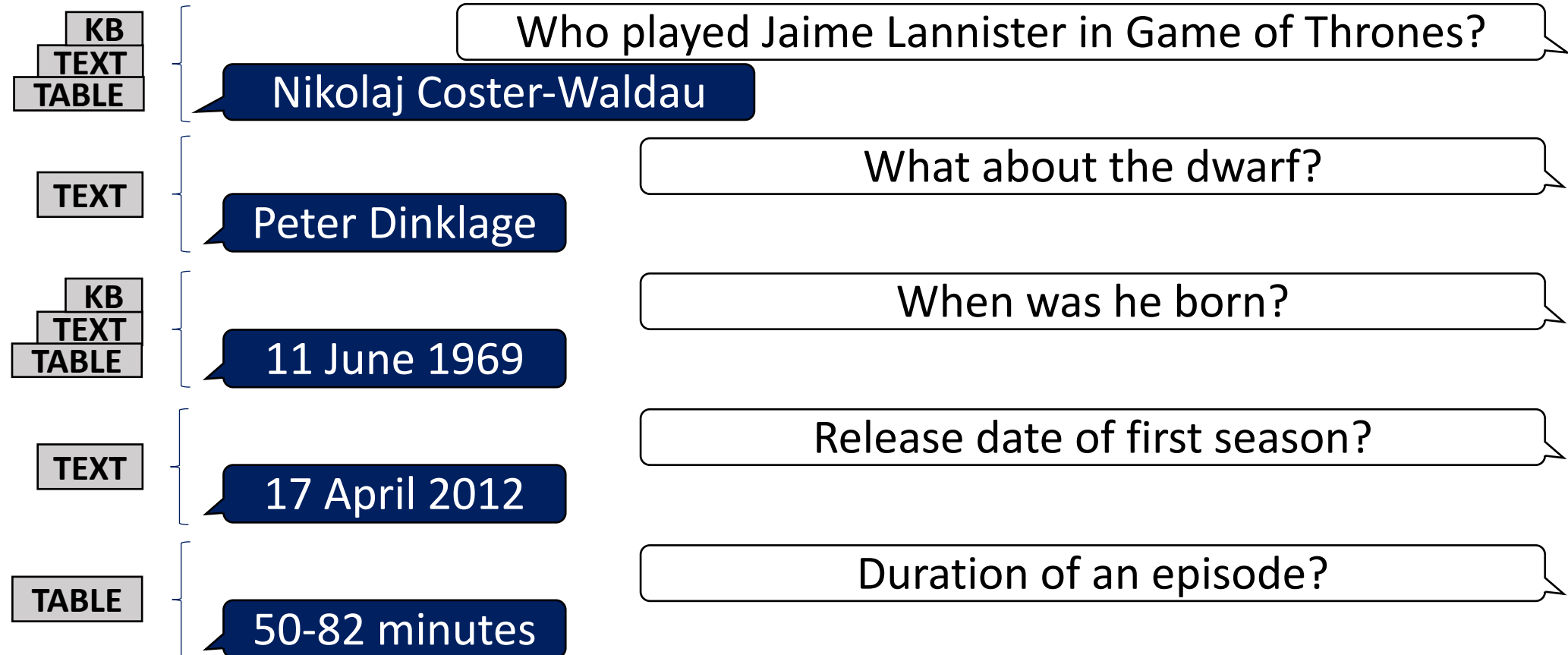
Duration of an episode?

