TopiOCQA: Open-domain Conversational Question Answering with Topic Switching

Vaibhav Adlakha[1,4], Shehzaad Dhuliawala[2], Kaheer Suleman[3], Harm de Vries[4], Siva Reddy[1]

Mila/McGill University[1]   ETH Zürich[2]   Microsoft Montréal[3]   ServiceNow Research[4]

Conversational Question Answering

A dialogue sub-genre which is distinguished from other kinds of information-seeking dialogue (e.g.travel information) by its more exploratory and less task-oriented nature.
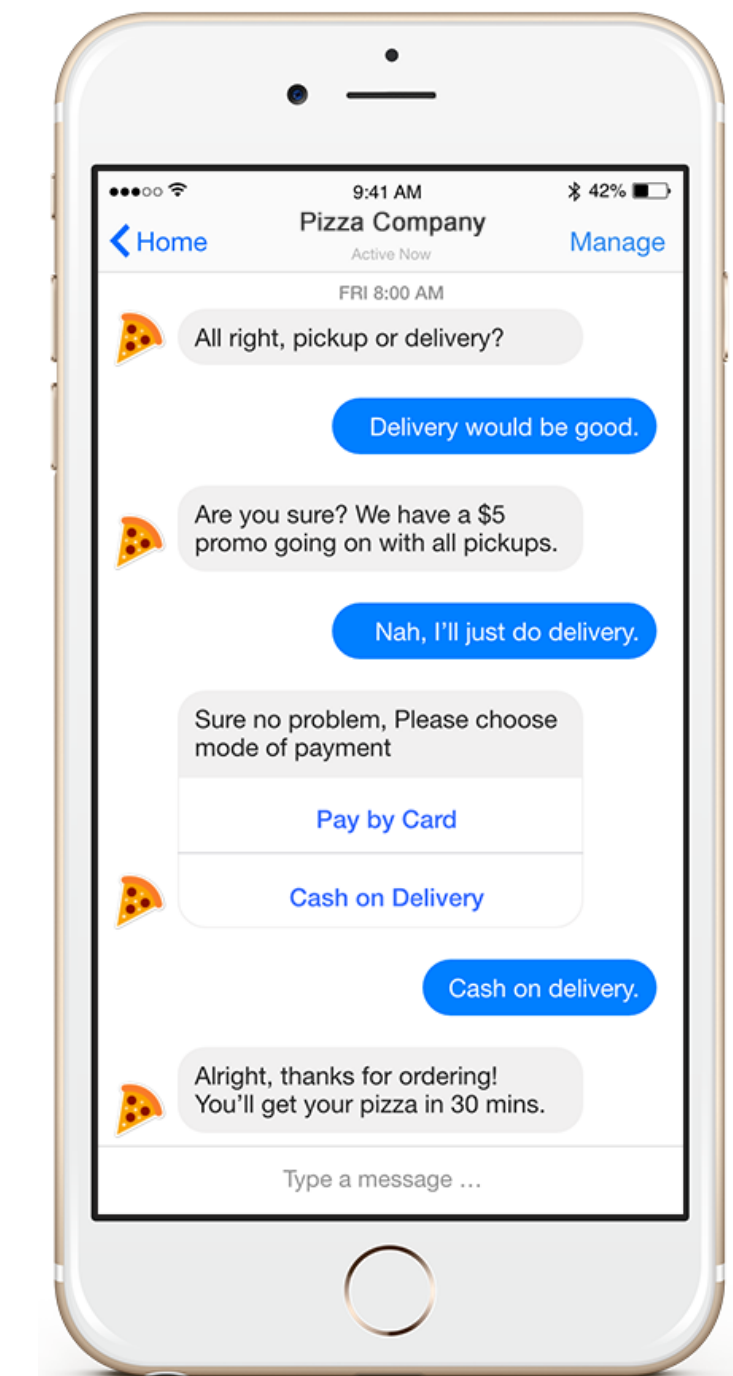
Stede and Schlangen, 2004

Open-ended

Task-oriented

Conversational Question Answering

```
dialogue 008
```

1	E:	Hello. This is Potsdam Tourist Information. How may I help you?
2	U:	Well, I'd like to know whether a weekend in Potsdam is worthwhile?

Conversational Question Answering

```
dialogue 008
```

1	E:	Hello. This is Potsdam Tourist Information. How may I help you?
2	U:	Well, I'd like to know whether a weekend in Potsdam is worthwhile?
3	E:	A weekend in Potsdam is always worth your while, there are lots of nice things you can do here.
4	U:	Like what?

Conversational Question Answering

dialogue 008

1	E:	Hello. This is Potsdam Tourist Information. How may I help you?
2	U:	Well, I'd like to know whether a weekend in Potsdam is worthwhile?
3	E:	A weekend in Potsdam is always worth your while, there are lots of nice things you can do here.
4	U:	Like what?
5	E:	For example, you can make a boat trip, or a sightseeing tour. Always popular with tourists is Park Sanssouci.
6	U:	Is all of this in the inner city? Can you walk to these things, or do you have to use some kind of transport?

Overview of contributions

- A new dataset - TopiOCQA - for open-domain conversational question-answering (QA)

  - More **challenging** and **closer to real-world** scenario than previous datasets

  - Incorporates **multiple topics** in each conversation

Overview of contributions

- A new dataset - TopiOCQA - for open-domain conversational question-answering (QA)

  ▪ More **challenging** and **closer to real-world** scenario than previous datasets

  ▪ Incorporates **multiple topics** in each conversation

- Extensive experimentation on TopiOCQA to answer these questions:

  ▪ Do models with access to document corpus (open-book) perform better than the ones without access? — **Yes**

  ▪ What is the best representation of a conversational question? — **Entire dialog history**

  ▪ Where are the current best models limited? — **at the retrieval stage**

Conversational Question Answering

Existing datasets include a reference document

The Virginia governor's race, billed as the marquee battle of an otherwise anticlimactic 2013 election cycle, is shaping up to be a foregone conclusion. Democrat Terry McAuliffe, the longtime political fixer and moneyman, hasn't trailed in a poll since May. Barring a political miracle, Republican Ken Cuccinelli will be delivering a concession speech on Tuesday evening in Richmond. In recent ...

Q_1: What are the candidates **running** for?
A_1: Governor
R_1: The Virginia governor's race

Q_2: **Where**?
A_2: Virginia
R_2: The Virginia governor's race

Q_3: Who is the democratic candidate?
A_3: **Terry McAuliffe**
R_3: Democrat Terry McAuliffe

Reddy et al., 2019 (CoQA)

Section: Augusto Pinochet : Intellectual life...

