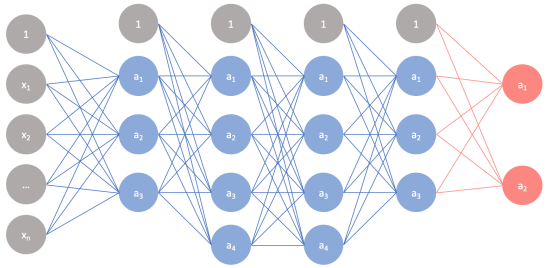
Answering Open-domain, Information-seeking Questions from Text

Sewon Min
University of Washington

KAIST AI, April 27th, 2021

Question answering

Q: Who is the bad guy in harry potter?

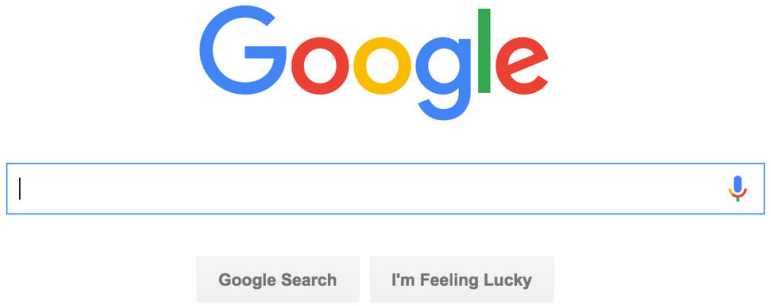


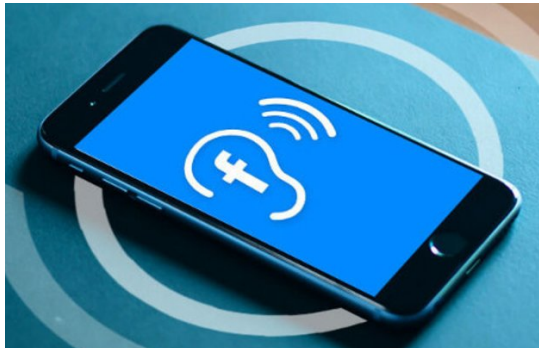
A: Lord Voldemort

A machine learning model

Question answering

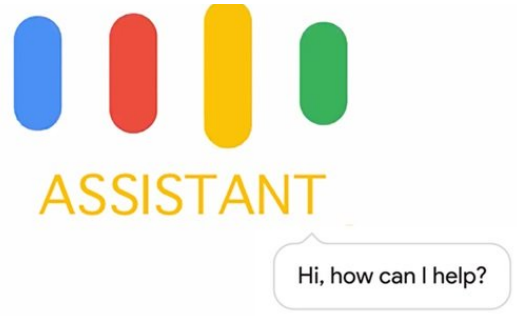
- Key element in a wide range of applications (search engines, personal assistants, …)





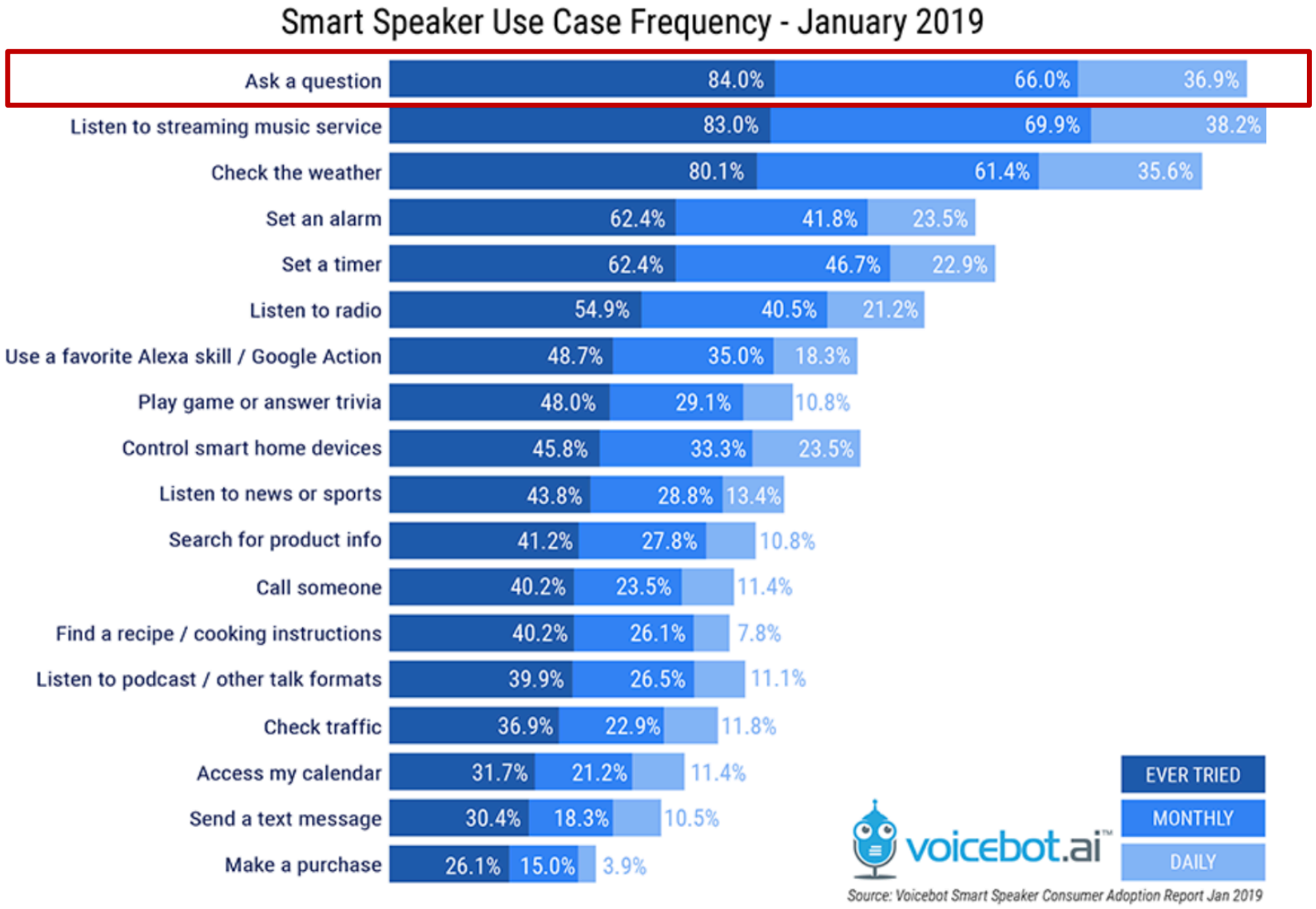
Hey Siri

amazon alexa



Question answering

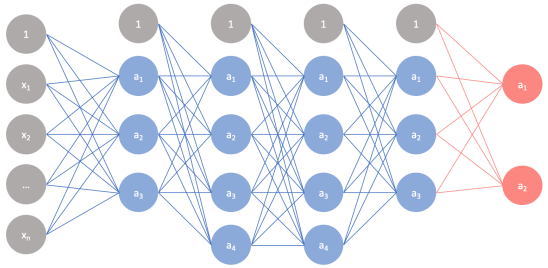
- Key element in a wide range of applications (search engines, personal assistants, …)



Question answering

- Key element in a wide range of applications (search engines, personal assistants, …)
- Evaluate the progress for natural language understanding

Q: Who is the bad guy in harry potter?



A: Lord Voldemort

Harry Potter

From Wikipedia, the free encyclopedia

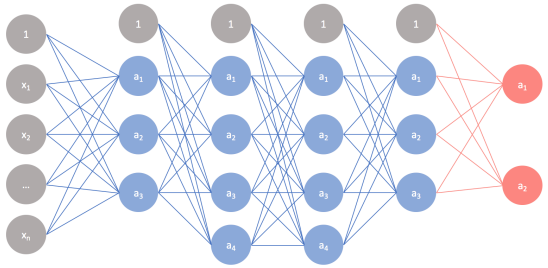
Harry Potter is a series of seven fantasy novels (…) The novels chronicle the lives of a young wizard, Harry Potter (…) The main story arc concerns his struggle against a villain **Lord Voldemort**, a dark wizard who intends to become immortal, overthrow the wizard governing body (…)

Question answering

- Key element in a wide range of applications (search engines, personal assistants, …)
- Evaluate the progress for natural language understanding

Q: Who is the bad guy in harry potter?