50-82 minutes

⇒ Information left **implicit**

⇒ Must be derived from **conversational context**

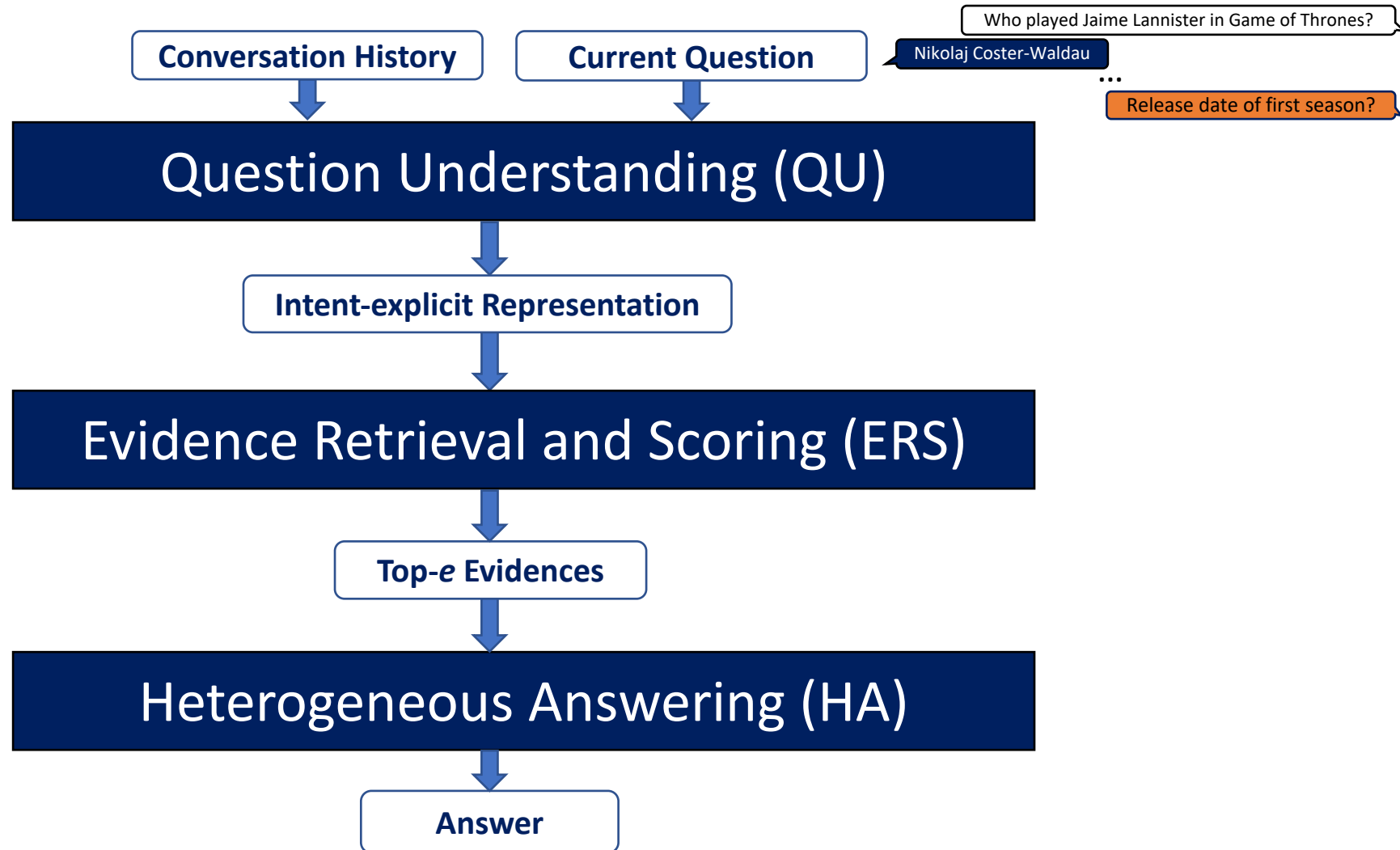
# Conversational Question Answering



⇒ Heterogeneous information sources **for broader coverage**

⇒ **Leverage redundancy** of information sources

# Pipeline Overview



# Desiderata and Contributions

---

Leverage **heterogeneous** information sources for conversational QA

★ General 3-staged pipeline for this task

No restriction to specific information source

★ **CONVINSE**: Intent-explicit **structured representation** of questions

Make process **explainable**

★ **Intermediate** representations in **symbolic** space

Avoid expensive annotations

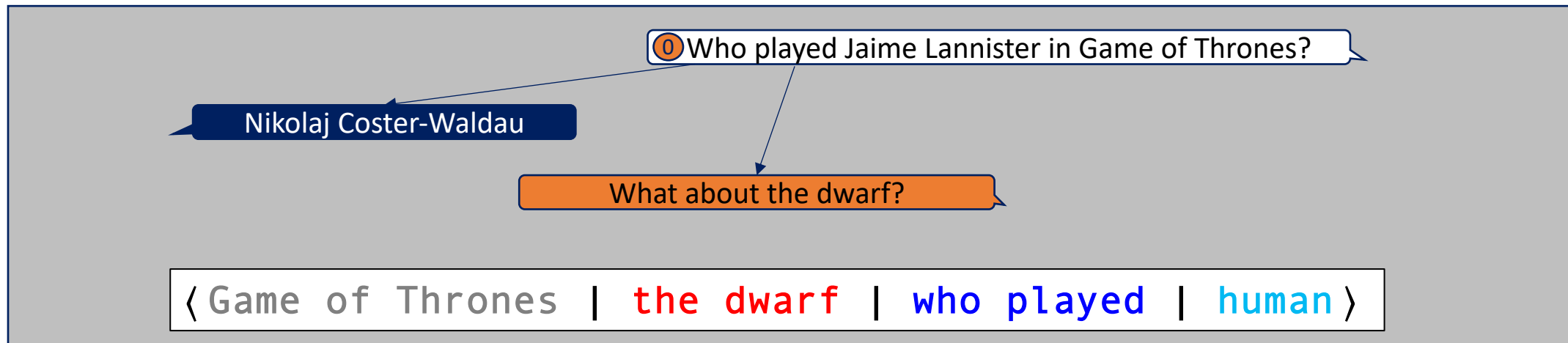
★ **Distant supervision** from **raw** sequences of **QA-pairs**

# Structured Representation (SR)

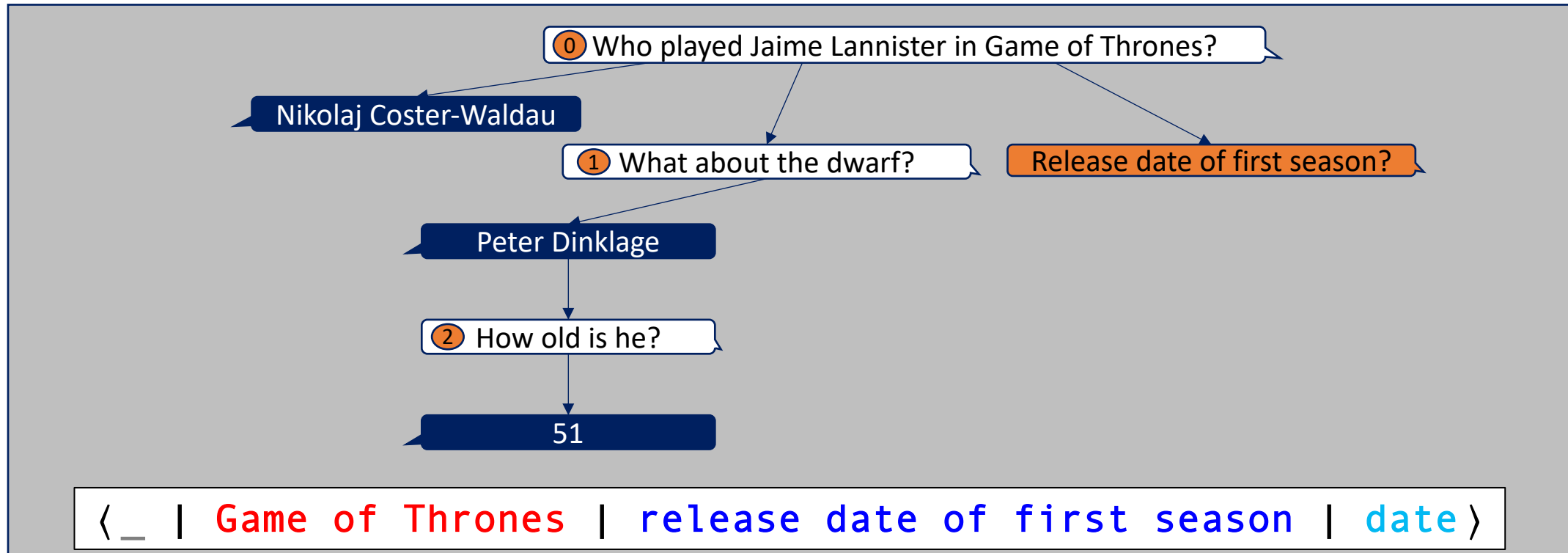
- ⇒ Capture information need in a **structured** way, based solely on the conversation
- ⇒ Relaxed **categorization** into