STUDENT: **Was he known for being intelligent?**
TEACHER: ↪ No, Pinochet was publicly known as a man with a lack of culture.
STUDENT: **why did people feel that way?**
TEACHER: ↪ reinforced by the fact that he also portrayed himself as a common man
STUDENT: **did he have any hobbies?**
TEACHER: ↪ Yes, Before wresting power from Allende, Pinochet had written two books.
STUDENT: **what is the name of a book written by him?**
TEACHER: ↪ Geopolitica (1968) and Campana de Tarapaca (1972).
STUDENT: **what were the books about?**
TEACHER: ↪ Chile's military literature.
STUDENT: **was there anything noteworthy regarding his books?**
TEACHER: ↪ Yes, In Geopolitica Pinochet plagiarized (...) Gregorio Rodriguez Tascon

Choi et al., 2018 (QuAC)

(Open-domain) Conversational Question Answering

Few recent approaches adapt existing datasets by disambiguating first question

The Virginia governor's race, billed as the marquee battle of an otherwise anticlimactic 2013 election cycle, is shaping up to be a foregone conclusion. Democrat Terry McAuliffe, the longtime political fixer and moneyman, hasn't trailed in a poll since May. Barring a political miracle, Republican Ken Cuccinelli will be delivering a concession speech on Tuesday evening in Richmond. In recent ...

Q: What are the candidates in **Virginia governor's race** running for?

~~Q_1: What are the candidates **running** for?~~

A_1: Governor
R_1: The Virginia governor's race

Q_2: **Where**?
A_2: Virginia
R_2: The Virginia governor's race

Q_3: Who is the democratic candidate?
A_3: **Terry McAuliffe**
R_3: Democrat Terry McAuliffe

(Open-domain) Conversational Question Answering

But the underlying reference passage is still same for entire conversation

The Virginia governor's race, billed as the marquee battle of an otherwise anticlimactic 2013 election cycle, is shaping up to be a foregone conclusion. Democrat Terry McAuliffe, the longtime political fixer and moneyman, hasn't trailed in a poll since May. Barring a political miracle, Republican Ken Cuccinelli will be delivering a concession speech on Tuesday evening in Richmond. In recent ...

Q: What are the candidates in **Virginia governor's race** running for?

~~Q_1: What are the candidates **running** for?~~

A_1: Governor
R_1: The Virginia governor's race

Q_2: **Where**?
A_2: Virginia
R_2: The Virginia governor's race

Q_3: Who is the democratic candidate?
A_3: **Terry McAuliffe**
R_3: Democrat Terry McAuliffe

(Open-domain) Conversational Question Answering

- Conversations in open-domain span multiple topics/documents

  Users search for related topics in the same session (Spink et al, 2002)

  Topic transitions are frequently observed in human-human conversations (Paranjape and Manning, 2021)

  Hence, we cannot adapt from existing datasets

TopiOCQA

- **50,466** turns (QA Pairs)

- **3,920** conversations

- Average of **13** turns per conversation

- Average of **4** different topics per conversation

Q_1: **when was the byzantine empire born what was it originally called?**
A_1: 5th century AD and was called Eastern Roman Empire, or Byzantium
Topic: Byzantine Empire
. . . . .

Q_3: **which battle or event marked the fall of this empire?**
A_3: A six-year-long civil war followed by attack from Sultan Mehmed's army
Topic: Byzantine Empire

Q_4: **did he conquer other territories as well?**
A_4: Yes. Anatolia and in Southeast Europe as far west as Bosnia
Topic: Mehmed the Conqueror

Q_5: **where is the first area located in present day terms?**
A_5: Turkey
Topic: Anatolia
. . . . .

Q_7: **what is the present day capital of the country?**
A_7: Ankara
Topic: Turkey

Q_8: **can you name some of the other major cities here?**
A_8: Istanbul
Topic: Turkey

Q_9: **were any of these cities associated with the first empire you were discussing?**
A_9: The Ottomans made the city of Ankara the capital first of the Anatolia Eyalet and then the Angora Vilayet
Topic: Ankara

A sample conversation from TopiOCQA

TopiOCQA

☑ Open-domain

- No reference document provided

Q1: **when was the byzantine empire born what was it originally called?**
A1: 5th century AD and was called Eastern Roman Empire, or Byzantium
Topic: Byzantine Empire

. . . . .

Q3: **which battle or event marked the fall of this empire?**
A3: A six-year-long civil war followed by attack from Sultan Mehmed's army
Topic: Byzantine Empire

Q4: **did he conquer other territories as well?**
A4: Yes. Anatolia and in Southeast Europe as far west as Bosnia
Topic: Mehmed the Conqueror

Q5: **where is the first area located in present day terms?**
A5: Turkey
Topic: Anatolia

. . . . .

Q7: **what is the present day capital of the country?**
A7: Ankara
Topic: Turkey

Q8: **can you name some of the other major cities here?**
A8: Istanbul
Topic: Turkey

Q9: **were any of these cities associated with the first empire you were discussing?**
A9: The Ottomans made the city of Ankara the capital first of the Anatolia Eyalet and then the Angora Vilayet
Topic: Ankara

A sample conversation from TopiOCQA

TopiOCQA

☑ Open-domain

- No reference document provided

☑ Free-form answers

- Not restricted to spans in a document

Q_1: **when was the byzantine empire born what was it originally called?**
A_1: 5th century AD and was called Eastern Roman Empire, or Byzantium
Topic: Byzantine Empire

. . . . .

Q_3: **which battle or event marked the fall of this empire?**
A_3: A six-year-long civil war followed by attack from Sultan Mehmed's army
Topic: Byzantine Empire

Q_4: **did he conquer other territories as well?**
A_4: Yes. Anatolia and in Southeast Europe as far west as Bosnia
Topic: Mehmed the Conqueror

Q_5: **where is the first area located in present day terms?**
A_5: Turkey
Topic: Anatolia

. . . . .

Q_7: **what is the present day capital of the country?**
A_7: Ankara
Topic: Turkey

Q_8: **can you name some of the other major cities here?**
A_8: Istanbul
Topic: Turkey

Q_9: **were any of these cities associated with the first empire you were discussing?**
A_9: The Ottomans made the city of Ankara the capital first of the Anatolia Eyalet and then the Angora Vilayet
Topic: Ankara

A sample conversation from TopiOCQA

TopiOCQA

☑ Open-domain

• No reference document provided

☑ Free-form answers

• Not restricted to spans in a document

☑ Incorporates topic switching

• Conversation spans multiple related topics

Q_1: **when was the byzantine empire born what was it originally called?**
A_1: 5th century AD and was called Eastern Roman Empire, or Byzantium
`Topic`: Byzantine Empire

. . . . .

Q_3: **which battle or event marked the fall of this empire?**
A_3: A six-year-long civil war followed by attack from Sultan Mehmed's army
`Topic`: Byzantine Empire

Q_4: **did he conquer other territories as well?**
A_4: Yes. Anatolia and in Southeast Europe as far west as Bosnia
`Topic`: Mehmed the Conqueror

Q_5: **where is the first area located in present day terms?**
A_5: Turkey
`Topic`: Anatolia

. . . . .

Q_7: **what is the present day capital of the country?**
A_7: Ankara
`Topic`: Turkey

Q_8: **can you name some of the other major cities here?**
A_8: Istanbul
`Topic`: Turkey

Q_9: **were any of these cities associated with the first empire you were discussing?**
A_9: The Ottomans made the city of Ankara the capital first of the Anatolia Eyalet and then the Angora Vilayet
`Topic`: Ankara

A sample conversation from TopiOCQA

TopiOCQA

☑ Open-domain

• No reference document provided

☑ Free-form answers

• Not restricted to spans in a document

☑ Incorporates topic switching

• Conversation spans multiple related topics

• Exhibits complex topic switching phenomena

Q_1: **when was the byzantine empire born what was it originally called?**
A_1: 5th century AD and was called Eastern Roman Empire, or Byzantium
`Topic`: Byzantine Empire
. . . . .