Entity (novel series)



A: Lord Voldemort

Harry Potter

From Wikipedia, the free encyclopedia

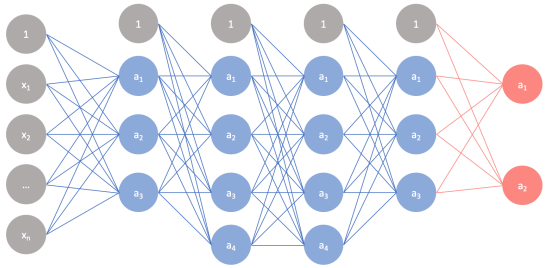
Harry Potter is a series of seven fantasy novels (…) The novels chronicle the lives of a young wizard, Harry Potter (…) The main story arc concerns his struggle against a villain **Lord Voldemort**, a dark wizard who intends to become immortal, ove *Entity (person)* erning body (…)

Question answering

- Key element in a wide range of applications (search engines, personal assistants, …)
- Evaluate the progress for natural language understanding

Q: Who is the bad guy in harry potter?

Entity (novel series)



A: Lord Voldemort

Harry Potter

From Wikipedia, the free encyclopedia

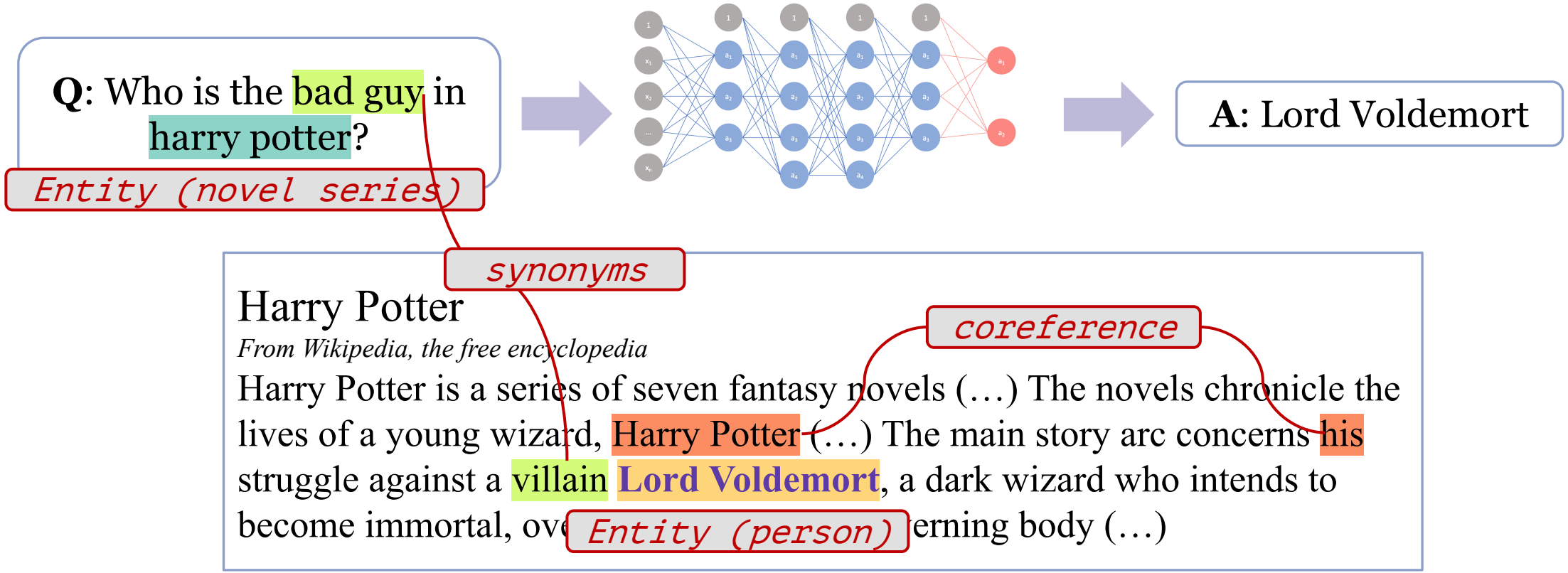
Harry Potter is a series of seven fantasy novels (…) The novels chronicle the lives of a young wizard, Harry Potter (…) The main story arc concerns his struggle against a villain **Lord Voldemort**, a dark wizard who intends to become immortal, ove erning body (…)

coreference

Entity (person)

Question answering

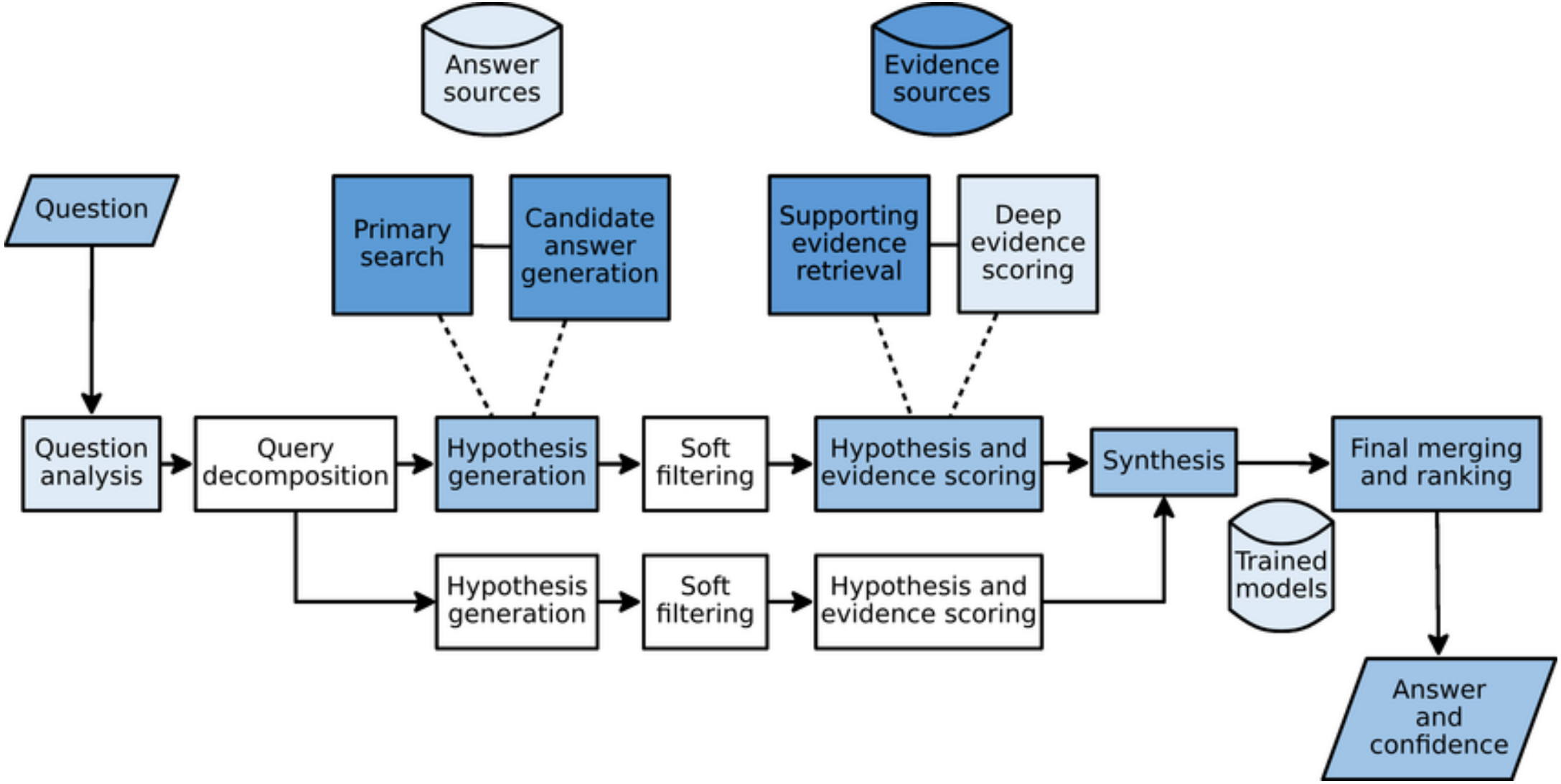
- Key element in a wide range of applications (search engines, personal assistants, …)
- Evaluate the progress for natural language understanding



IBM Watson

Two “Jeopardy!” champions, Ken Jennings, left, and Brad Rutter, competed against a computer named Watson, which proved adept at buzzing in quickly. Carol Kaelson/Jeopardy Productions Inc., via Associated

IBM Watson


Answer sources
Evidence sources
Question
Primary search
Candidate answer generation
Supporting evidence retrieval
Deep evidence scoring
Question analysis
Query decomposition
Hypothesis generation
Soft filtering
Hypothesis and evidence scoring
Synthesis
Final merging and ranking
Hypothesis generation
Soft filtering
Hypothesis and evidence scoring
Trained models
Answer and confidence

Question Answering in deep learning era

Today: Open-domain Textual QA

Q: Who is the bad guy in harry potter?