`< context | entity | relation | answer type >`

- ⇒ **Context** makes the entity more **explicit**



# Structured Representation (SR)

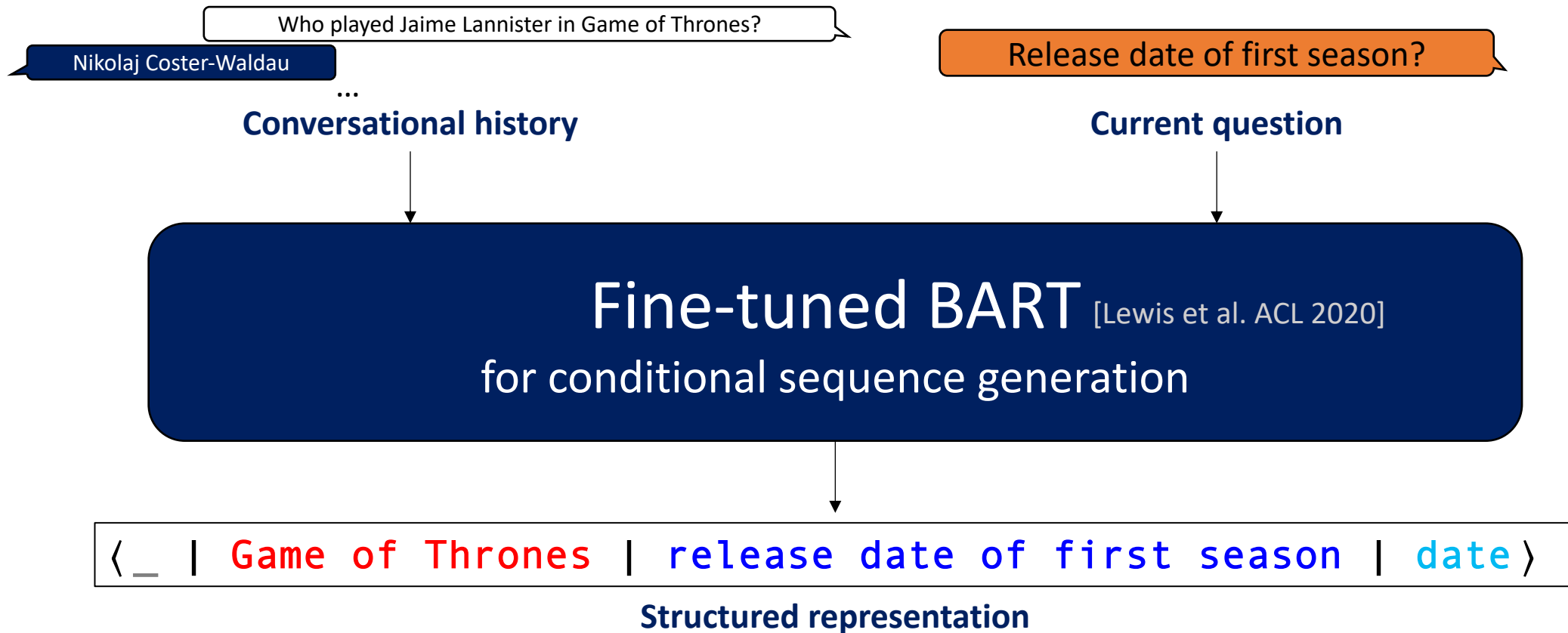


⇒ **Context** not always required

⇒ **Relaxed** notion of a relation

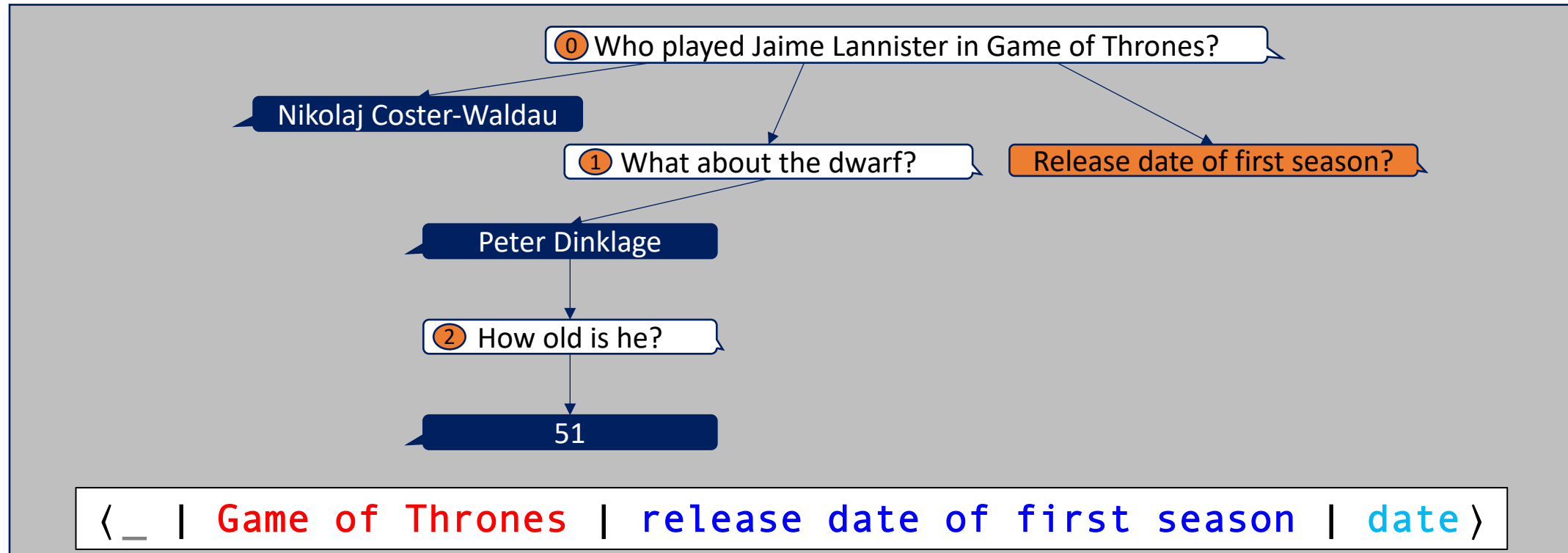
# Structured Representation Module

⇒ Sequence generation task



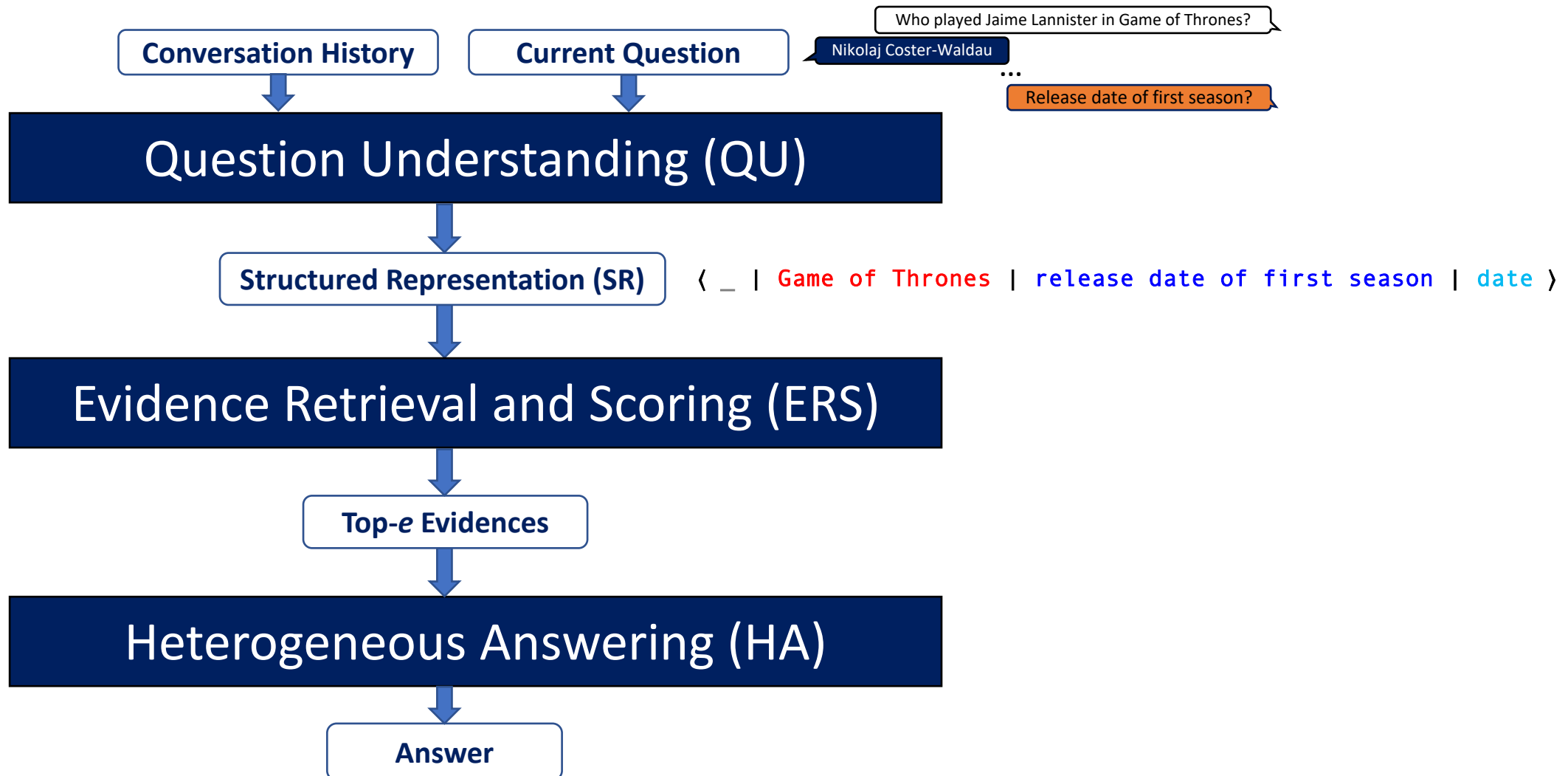


# Conversational Flow Graph (CFG)

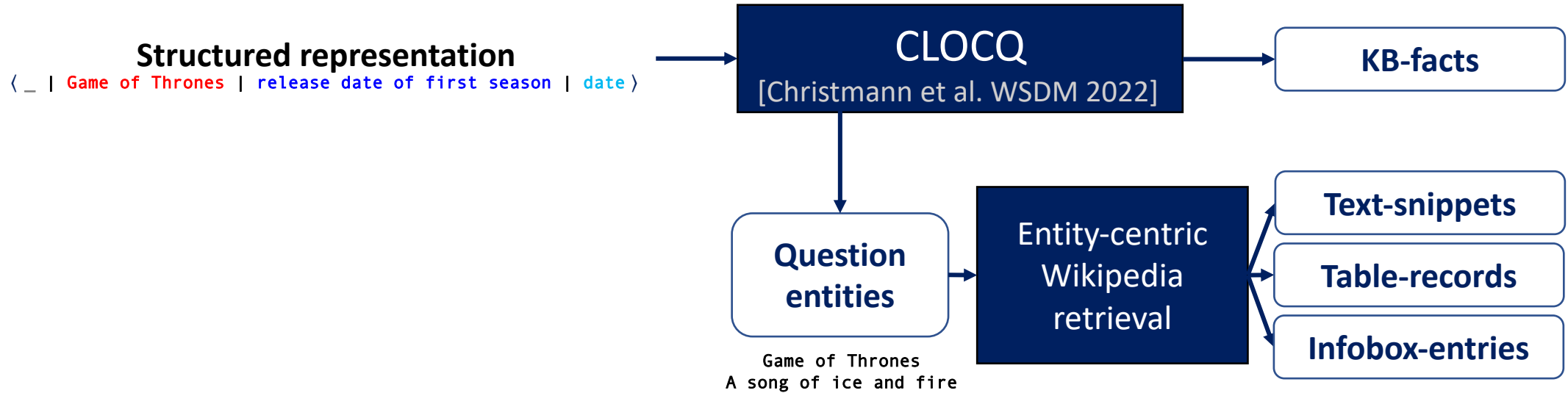


⇒ Adds **explainability** to **developers** and **end users**

# System Overview



# Evidence Retrieval



# Evidence Retrieval

## Original surface form

### KB

Game of Thrones (Q23572)
American fantasy drama television series
GoT   GOT
Statements
instance of  television series

## Verbalized/processed form

Game of Thrones, instance of, **television series**

### Text

**Game of Thrones** is an American **fantasy drama** television series created by **David Benioff** and **D. B. Weiss** for **HBO**. ...