Q_3: **which battle or event marked the fall of this empire?**
A_3: A six-year-long civil war followed by attack from Sultan Mehmed's army
`Topic`: Byzantine Empire

Q_4: **did he conquer other territories as well?**
A_4: Yes. Anatolia and in Southeast Europe as far west as Bosnia
`Topic`: Mehmed the Conqueror

Q_5: **where is the first area located in present day terms?**
A_5: Turkey
`Topic`: Anatolia
. . . . .

Q_7: **what is the present day capital of the country?**
A_7: Ankara
`Topic`: Turkey

Q_8: **can you name some of the other major cities here?**
A_8: Istanbul
`Topic`: Turkey

Q_9: **were any of these cities associated with the first empire you were discussing?**
A_9: The Ottomans made the city of Ankara the capital first of the Anatolia Eyalet and then the Angora Vilayet
`Topic`: Ankara

A sample conversation from TopiOCQA

TopiOCQA

☑ Open-domain

- No reference document provided

☑ Free-form answers

- Not restricted to spans in a document

☑ Incorporates topic switching

- Conversation spans multiple related topics

- Exhibits complex topic switching phenomena

Q_1: **when was the byzantine empire born what was it originally called?**
A_1: 5th century AD and was called Eastern Roman Empire, or Byzantium
Topic: Byzantine Empire
. . . . .

Q_3: **which battle or event marked the fall of this empire?**
A_3: A six-year-long civil war followed by attack from Sultan Mehmed's army
Topic: Byzantine Empire

Q_4: **did he conquer other territories as well?**
A_4: Yes. Anatolia and in Southeast Europe as far west as Bosnia
Topic: Mehmed the Conqueror

Q_5: **where is the first area located in present day terms?**
A_5: Turkey
Topic: Anatolia
. . . . .

Q_7: **what is the present day capital of the country?**
A_7: Ankara
Topic: Turkey

Q_8: **can you name some of the other major cities here?**
A_8: Istanbul
Topic: Turkey

Q_9: **were any of these cities associated with the first empire you were discussing?**
A_9: The Ottomans made the city of Ankara the capital first of the Anatolia Eyalet and then the Angora Vilayet
Topic: Ankara

A sample conversation from TopiOCQA

TopiOCQA

☑ Open-domain

- No reference document provided

☑ Free-form answers

- Not restricted to spans in a document

☑ Incorporates topic switching

- Conversation spans multiple related topics

- Exhibits complex topic switching phenomena

Q_1: **when was the byzantine empire born what was it originally called?**
A_1: 5th century AD and was called Eastern Roman Empire, or Byzantium
Topic: Byzantine Empire
. . . . .

Q_3: **which battle or event marked the fall of this empire?**
A_3: A six-year-long civil war followed by attack from Sultan Mehmed's army
Topic: Byzantine Empire

Q_4: **did he conquer other territories as well?**
A_4: Yes. Anatolia and in Southeast Europe as far west as Bosnia
Topic: Mehmed the Conqueror

Q_5: **where is the first area located in present day terms?**
A_5: Turkey
Topic: Anatolia
. . . . .

Q_7: **what is the present day capital of the country?**
A_7: Ankara
Topic: Turkey

Q_8: **can you name some of the other major cities here?**
A_8: Istanbul
Topic: Turkey

Q_9: **were any of these cities associated with the first empire you were discussing?**
A_9: The Ottomans made the city of Ankara the capital first of the Anatolia Eyalet and then the Angora Vilayet
Topic: Ankara

A sample conversation from TopiOCQA

TopiOCQA

☑️ Open-domain

- No reference document provided

☑️ Free-form answers

- Not restricted to spans in a document

☑️ Incorporates topic switching

- Conversation spans multiple related topics

- Exhibits complex topic switching phenomena

Q_1: **when was the byzantine empire born what was it origi-nally called?**
A_1: 5th century AD and was called Eastern Roman Empire, or Byzantium
`Topic`: Byzantine Empire
. . . . .

Q_3: **which battle or event marked the fall of this empire?**
A_3: A six-year-long civil war followed by attack from Sultan Mehmed's army
`Topic`: Byzantine Empire

Q_4: **did he conquer other territories as well?**
A_4: Yes. Anatolia and in Southeast Europe as far west as Bosnia
`Topic`: Mehmed the Conqueror

Q_5: **where is the first area located in present day terms?**
A_5: Turkey
`Topic`: Anatolia
. . . . .

Q_7: **what is the present day capital of the country?**
A_7: Ankara
`Topic`: Turkey

Q_8: **can you name some of the other major cities here?**
A_8: Istanbul
`Topic`: Turkey