Harry Potter
From Wikipedia, the free encyclopedia
Harry Potter is a series of seven fantasy novels … Harry's struggle against **Lord Voldemort**, a dark

Today: Open-domain Textual QA

Q: Who is the bad guy in harry potter?

Harry Potter
From Wikipedia, the free encyclopedia
Harry Potter is a series of seven fantasy novels … Harry's struggle against **Lord Voldemort**, a dark

- Models find the answer from a large collection of documents (e.g. Wikipedia)

Today: Open-domain Textual QA

Q: Who is the bad guy in harry potter?

Harry Potter
From Wikipedia, the free encyclopedia
Harry Potter is a series of seven fantasy novels … Harry's struggle against **Lord Voldemort**, a dark

- Models find the answer from a large collection of documents (e.g. Wikipedia)
- We care about a factoid question in natural language

Today: Open-domain Textual QA

Q: Who is the bad guy in harry potter?

Harry Potter
From Wikipedia, the free encyclopedia
Harry Potter is a series of seven fantasy novels … Harry's struggle against **Lord Voldemort**, a dark

- Models find the answer from a large collection of documents (e.g. Wikipedia)
- We care about a factoid question in natural language **that naturally occurs**

a.k.a. information-seeking

Source of Questions

Naturally-occurring Qs

Who is the bad guy in harry potter?

Less studied before 2019

Source of Questions

Naturally-occurring Qs

Who is the bad guy in harry potter?

Less studied before 2019

Annotated Qs

Harry Potter is a series of seven fantasy novels … The main story arc concerns his struggle against a villain Lord Voldemort, a dark wizard …

Q: Who is the bad guy in harry potter?
A: Lord Voldemort

Source of Questions

Naturally-occurring Qs

Who is the bad guy in harry potter?

Less studied before 2019

Annotated Qs

Harry Potter is a series of seven fantasy novels … The main story arc concerns his struggle against a villain Lord Voldemort, a dark wizard …

Q: Who is the bad guy in harry potter?
A: Lord Voldemort

Dominant approach in modern large-scale data collection
e.g. SQuAD (Rajpurkar et al 2017)

Source of Questions

Naturally-occurring Qs

Who is the bad guy in harry potter?

Less studied before 2019

Annotated Qs

Harry Potter is a series of seven fantasy novels … The main story arc concerns his struggle against a villain Lord Voldemort, a dark wizard …

Q: Who is the bad guy in harry potter?
A: Lord Voldemort

Dominant approach in modern large-scale data collection
e.g. SQuAD (Rajpurkar et al 2017)

Great advances have made!
(Seo et al 2017, Yu et al 2018, Devlin et al 2019)

Source of Questions

Naturally-occurring Qs

Who is the bad guy in harry potter?

Less studied before 2019

Annotated Qs

Harry Potter is a series of seven fantasy novels … The main story arc concerns his struggle against a villain Lord Voldemort, a dark wizard …

Q: Who is the bad guy in harry potter?
A: Lord Voldemort

Dominant approach in modern large-scale data collection
)

Great advances have made!
(Seo et al 2017, Yu et al 2018, Devlin et al 2019)

Difference in question distributions

Problem 1 – little lexical cues

Harry Potter is a series of seven fantasy novels … The main story arc concerns his struggle against a villain Lord Voldemort, a dark wizard …

→ "Who is the villain in Harry Potter?"

Easy

- Lexical cues in annotated Qs allow *bypassing* natural language understanding

(Jia and Liang 2017, Kaushik & Lipton, 2018, Min et al 2019, Gardner et al 2019)

Problem 1 – little lexical cues

Harry Potter is a series of seven fantasy novels … The main story arc concerns his struggle against a villain Lord Voldemort, a dark wizard …

"Who is the villain in Harry Potter?"

Easy

"Who is the bad guy in Harry Potter?"

- Lexical cues in annotated Qs allow *bypassing* natural language understanding

(Jia and Liang 2017, Kaushik & Lipton, 2018, Min et al 2019, Gardner et al 2019)

Problem 1 – little lexical cues

Harry Potter is a series of seven fantasy novels … The main story arc concerns his struggle against a villain Lord Voldemort, a dark wizard …

"Who is the villain in Harry Potter?"

Easy

"Who is the bad guy in Harry Potter?"

Hard

- Lexical cues in annotated Qs allow *bypassing* natural language understanding
- State-of-the-art models poorly perform on naturally-occurring Qs with little lexical cues (accuracy of 26%* as of early 2019, even with BERT!)

(Jia and Liang 2017, Kaushik & Lipton, 2018, Min et al 2019, Gardner et al 2019)

* on NQ, from Lee et al 2019

Problem 2 – no single definite answer

“When did Harry Potter and the Sorcerer’s stone movie come out?”

Problem 2 – no single definite answer

“When did Harry Potter and the Sorcerer’s stone movie come out?”

> Harry Potter and the Philosopher's Stone … had its world premiere at the Odeon Leicester Square in London on 4 November 2001 (…) released to cinemas in the United Kingdom and United States on 16 November 2001.

Problem 2 – no single definite answer

"When did Harry Potter and the Sorcerer's stone movie come out?"

> Harry Potter and the Philosopher's Stone ... had its world premiere at the Odeon Leicester Square in London on 4 November 2001 (...) released to cinemas in the United Kingdom and United States on 16 November 2001.

- Annotated Qs are based on a strong assumption that every question has a single clear answer

Problem 2 – no single definite answer

"When did Harry Potter and the Sorcerer's stone movie come out?"

> Harry Potter and the Philosopher's Stone ... had its world premiere at the Odeon Leicester Square in London on 4 November 2001 (...) released to cinemas in the United Kingdom and United States on 16 November 2001.

- Annotated Qs are based on a strong assumption that every question has a single clear answer
- Often not the case in real questions

Problem 2 – no single definite answer

"When did Harry Potter and the Sorcerer's stone movie come out?"

> Harry Potter and the Philosopher's Stone ... had its world premiere at the Odeon Leicester Square in London on 4 November 2001 (...) released to cinemas in the United Kingdom and United States on 16 November 2001.

- Annotated Qs are based on a strong assumption that every question has a single clear answer
- Often not the case in real questions
  - *An intrinsic problem: question writers do not have full background knowledge*

Overview

I. Beyond questions with lexical cues

Advanced SOTA: 26 → 83

GraphRetriever (Min et al. 2020, Li et al. EMNLP 2020)
DPR (Karpukhin et al. EMNLP 2020)
EfficientQA competition (Min et al. 2021, Submitted to PMLR 2021)

II. Beyond unambiguous questions

Identified problem for the first time

AmbigQA (Min et al. EMNLP 2020)
Joint Passage Retrieval (Min et al. 2021)

III. Future directions

Overview

I. Beyond questions with lexical cues

Advanced SOTA: 26 → 83

GraphRetriever (Min et al. 2020, Li et al. EMNLP 2020)

DPR (Karpukhin et al. EMNLP 2020)

EfficientQA competition (Min et al. 2021, Submitted to PMLR 2021)

II. Beyond unambiguous questions

Identified problem for the first time

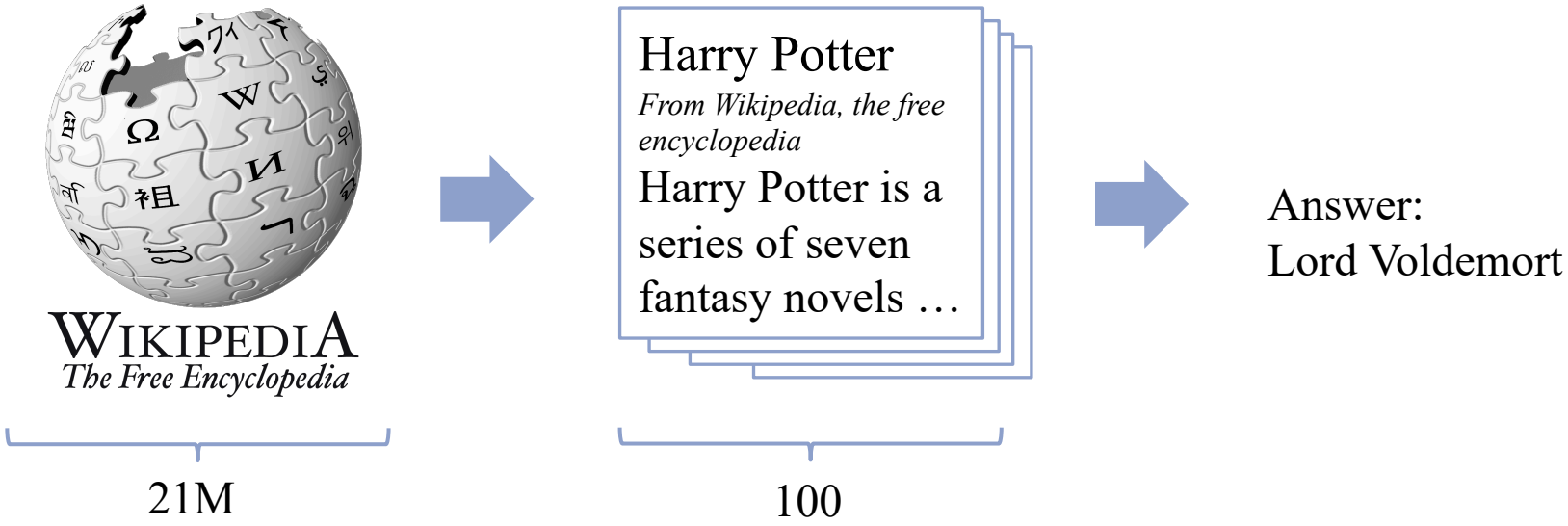
AmbigQA (Min et al. EMNLP 2020)