Game of Thrones, Game of Thrones is an American **fantasy drama** television series created by **David Benioff** and **D. B. Weiss** for **HBO**.

### Table

Game of Thrones adaptation and season schedule					
Season	Ordered	Filming	First aired	Last aired	Novel(s) adapted
Season 1	March 2, 2010	Second half of 2010	April 17, 2011	June 19, 2011	<i>A Game of Thrones</i>
Season 2	April 19, 2011	Second half of 2011	April 1, 2012	June 3, 2012	<i>A Clash of Kings</i>
...					

Game of Thrones, Season *is* **Season 1**, Ordered *is* **March 2, 2010**, Filming *is* Second half of **2010**...

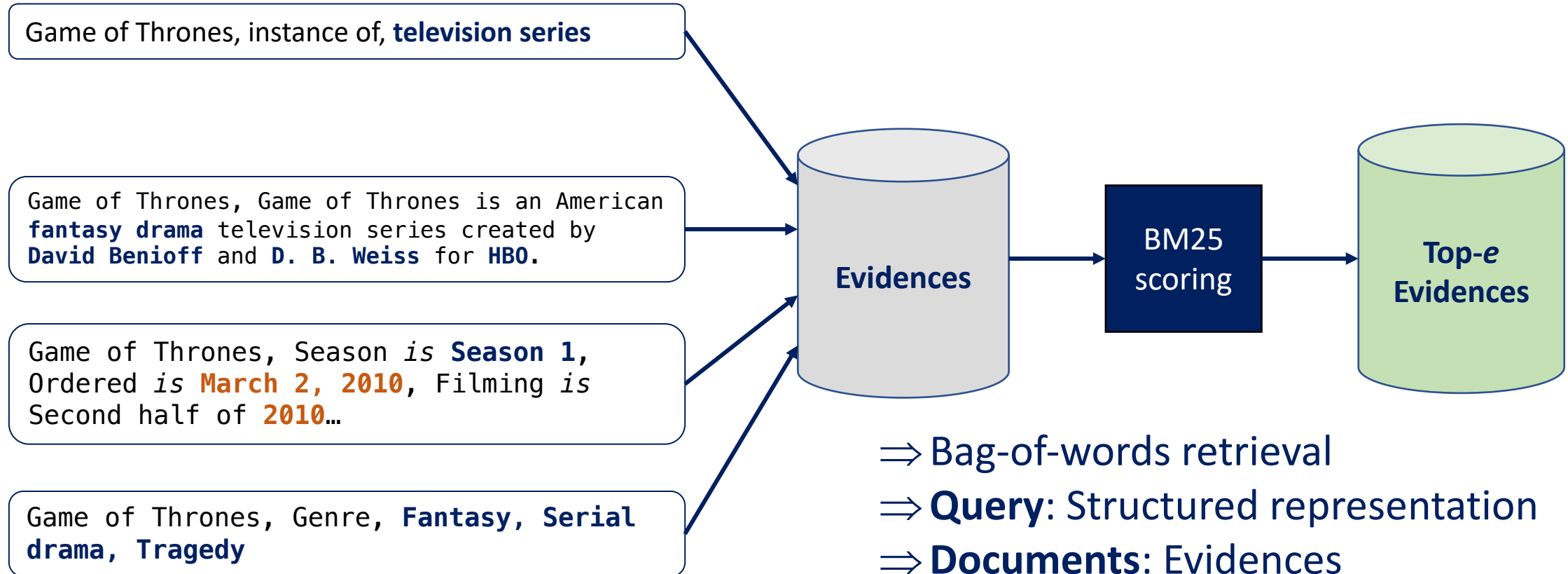
### Infobox

Game of Thrones	
	
Genre	<b>Fantasy</b> <sup>[1]</sup> <b>Serial drama</b> <sup>[2]</sup> <b>Tragedy</b> <sup>[3][4]</sup>
Created by	<b>David Benioff</b> <b>D. B. Weiss</b>

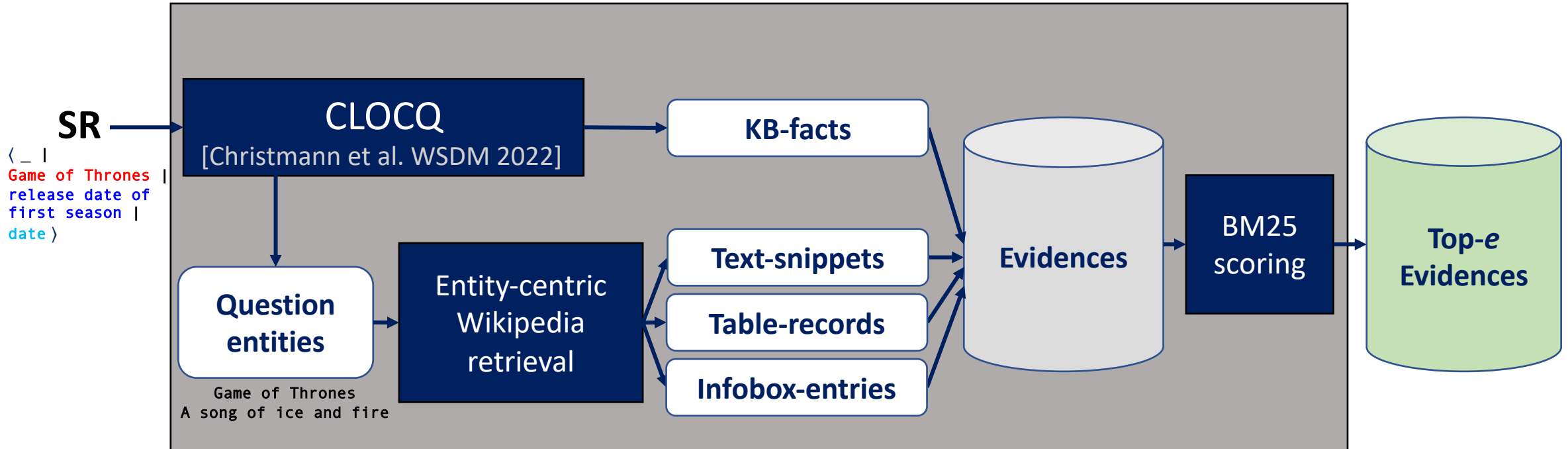
Game of Thrones, Genre, **Fantasy**, **Serial drama**, **Tragedy**