Q_9: **were any of these cities associated with the first empire you were discussing?**
A_9: The Ottomans made the city of Ankara the capital first of the Anatolia Eyalet and then the Angora Vilayet
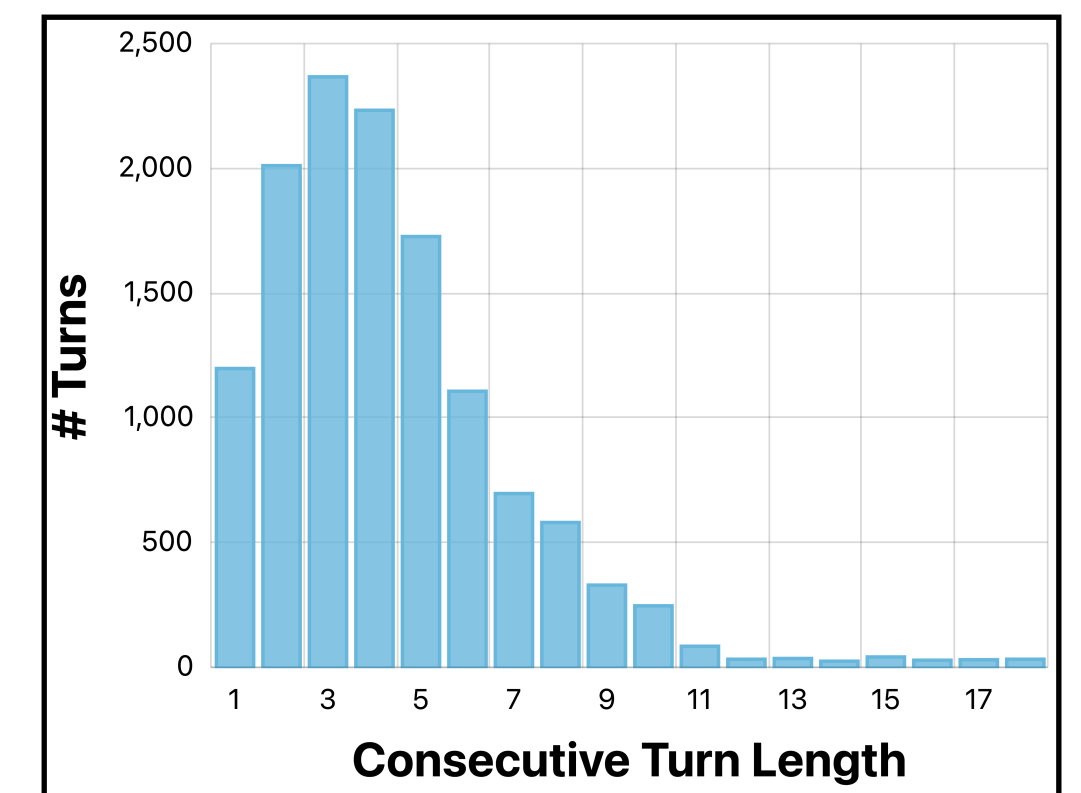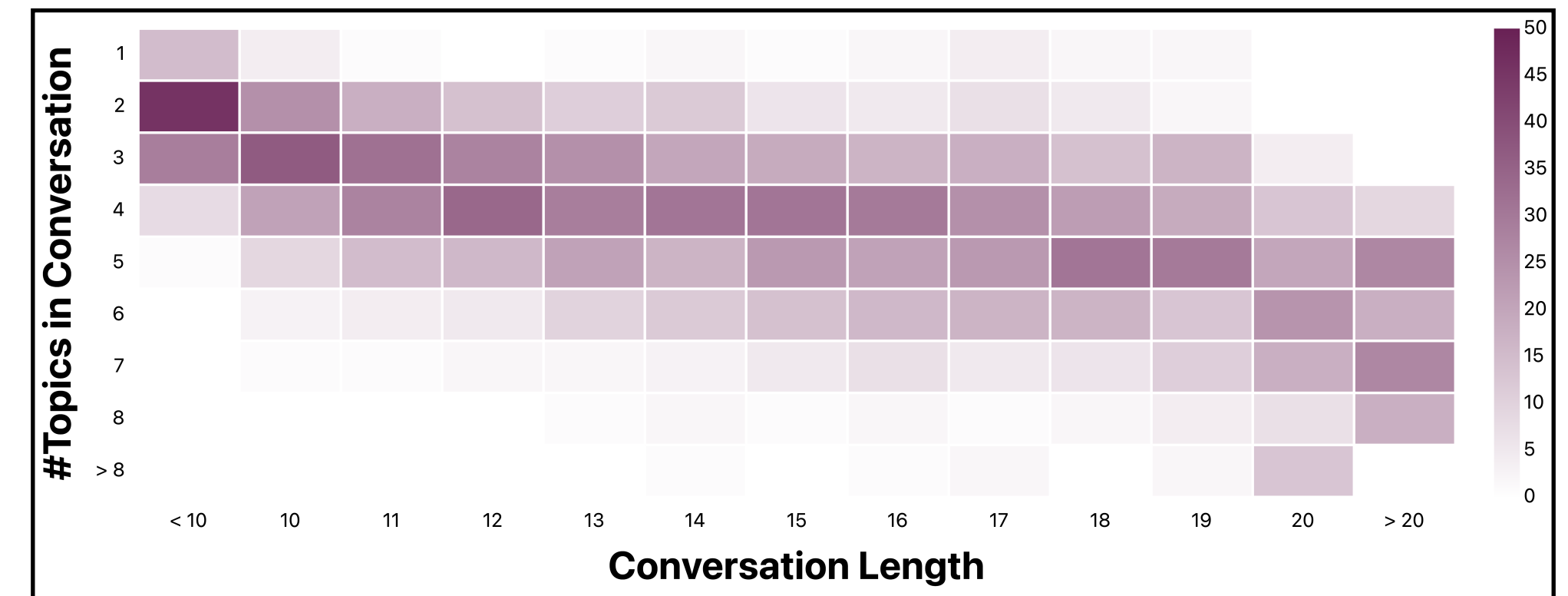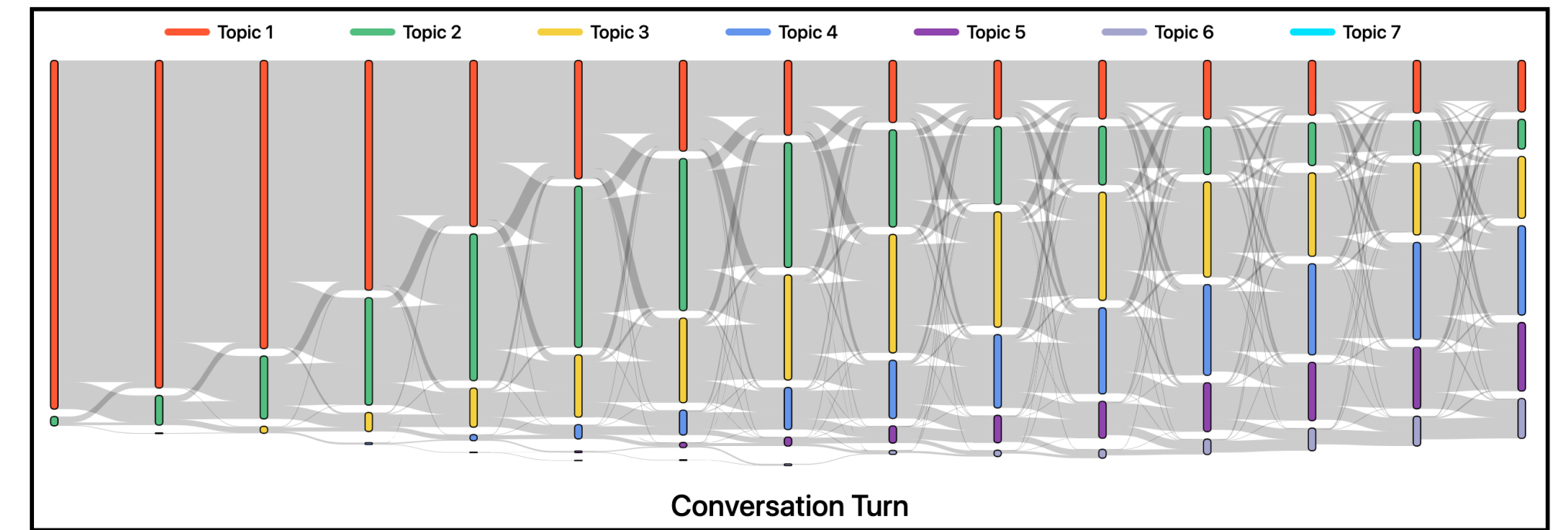`Topic`: Ankara

A sample conversation from TopiOCQA

Dataset Analysis

- How long annotators stay on same topic?

- Do they revisit old topics?

- What prompts a topic switch?

- What type of questions can be expected on switching to a new topic?

Refer to the paper for detailed analysis of TopiOCQA

How to model TopiOCQA?

TopiOCQA consists of

- **Gold document** for each turn

- Set of **gold answers** for each turn

Q_1: who is lead singer of rage against the machine?
A_1: Zack de la Rocha

Q_2: when was it formed?
A_2: 1991

Q_3: was it nominated for any award?

GOLD ANSWER: It has won two Grammy Awards with six nominations altogether. The band has also received three nominations from the MTV Video Music Awards, and another three for Rock and Roll Hall of Fame.