Joint Passage Retrieval (Min et al. 2021)

III. Future directions

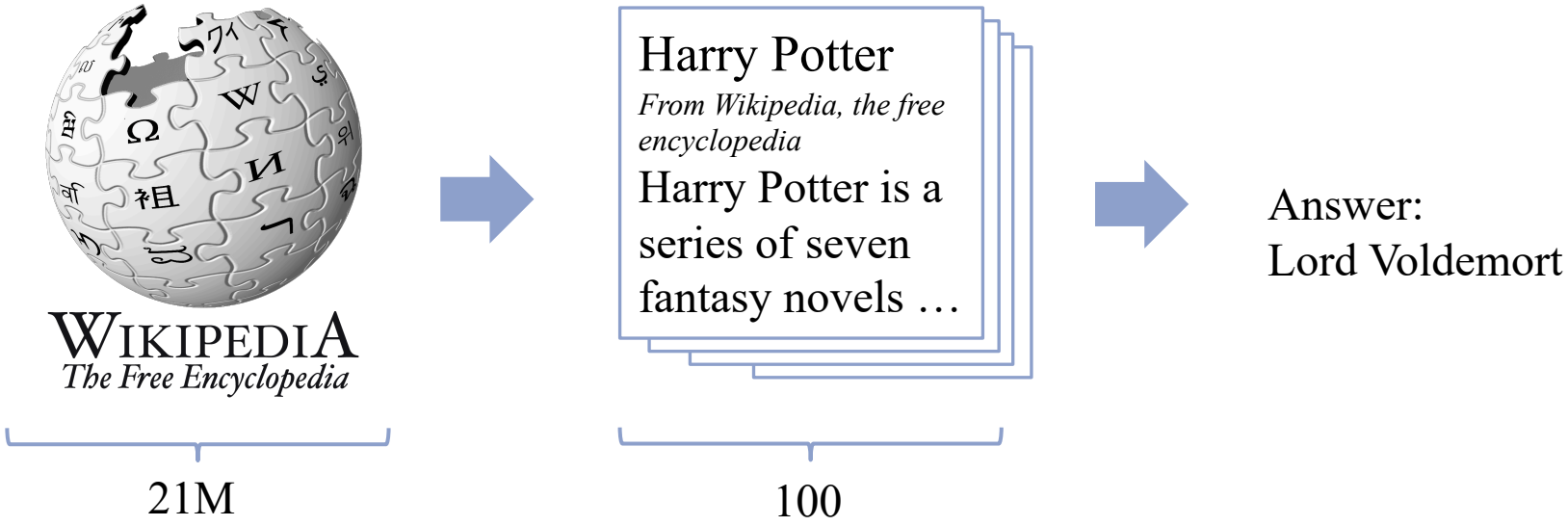
Goal

- QA models for answering naturally-occurring questions with little lexical cues



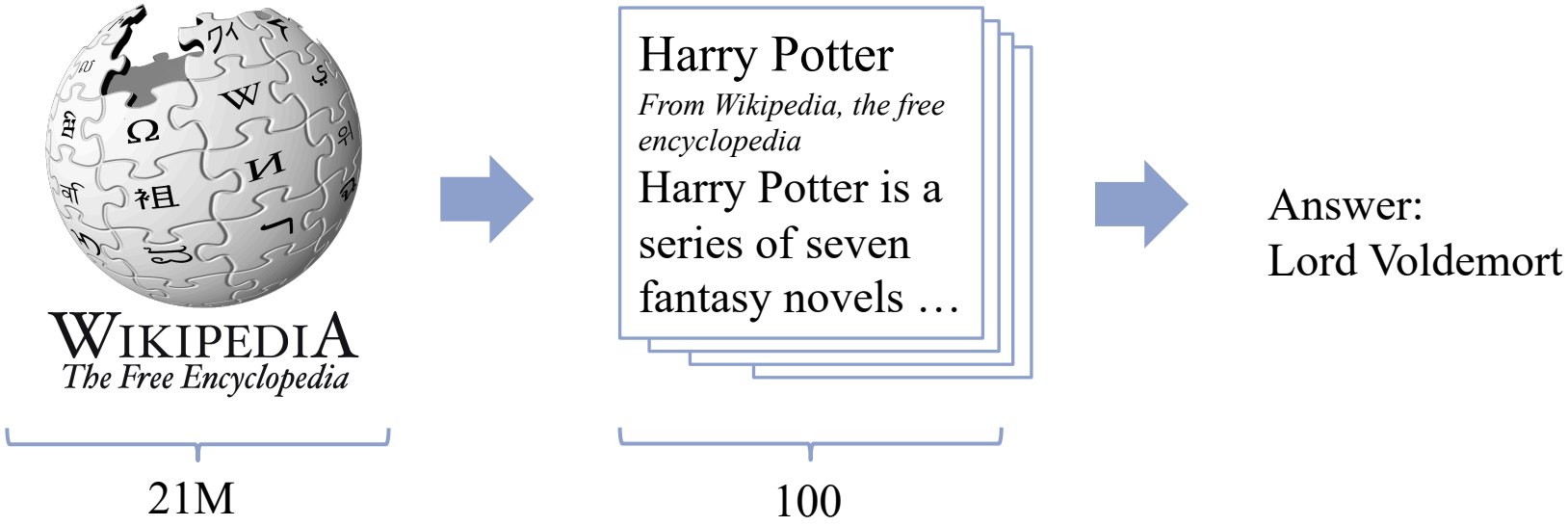
Goal

- QA models for answering naturally-occurring questions with little lexical cues, *specifically focusing on retrieval*



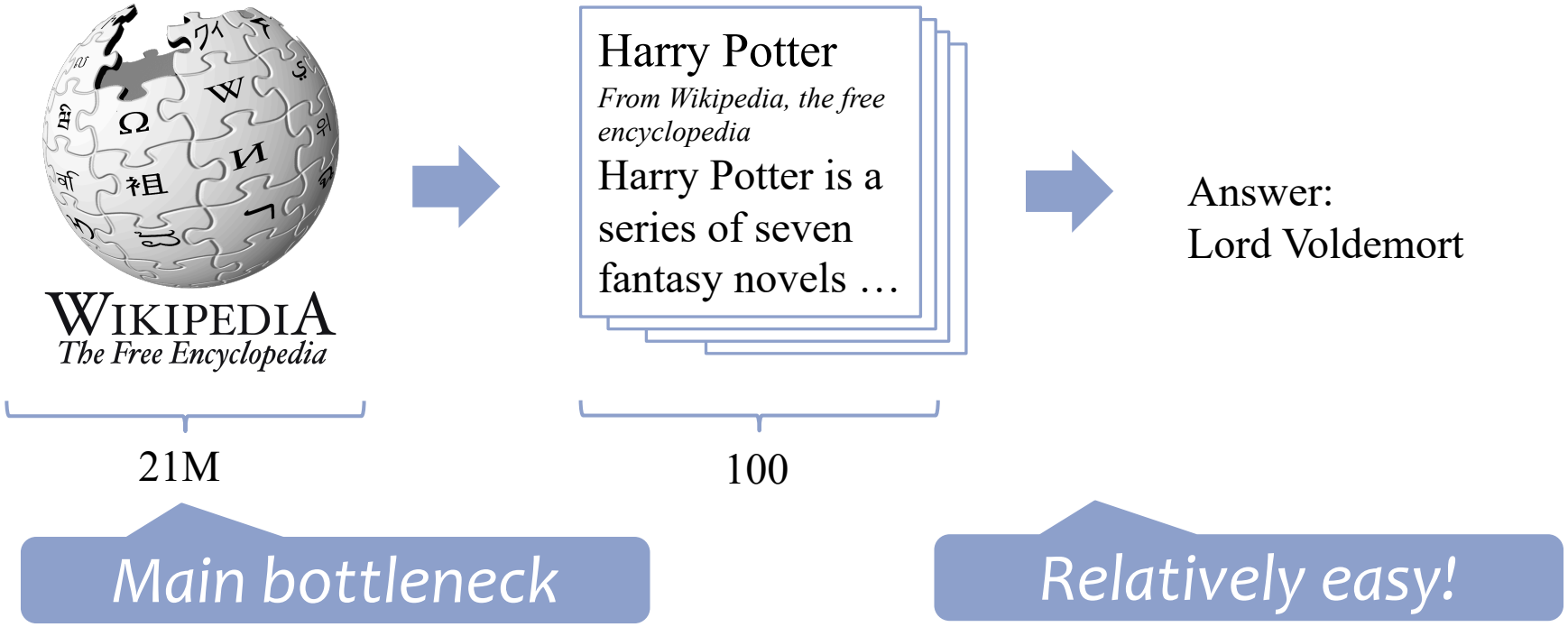
Goal

- QA models for answering naturally-occurring questions with little lexical cues, *specifically focusing on retrieval*
- Input: a question