# Evidence Retrieval and Scoring

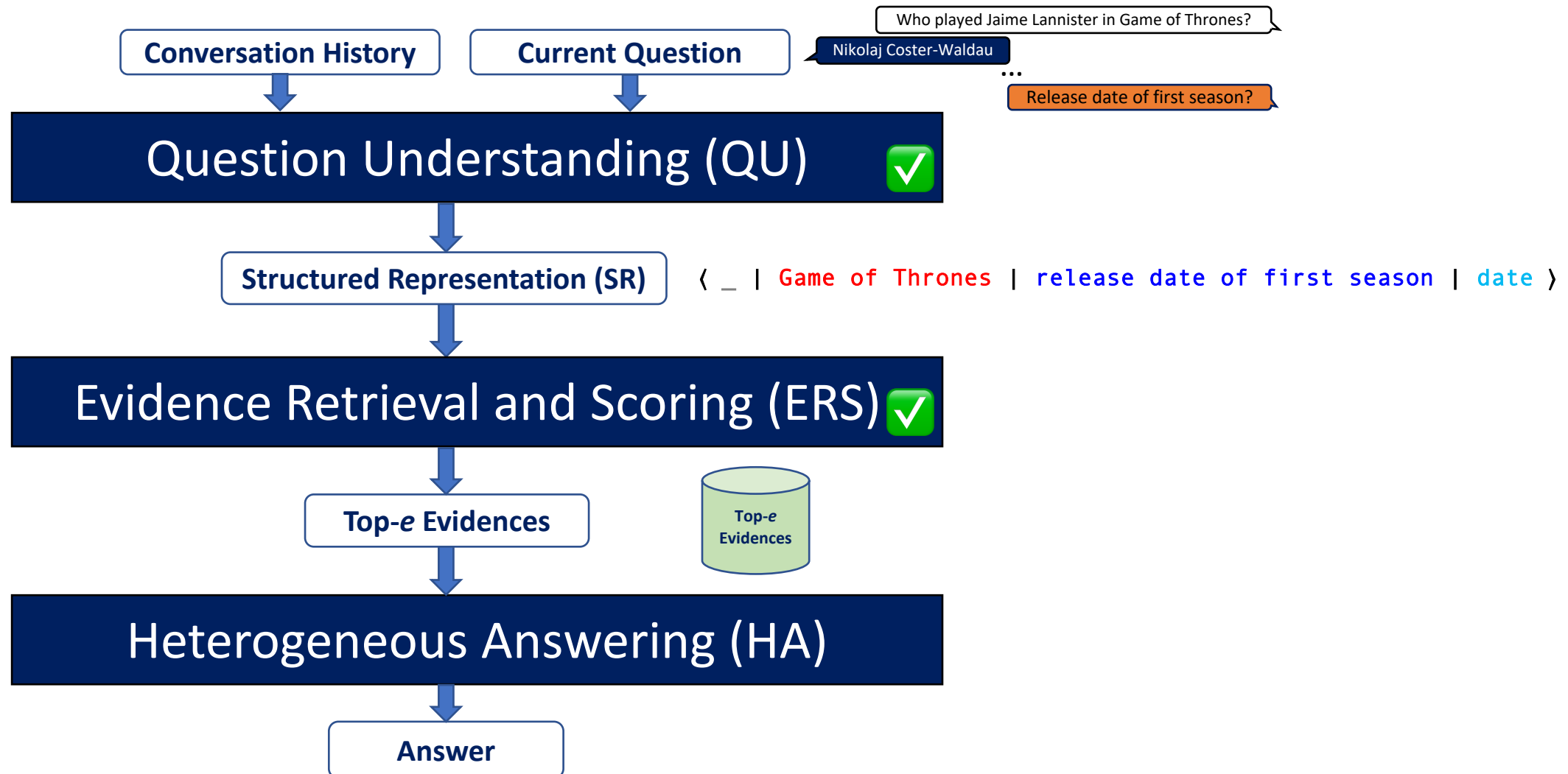
## Verbalized/processed form



# Evidence Retrieval and Scoring



# System Overview

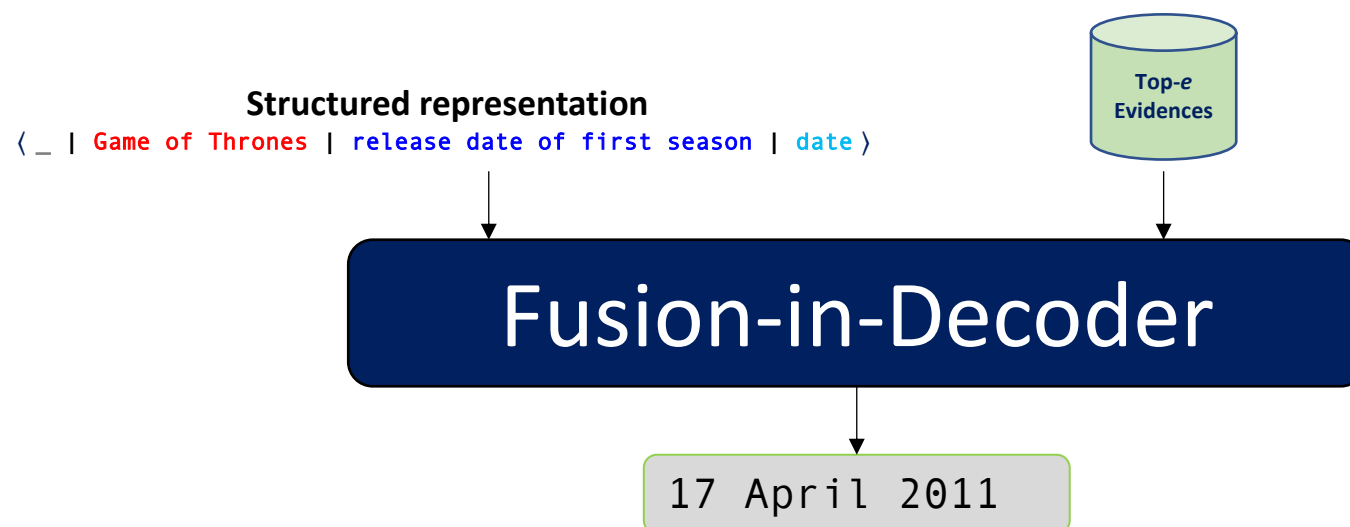


# Heterogeneous Answering

---

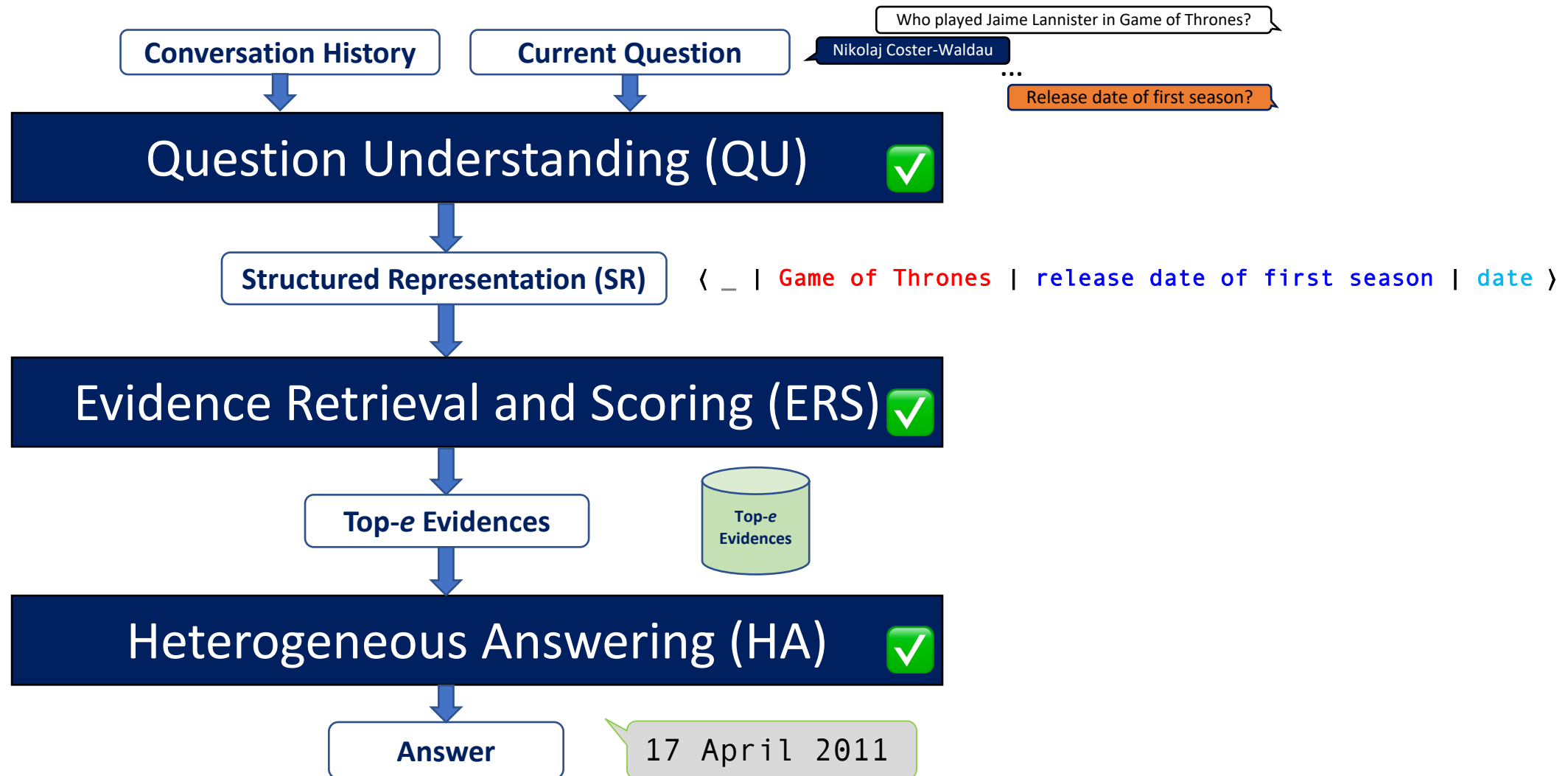
## Fusion-in-Decoder [Izacard et al. EACL 2021]