GOLD DOCUMENT
Title: Rage Against the Machine
Section Title: Awards and nominations
Content: Rage Against the Machine has won two Grammy Awards with six nominations altogether. During 2008, the band was also inducted into the Kerrang! "Hall of Fame". As well during 2010, Rage Against the Machine were bestowed with NME's Heroes of the Year Award. The band has also received three nominations from the MTV Video Music Awards, but has yet to win an award. Rage Against The Machine have been nominated for the Rock and Roll Hall of Fame in 2018, 2019 and 2021.
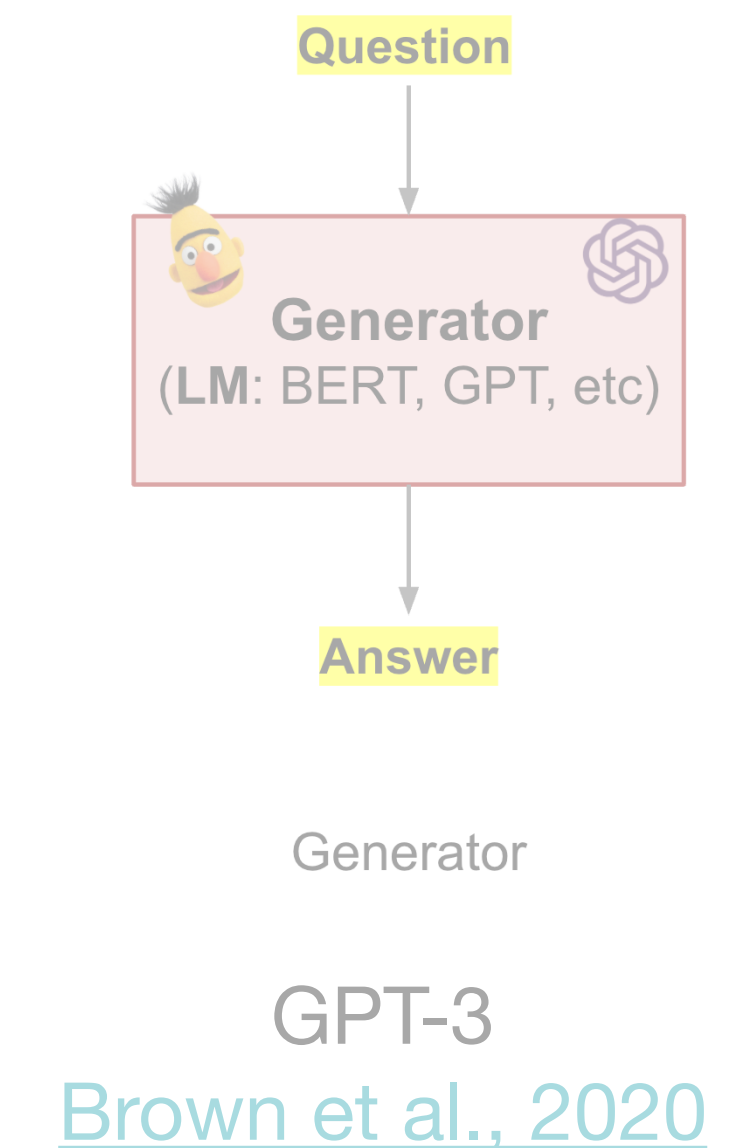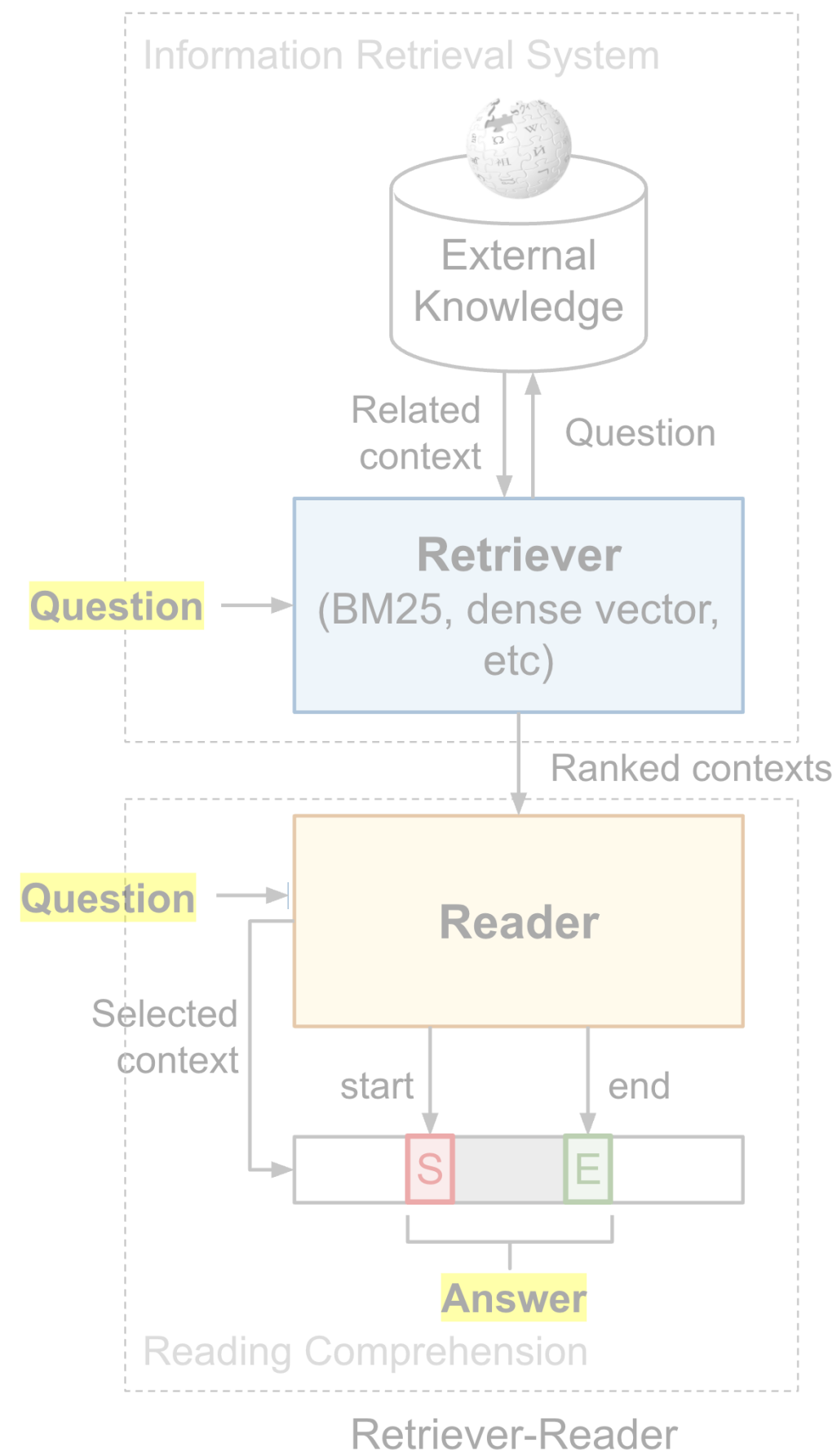
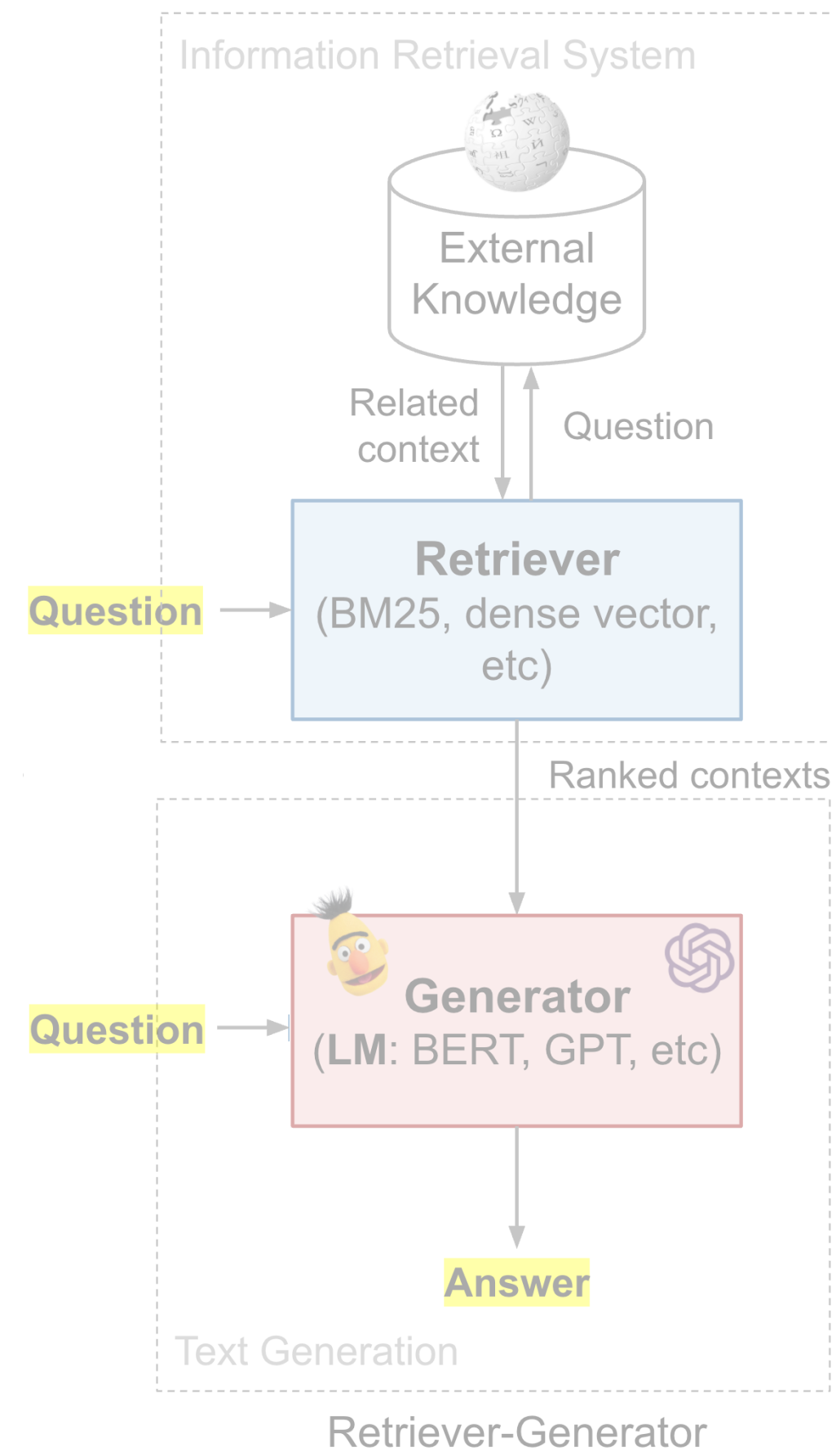# Open-book Models