Goal

- QA models for answering naturally-occurring questions with little lexical cues, *specifically focusing on retrieval*
- Input: a question
- Output: a small number of passages, retrieved from Wikipedia



Goal

- QA models for answering naturally-occurring questions with little lexical cues, *specifically focusing on retrieval*
- Input: a question
- Output: a small number of passages, retrieved from Wikipedia

Retrieved passages should contain & support the answer to the question

Efficiently operate over millions / billions of passages

State-of-the-art in early 2019

- Lexical-matching retrieval (TF-IDF/BM25)

Q: Who is the author of harry potter and the sorcerer's stone?

Harry Potter and the Philosopher's Stone is a fantasy novel written by British author J.K. Rowling.

Strong baseline on annotated data (recall of 82% on SQuAD*)

* from Ma et al 2021 “A Replication Study of Dense Passage Retriever”

State-of-the-art in early 2019 - limitations

- Lexical-matching retrieval (TF-IDF/BM25)

Q: Who is the bad guy in harry potter?

The main story arc concerns Harry's struggle against a villain **Lord Voldemort**.

State-of-the-art in early 2019 - limitations

- Lexical-matching retrieval (TF-IDF/BM25)

Q: Who is the bad guy in harry potter?

The main story arc concerns Harry's struggle against a villain **Lord Voldemort**.

Q: Who sang More than a Feeling by Boston?

"More Than a Feeling" is a song by the American rock band Boston. Written by Tom Scholz, …

Boston (album) is a debut studio album by American rock band Boston. … with singer **Brad Delp**.

State-of-the-art in early 2019 - limitations

- Lexical-matching retrieval (TF-IDF/BM25)

Q: Who is the bad guy in harry potter?

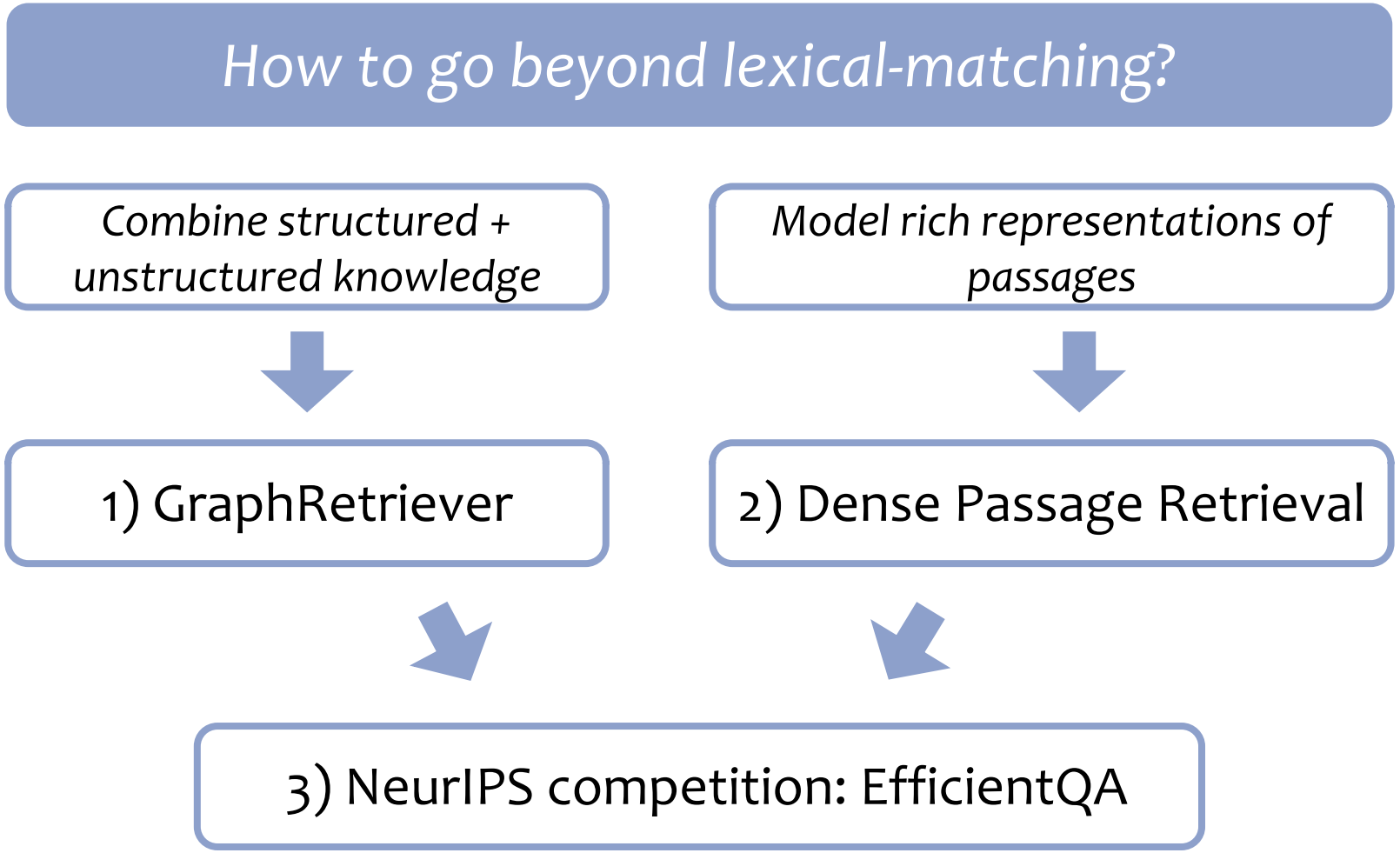
The main story arc concerns Harry's struggle against a villain **Lord Voldemort**.

Q: Who sang More than a Feeling by Boston?