- ★ **Generative answering:** seq2seq model that encodes evidence-SR pairs
- ★ Decodes **autoregressively** to generate answer
- ★ We **normalize to KB** for comparability





# System Overview

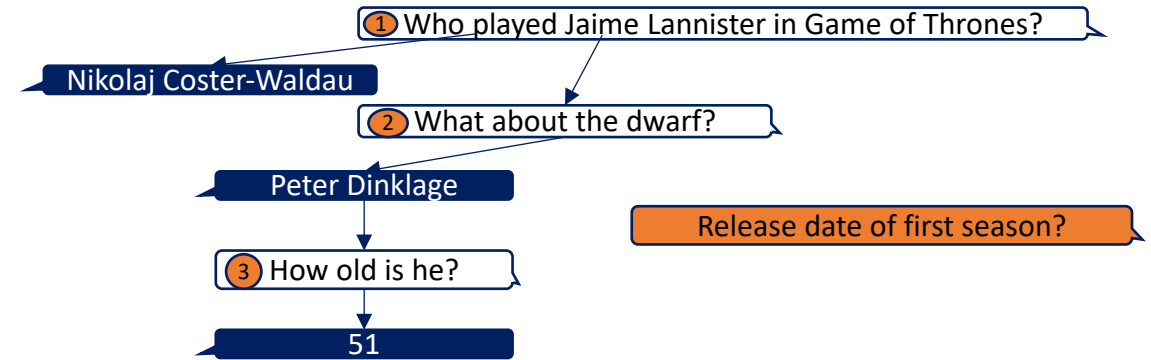


# Distant Supervision

Only **raw** sequence of **QA-pairs** given  
⇒ But **SRs required!**

**Idea:**

- ⇒ **Augment follow-up** with parts from history
- ⇒ **Check whether** this makes the information need **complete**

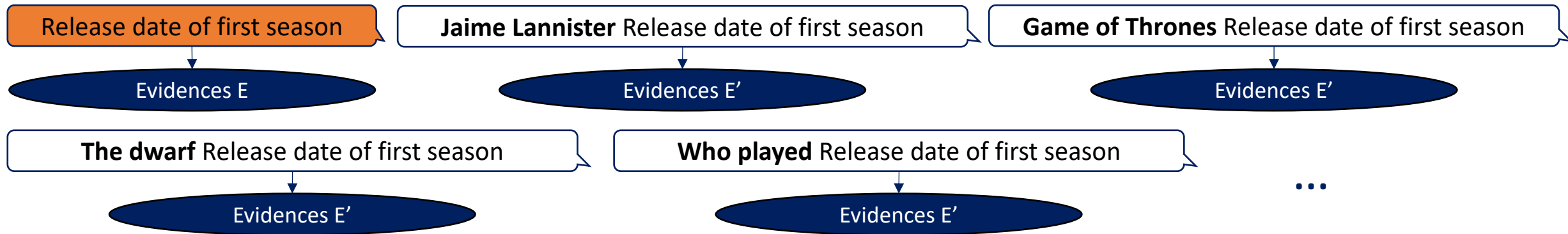
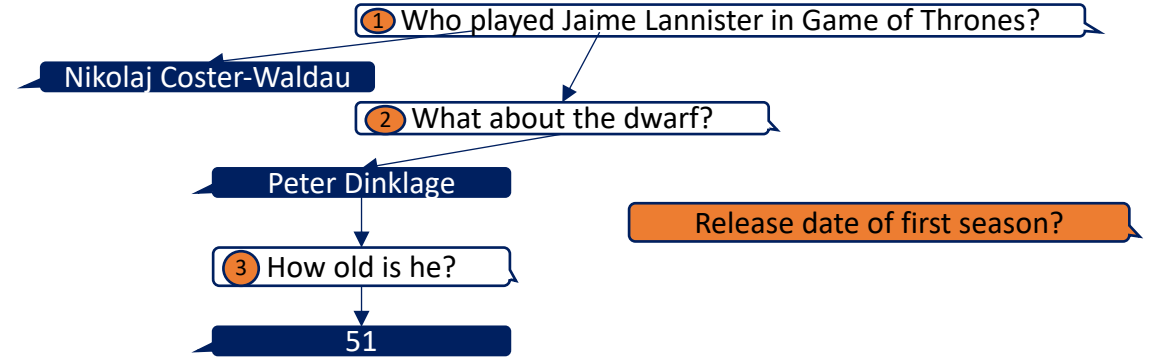


# Distant Supervision

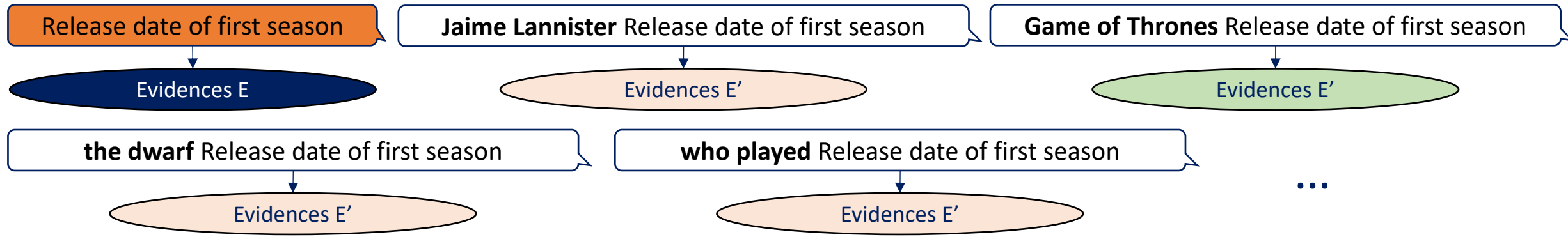
Only **raw** sequence of **QA-pairs** given  
⇒ But **SRs required!**

**Idea:**

- ⇒ **Augment follow-up** with parts from history
- ⇒ **Check whether** this makes the information need **complete**



# Distant Supervision



- ⇒ Detect **expansions** that trigger **new answering evidences**
- ⇒ **Merge** such expansions to get 'gold' **structured representation**

Using our **entity-centric retrieval**, we understand '**why**' evidences are retrieved

- ⇒ For which entity (all evidences), for which KB-item (KB-facts)
- ⇒ Allows for **structuring** the target **SR**