Dense Passage Retrieval
Karpukhin et al., 2020

Fusion-in-Decoder
Izacard and Grave, 2021

GPT-3
Brown et al., 2020

Image source: https://lilianweng.github.io/posts/2020-10-29-odqa

Closed-book Models

Dense Passage Retrieval
Karpukhin et al., 2020

Fusion-in-Decoder
Izacard and Grave, 2021

GPT-3
Brown et al., 2020

Image source: https://lilianweng.github.io/posts/2020-10-29-odqa

How to represent a conversational question?

Q_1: who is lead singer of rage against the machine?
A_1: Zack de la Rocha

Q_2: when was it formed?
A_2: 1991

Q_3: was it nominated for any award?

How to represent a conversational question?

Q_1: who is lead singer of rage against the machine?
A_1: Zack de la Rocha

Q_2: when was it formed?
A_2: 1991

Q_3: was it nominated for any award?

ORIGINAL: was it nominated for any award

How to represent a conversational question?

Q_1: who is lead singer of rage against the machine?
A_1: Zack de la Rocha

Q_2: when was it formed?
A_2: 1991

Q_3: was it nominated for any award?

ORIGINAL: was it nominated for any award

ALLHISTORY: who is lead singer of rage against the machine [SEP] Zack de la Rocha [SEP] when was it formed? [SEP] 1991 [SEP] was it nominated for any award

How to represent a conversational question?

Q_1: who is lead singer of rage against the machine?
A_1: Zack de la Rocha

Q_2: when was it formed?
A_2: 1991

Q_3: was it nominated for any award?

ORIGINAL: was it nominated for any award

ALLHISTORY: who is lead singer of rage against the machine [SEP] Zack de la Rocha [SEP] when was it formed? [SEP] 1991 [SEP] was it nominated for any award
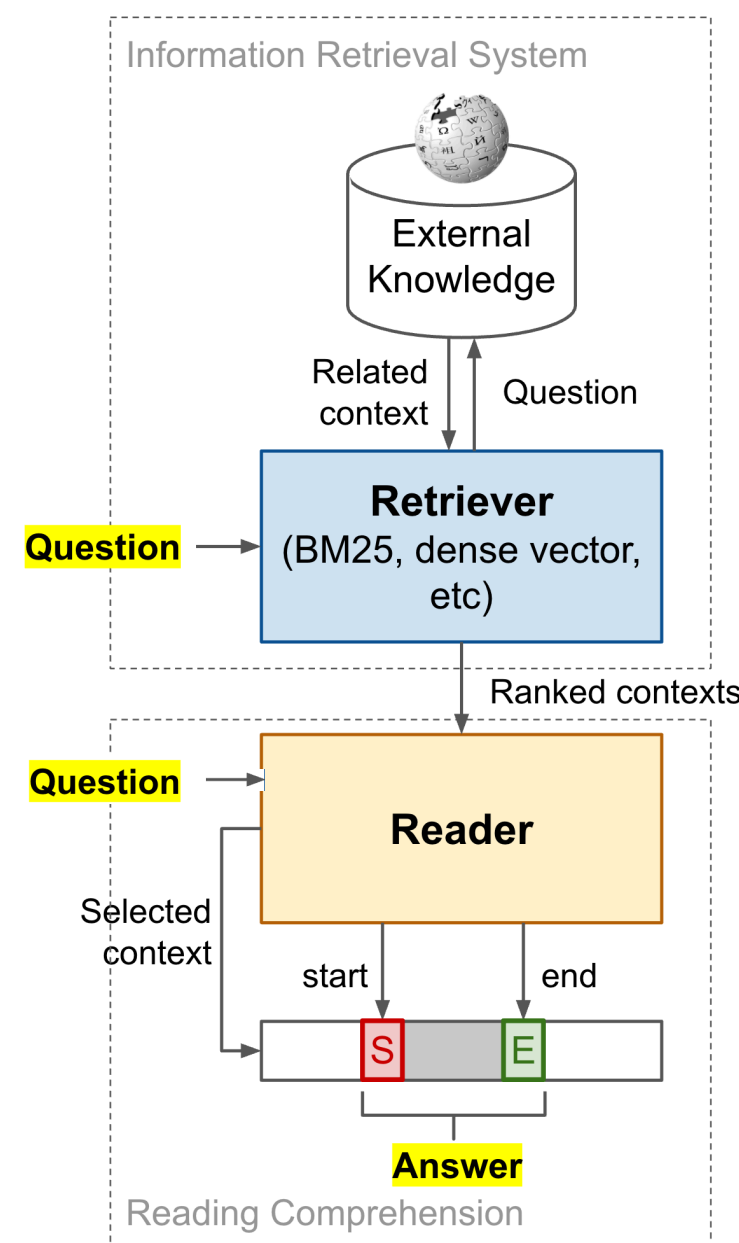
REWRITES: was rage against the machine nominated for any award