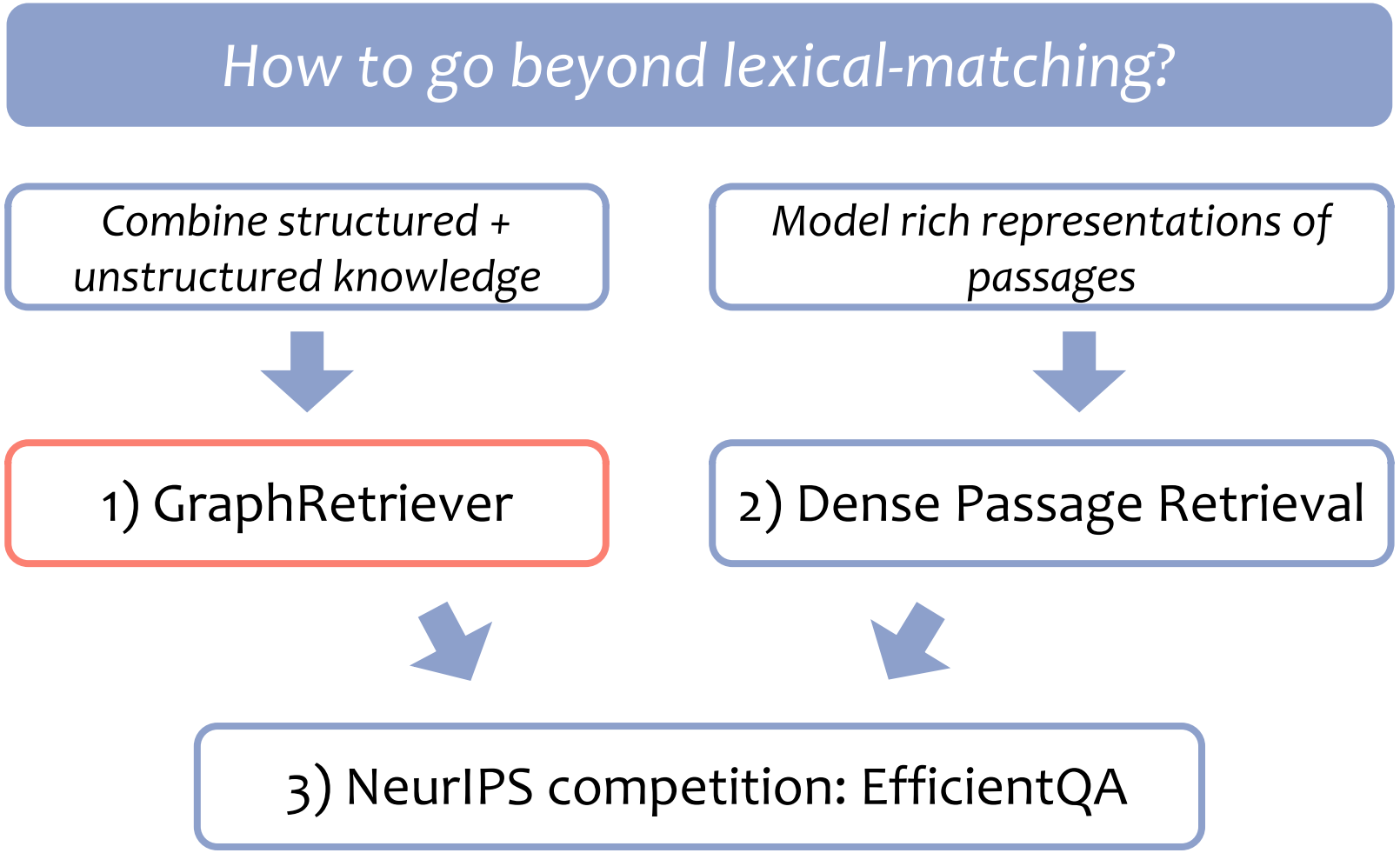
"More Than a Feeling" is a song by the American rock band Boston. Written by Tom Scholz, …

Boston (album) is a debut studio album by American rock band Boston. … with singer **Brad Delp**.

Our Work

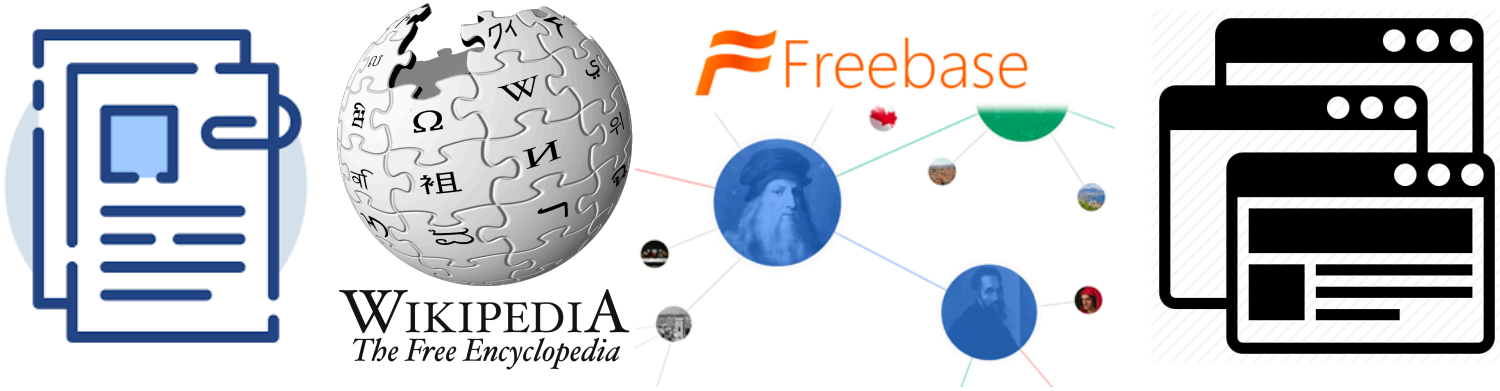

How to go beyond lexical-matching?
Combine structured + unstructured knowledge
Model rich representations of passages
1) GraphRetriever
2) Dense Passage Retrieval
3) NeurIPS competition: EfficientQA

Our Work


How to go beyond lexical-matching?
Combine structured + unstructured knowledge
Model rich representations of passages
1) GraphRetriever
2) Dense Passage Retrieval
3) NeurIPS competition: EfficientQA

GraphRetriever

Q: Who sang More than a Feeling by Boston?

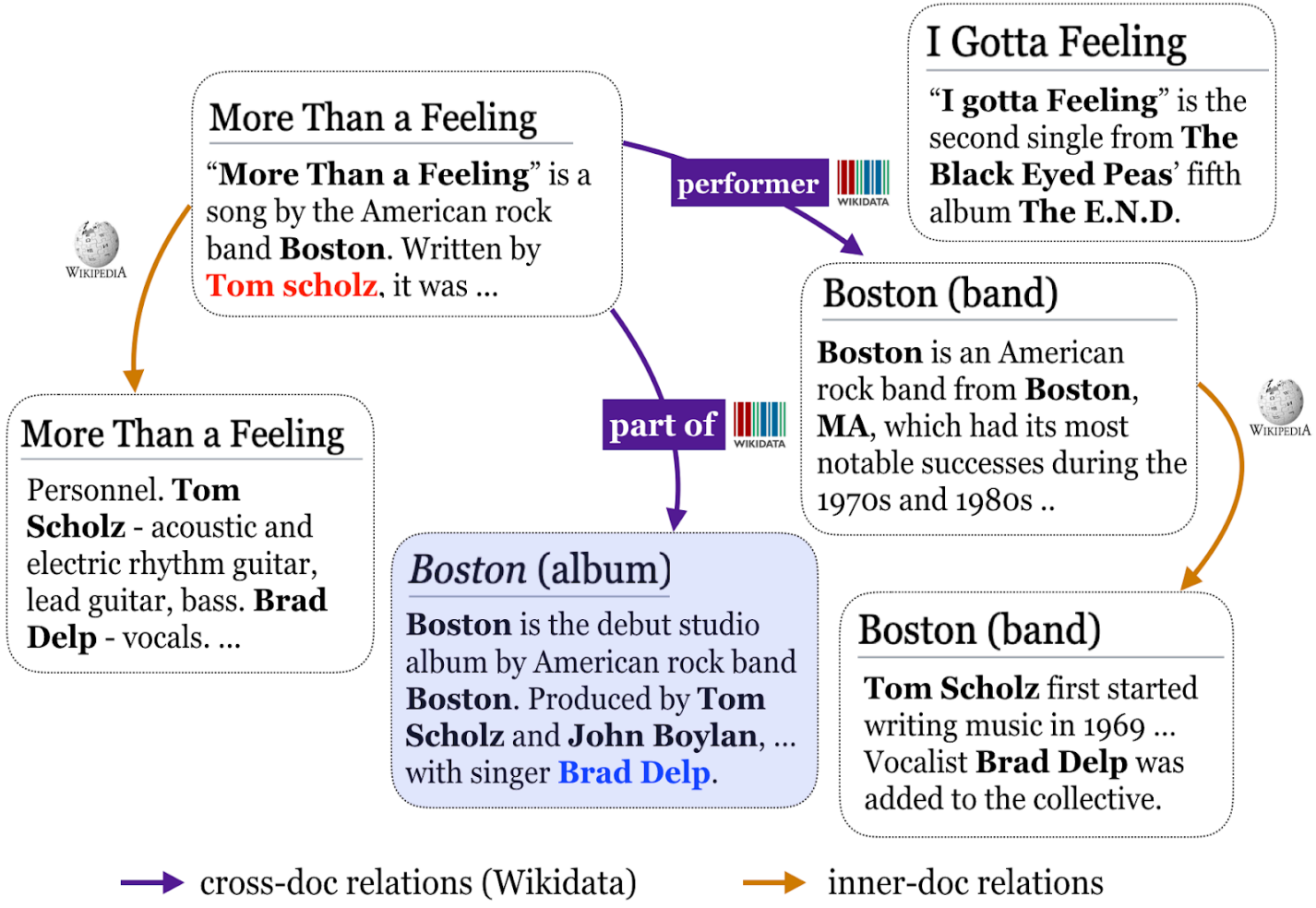


Combine structured + unstructured knowledge

Min, Chen, Zettlemoyer, Hajishirzi, "Knowledge Guided Text Retrieval and Reading for Open Domain Question Answering"