# ConvMix

---

**First dataset for conversational QA over heterogeneous sources**

⇒ 3,000 conversations

⇒ 16,000 questions

⇒ 5 domains: Books, Movies, Music, TV series, Soccer

Collected data:

- ★ Completed (intent-explicit) forms
- ★ Paraphrases
- ★ Question entities
- ★ Answers (free-text and normalized)
- ★ Answer source

# ConvMix

---

## Heterogeneity

- ★ Let users **detect answers** to their questions naturally
- ★ **Collect information source** they found the answer in
  - ⇒ Wikidata; Wikipedia Text, Table, or Infobox
- ★ Only ensure that users cover **each information source at least once** per task (5 conversations)

# Experimental Setup

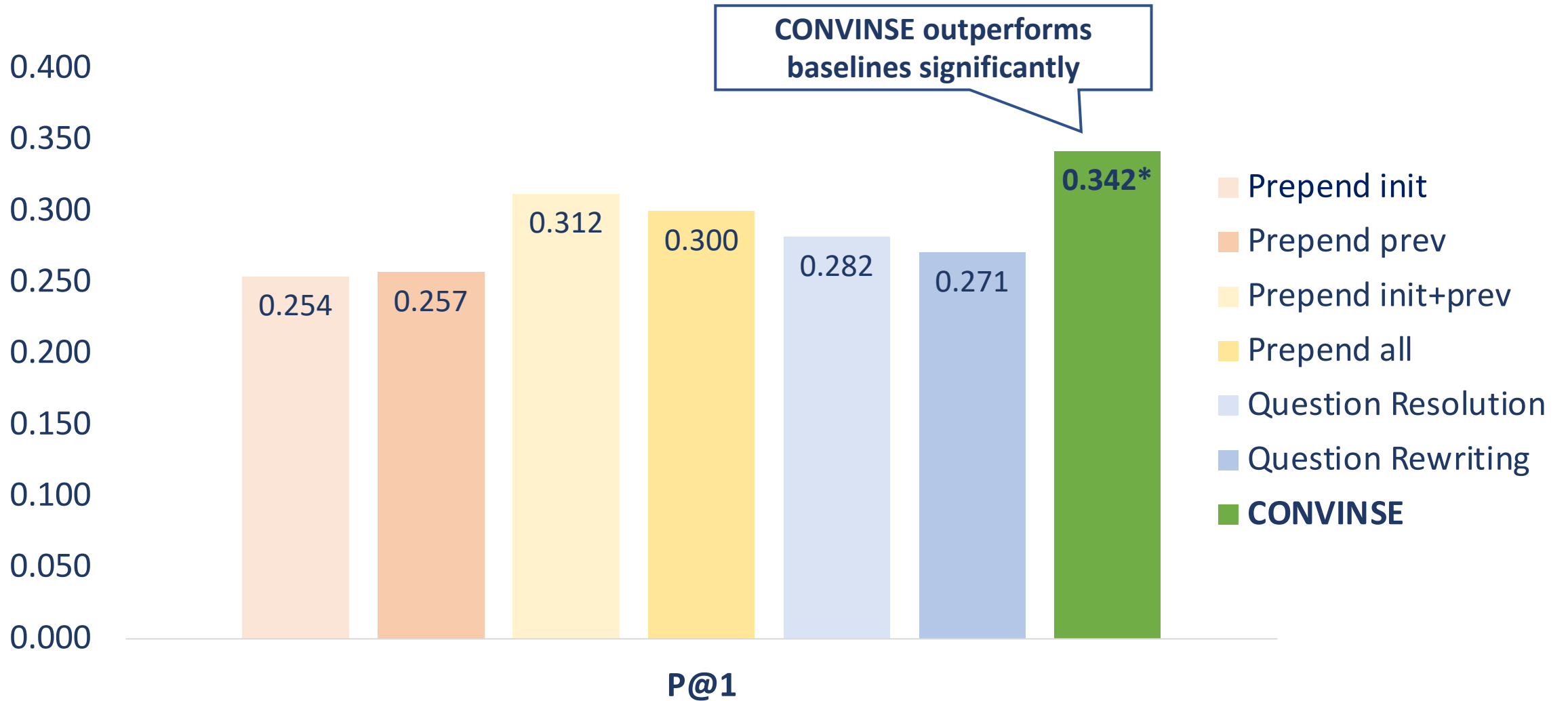
---

## **Baselines: Question Understanding**

- ☆ Prepend **initial** turn
- ☆ Prepend **last** turn
- ☆ Prepend **initial** and **last** turn
- ☆ Prepend **all** previous turns
- ☆ **Question Resolution** – Add relevant terms from history to query [Voskarides et al. SIGIR 2020]
- ☆ **Question Rewriting** – Rewrite to full-fledged grammatical question [Raposo et al. ECIR 2022]

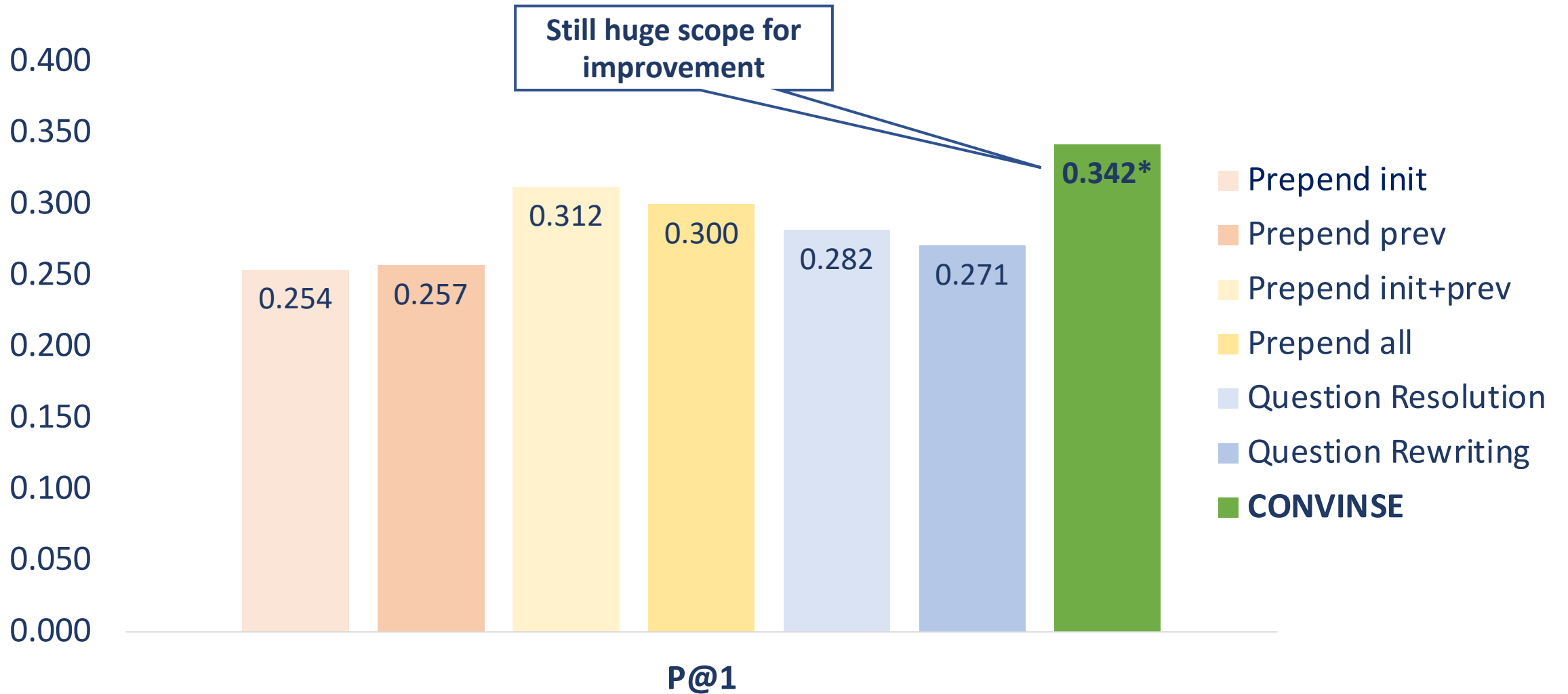
⇒ **ERS** and **HA** kept **static** across baselines (adapting baselines to task)