Experimental Results on TopiOCQA

- Are **open-book** models better than **closed-book** models?

- How good are current models at **retrieval in conversational settings**?

- Which **question representation** is well-suited for the task?

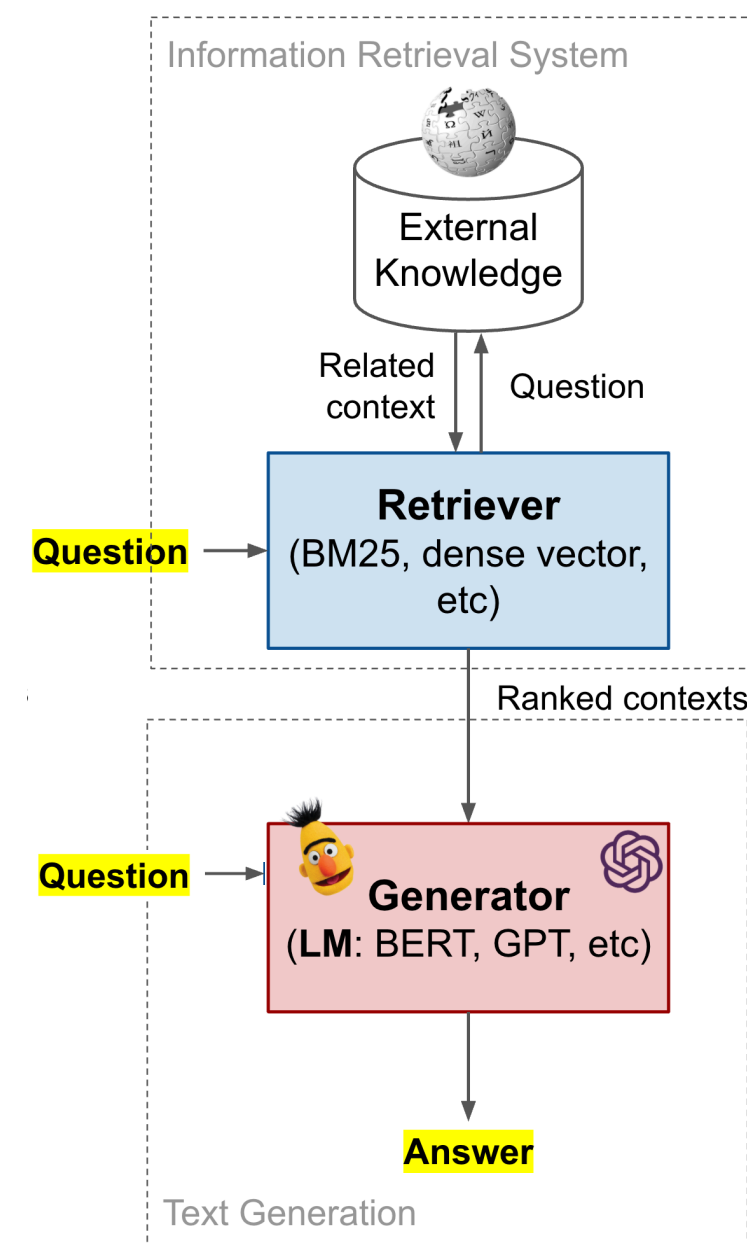- For open-book models, is the bottleneck **retriever** or **reader**?

Which class of models is better?
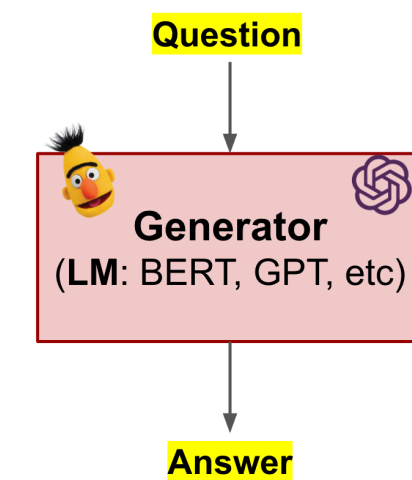
Open-book models outperform closed-book models

Dense Passage Retrieval
Karpukhin et al., 2020

26.7 F1 points below H.P.

Fusion-in-Decoder
Izacard and Grave, 2021

14.8 F1 points below H.P.

GPT-3
Brown et al., 2020

36.7 F1 points below H.P.

H.P. - Human Performance

How good are current models for conversational retrieval?

Dense retriever performs better than sparse retriever

Model	Question Rep	Dev	
		Top-20	Top-100
	ORIGINAL	5.2	9.1
BM25	ALLHISTORY	23.1	36.8
	REWRITES	32.5	49.2
	ORIGINAL	9.9	16.5
DPR Retriever	ALLHISTORY	**70.4**	**82.4**
	REWRITES	49.9	62.4

Which question representation is well-suited for the task?

Q_1: who is lead singer of rage against the machine?
A_1: Zack de la Rocha

Q_2: when was it formed?
A_2: 1991

Q_3: was it nominated for any award?