GraphRetriever

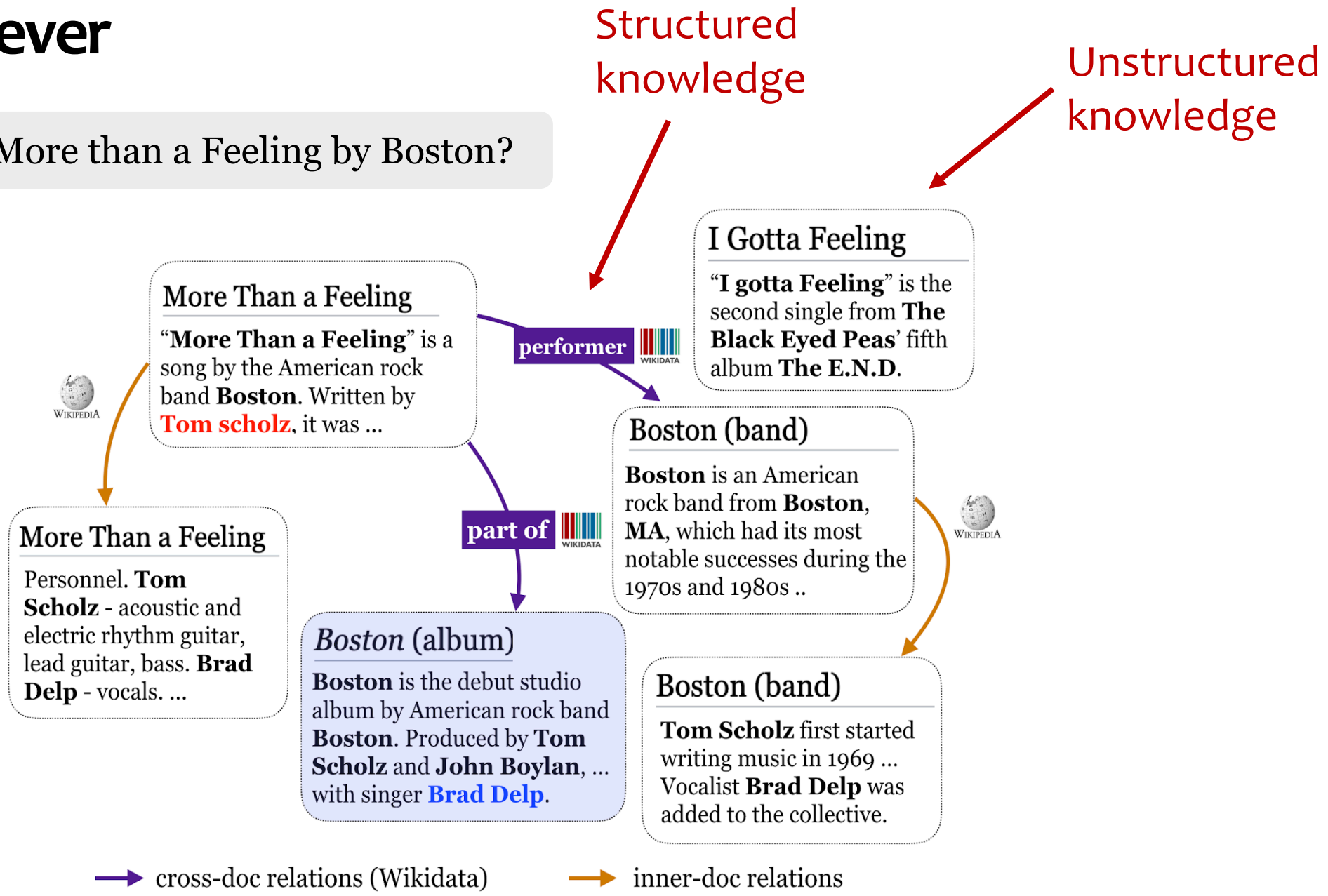
Q: Who sang More than a Feeling by Boston?



Min, Chen, Zettlemoyer, Hajishirzi, "Knowledge Guided Text Retrieval and Reading for Open Domain Question Answering"

GraphRetriever

Q: Who sang More than a Feeling by Boston?



Min, Chen, Zettlemoyer, Hajishirzi, "Knowledge Guided Text Retrieval and Reading for Open Domain Question Answering"

GraphRetriever

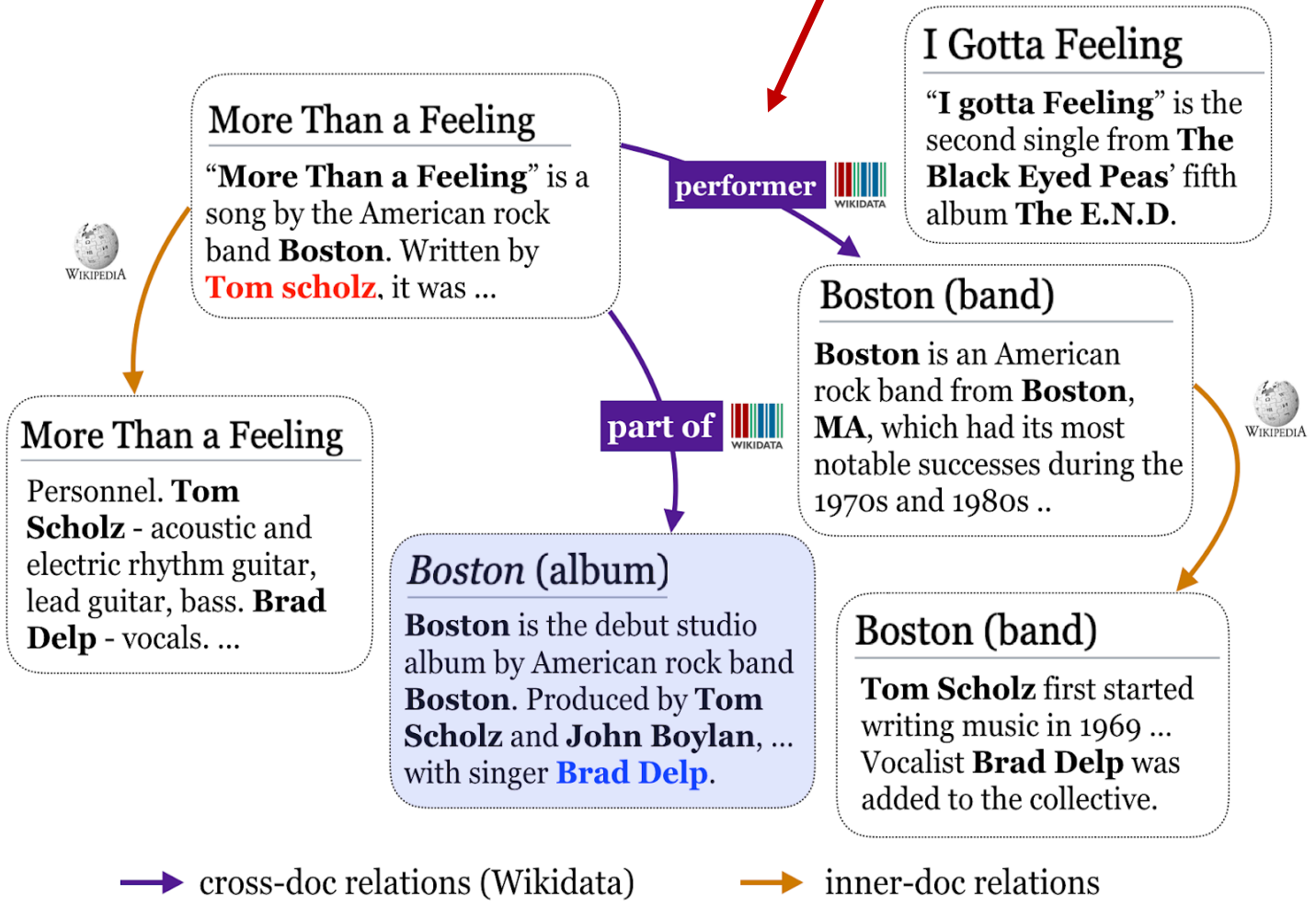
Q: Who sang More than a Feeling by Boston?