# Results – Precision at 1

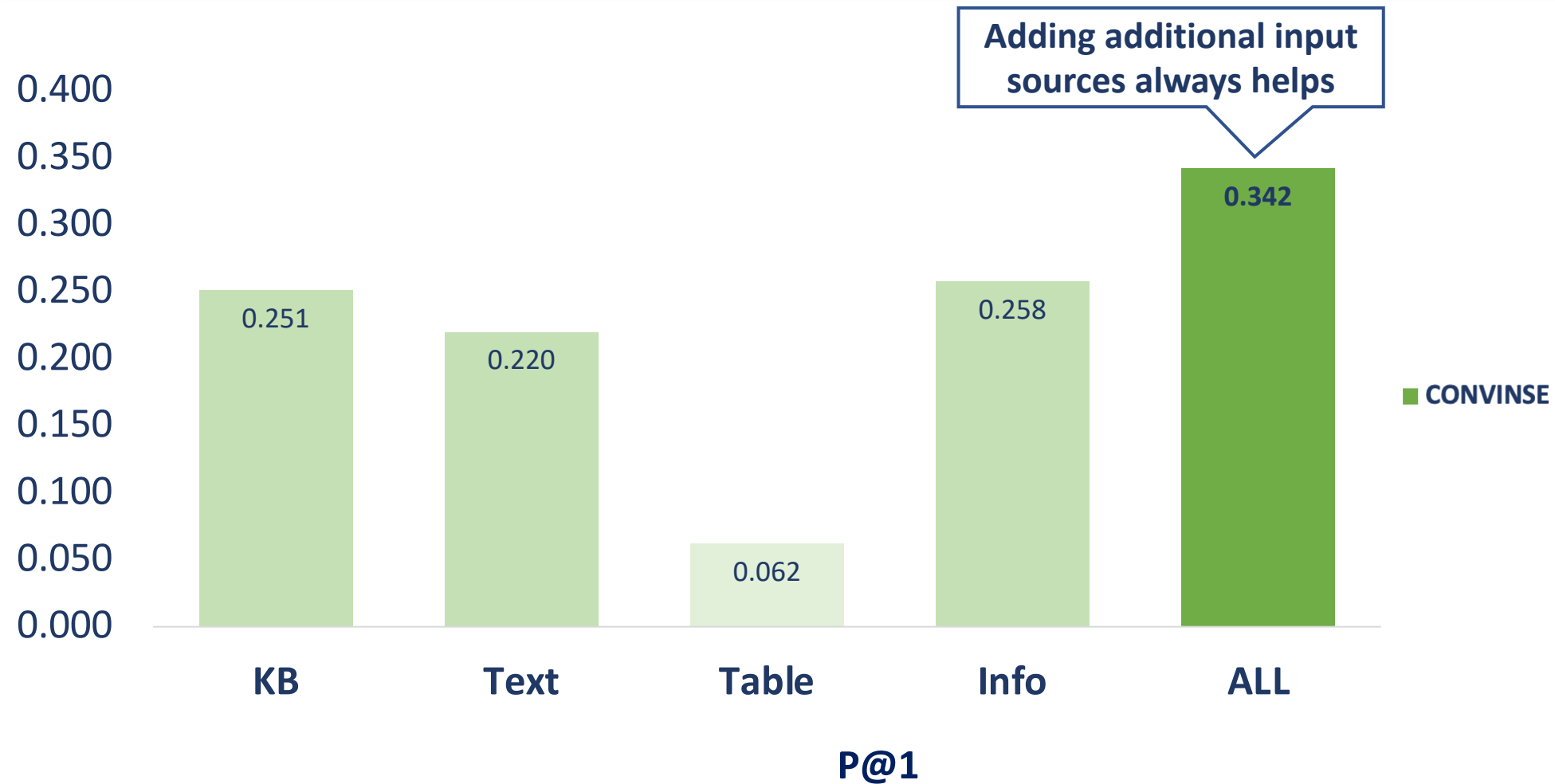




# Results – Precision at 1

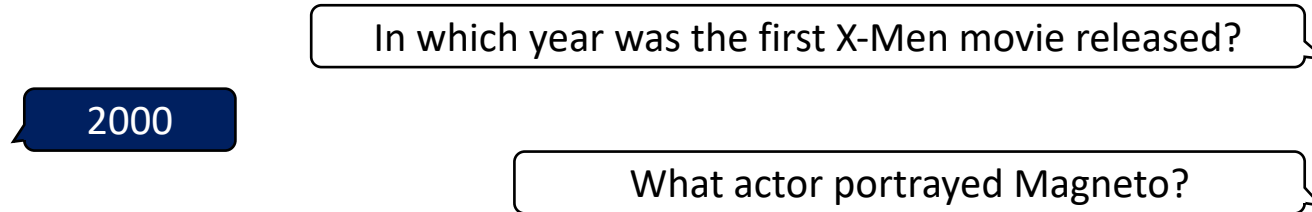


# Results – Per Information Source(s)



# Anecdotal Example

---



## Intent-explicit forms

- ☆ Question Resolution      What actor portrayed Magneto? x-men movie
- ☆ Question Rewriting      Which actor played the character Magneto in the X-Men movie?
- ★ CONVINSE      `< X-Men | Magneto | actor portrayed | human >`

# Conclusion

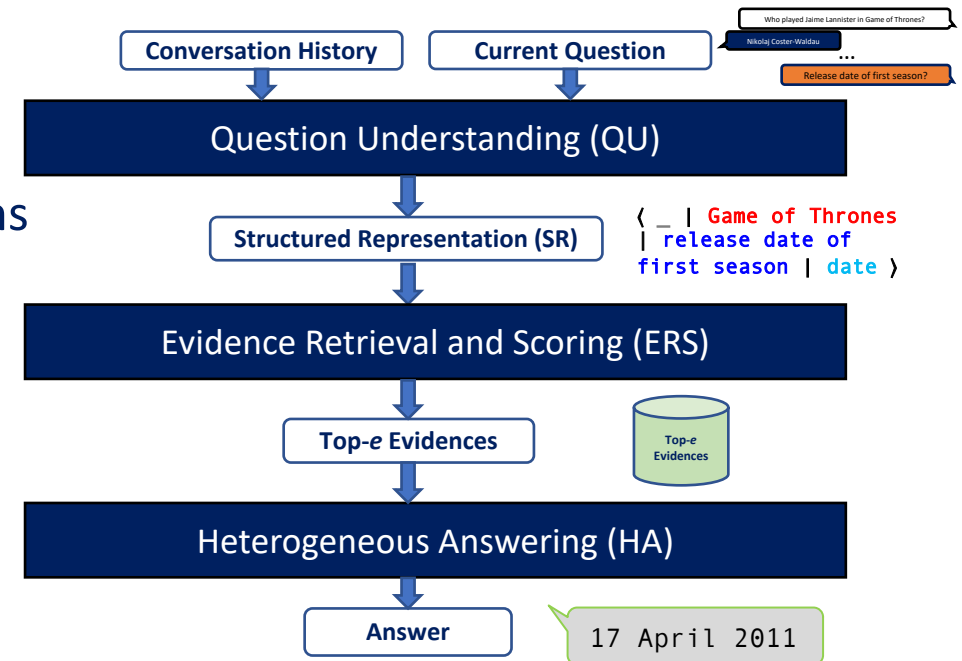
## CONVINSE: First method for ConvQA over heterogeneous sources

- ★ Propose a general 3-staged pipeline for this task
- ★ Intent-explicit **structured representation** of questions
- ★ **Distant supervision** from raw sequences of QA-pairs
- ★ Intermediate representations can add **explainability**

## ConvMix: First ConvQA benchmark for heterogeneous sources

- ★ 3,000 conversations
- ★ 16,000 questions

Code&Data: [convinse.mpi-inf.mpg.de](http://convinse.mpi-inf.mpg.de)



Thanks for the SIGIR student travel grant!

**Thank you!**