ORIGINAL: was it nominated for any award

ALLHISTORY: who is lead singer of rage against the machine [SEP] Zack de la Rocha [SEP] when was it formed? [SEP] 1991 [SEP] was it nominated for any award → Works best with dense retrievers
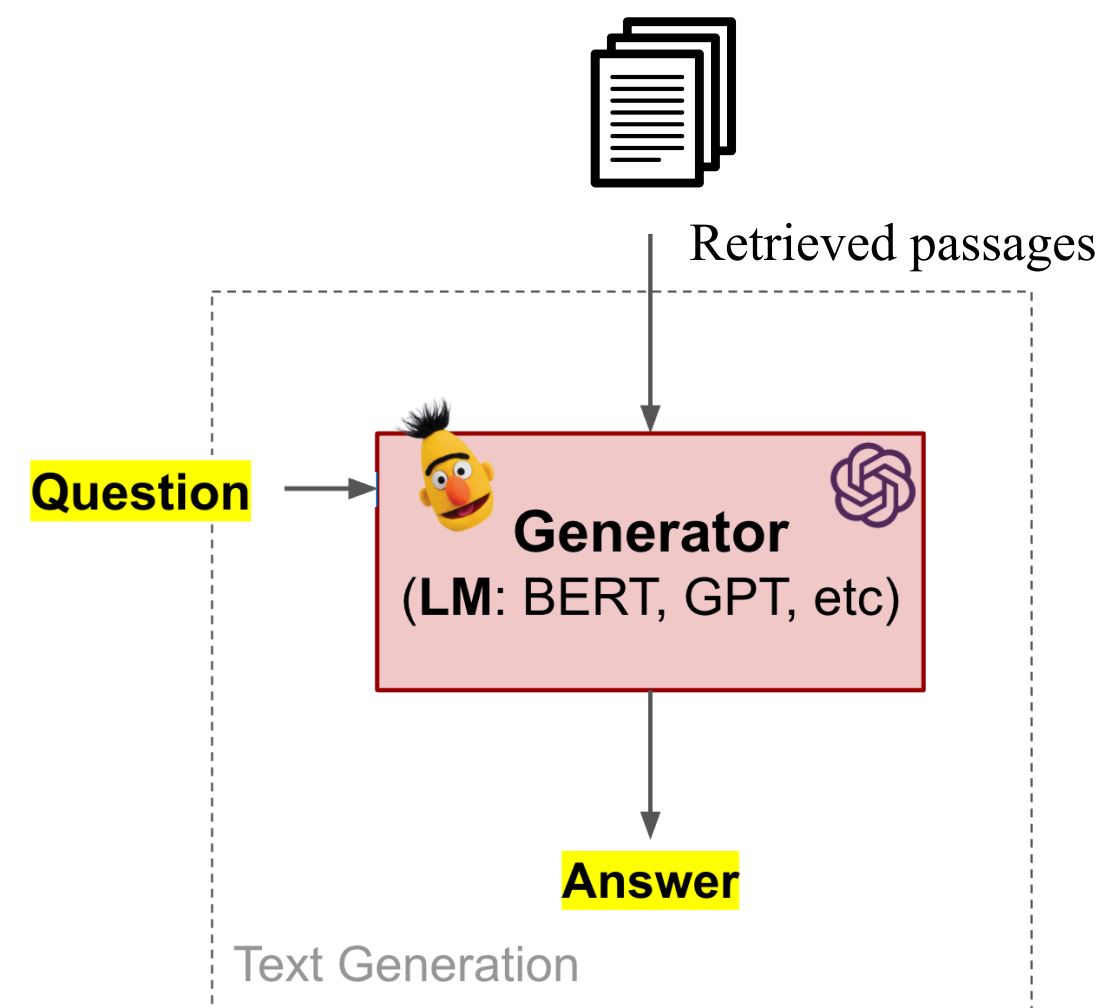
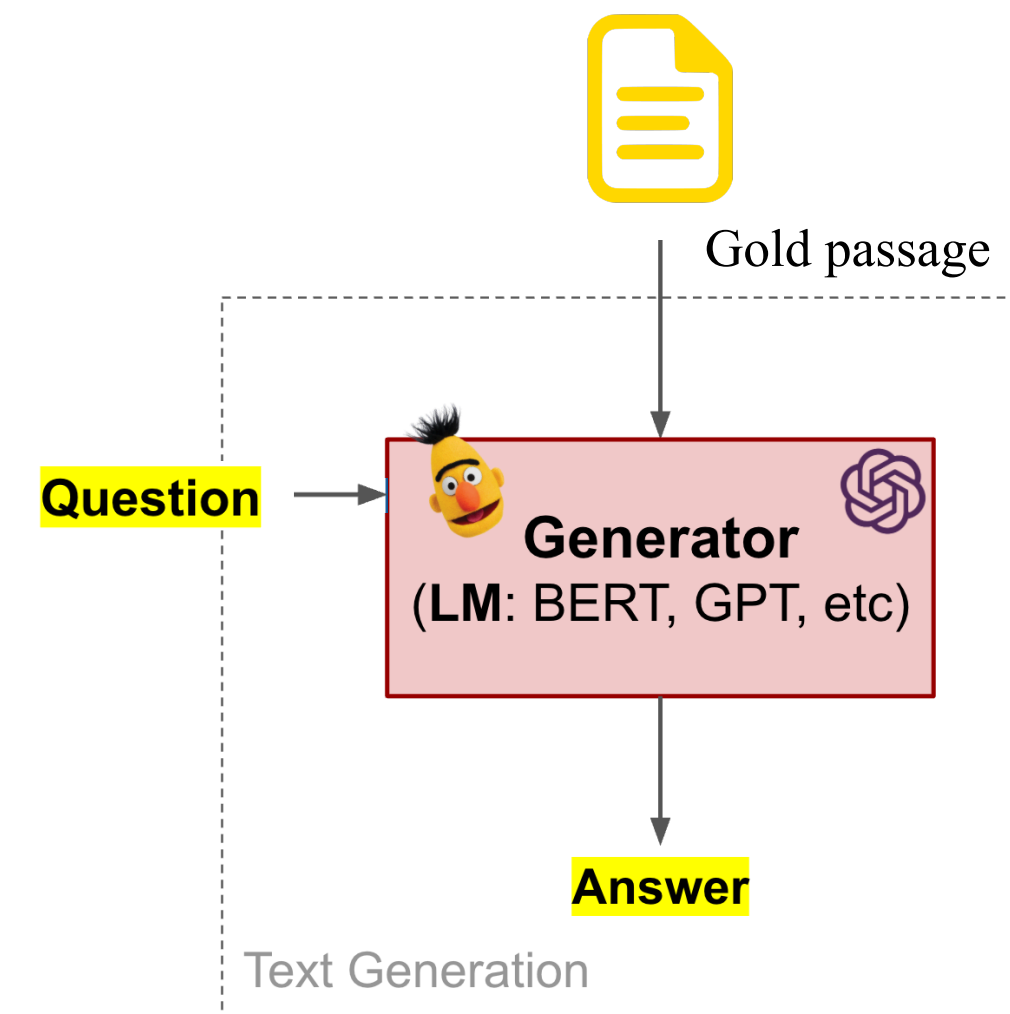REWRITES: was rage against the machine nominated for any award → Works best with sparse retrievers

What is the bottleneck for open-book models?

Retrieval is the bottleneck

14.8 F1 points below H.P.

4.6 F1 points below H.P.

Results correspond to Fusion-in-Decoder model   H.P. - Human Performance

Takeaways

- TopiOCQA: a new challenging dataset for open-domain conversational QA

  ▪ Has free-form answers

  ▪ Incorporates topic switching

  ▪ Has (QA turn, gold passage) and (QA turn, gold answers) pairs

Takeaways

- TopiOCQA: a new challenging dataset for open-domain conversational QA

  ▪ Has free-form answers

  ▪ Incorporates topic switching

  ▪ Has (QA turn, gold passage) and (QA turn, gold answers) pairs

▪ Modeling conversational QA

  ▪ Current models significantly fall short of human performance

  ▪ Open-book models perform better than closed-book models

  ▪ Retrieval is the bottleneck for open-book models

https://mcgill-nlp.github.io/topiocqa