Structured knowledge

Direct support

Unstructured knowledge

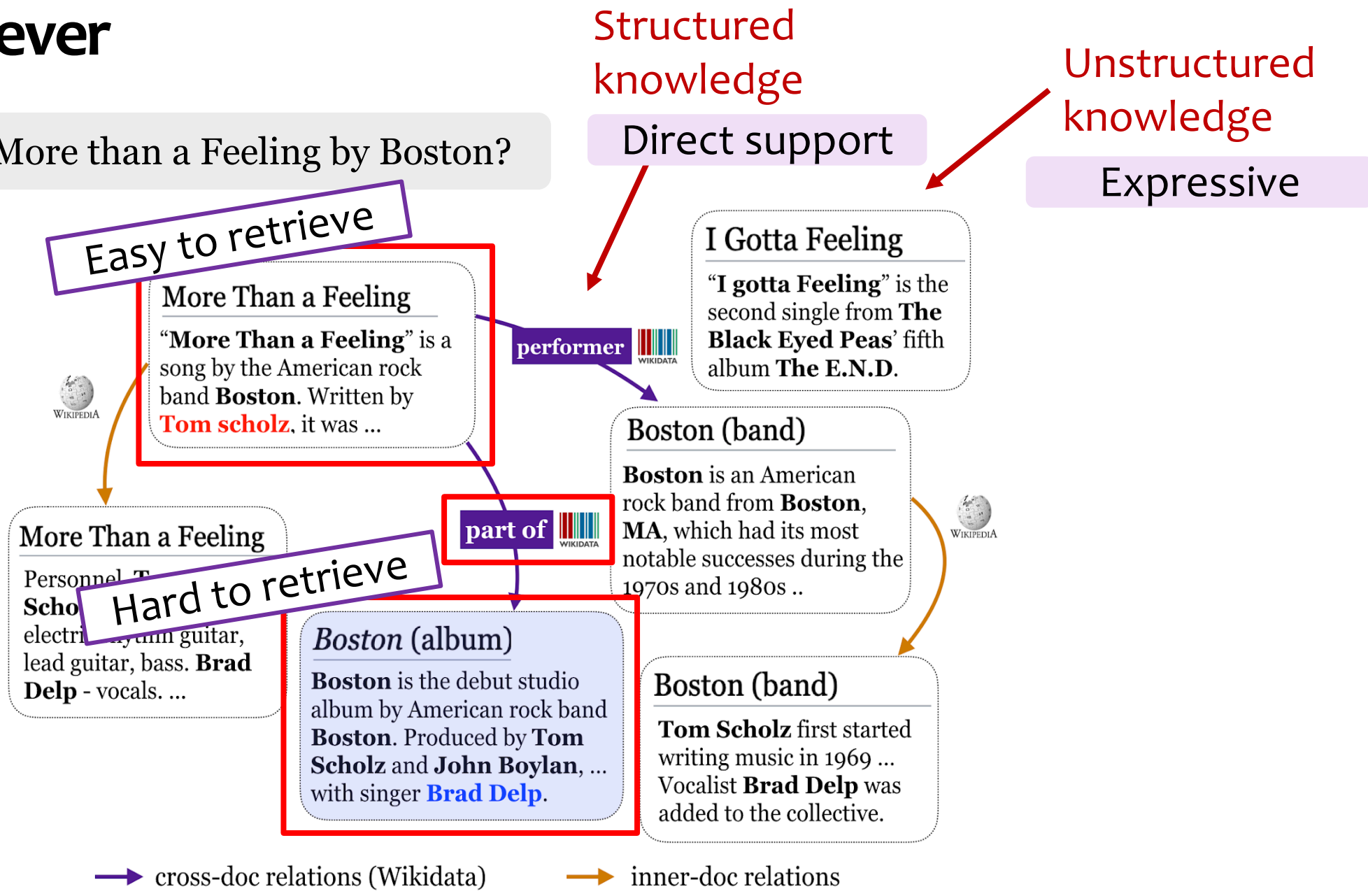
Expressive



Min, Chen, Zettlemoyer, Hajishirzi, "Knowledge Guided Text Retrieval and Reading for Open Domain Question Answering"

GraphRetriever

Q: Who sang More than a Feeling by Boston?



Min, Chen, Zettlemoyer, Hajishirzi, "Knowledge Guided Text Retrieval and Reading for Open Domain Question Answering"

Model – (1) GraphRetriever

Q: Who sang More than a Feeling by Boston?

Init: Seed passages $\mathcal{P}^{(0)}$

Recursive: $\mathcal{P}^{(m-1)} \rightarrow \mathcal{P}^{(m)}$

Min, Chen, Zettlemoyer, Hajishirzi, "Knowledge Guided Text Retrieval and Reading for Open Domain Question Answering"

Model – (1) GraphRetriever

Q: Who sang More than a Feeling by Boston?

Init: Seed passages $\mathcal{P}^{(0)}$

Recursive: $\mathcal{P}^{(m-1)} \rightarrow \mathcal{P}^{(m)}$

0-th iteration

Entity TF-IDF

More Than a Feeling
"More Than a Feeling" is a song by the American rock band Boston.

Entity TF-IDF

Boston (band)
Boston is an American rock band from Boston. Produced by Tom Scholz, ...

TF-IDF

I gotta Feeling
"I gotta Feeling" is the second single from The Black Eyed Peas' fifth ...

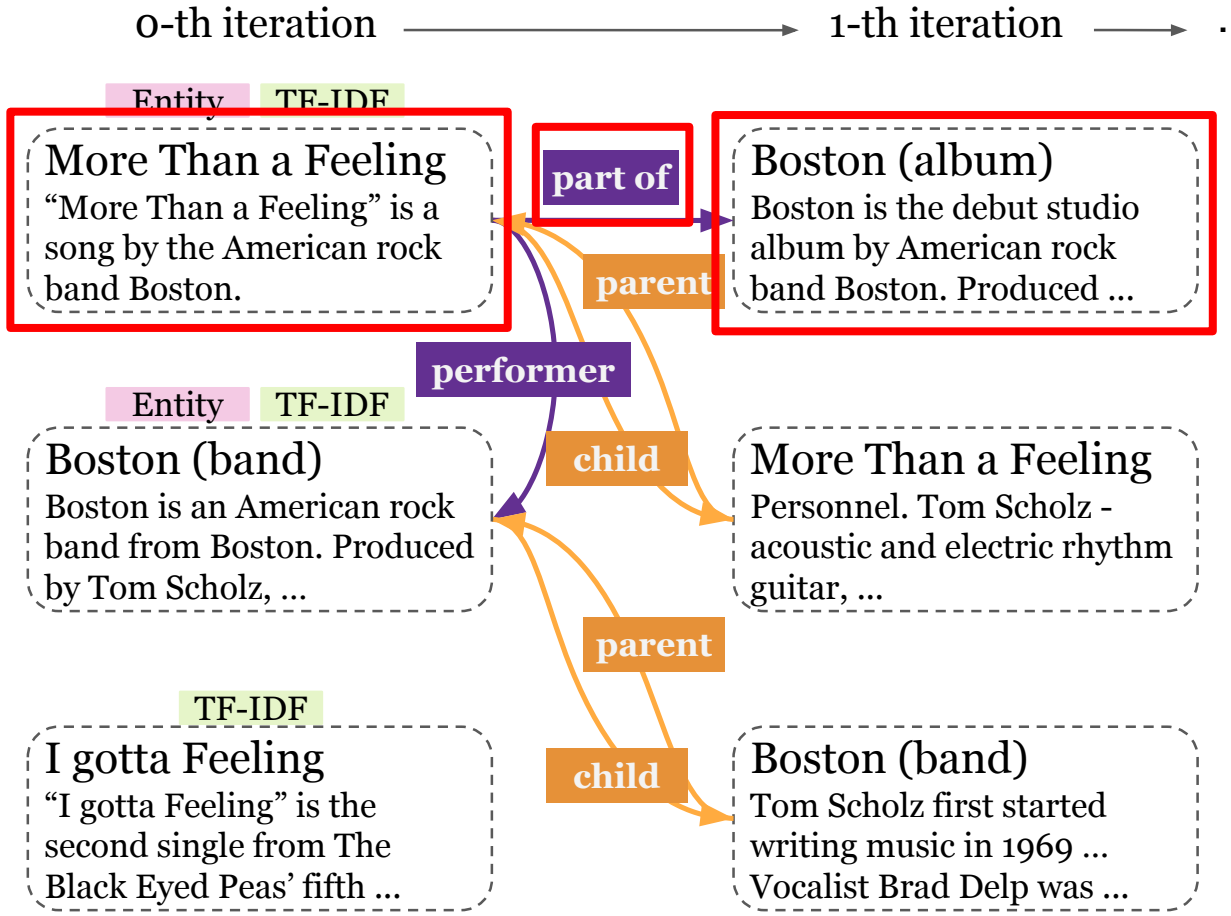
Min, Chen, Zettlemoyer, Hajishirzi, "Knowledge Guided Text Retrieval and Reading for Open Domain Question Answering"

Model – (1) GraphRetriever

Q: Who sang More than a Feeling by Boston?

Init: Seed passages $\mathcal{P}^{(0)}$

Recursive: $\mathcal{P}^{(m-1)} \rightarrow \mathcal{P}^{(m)}$



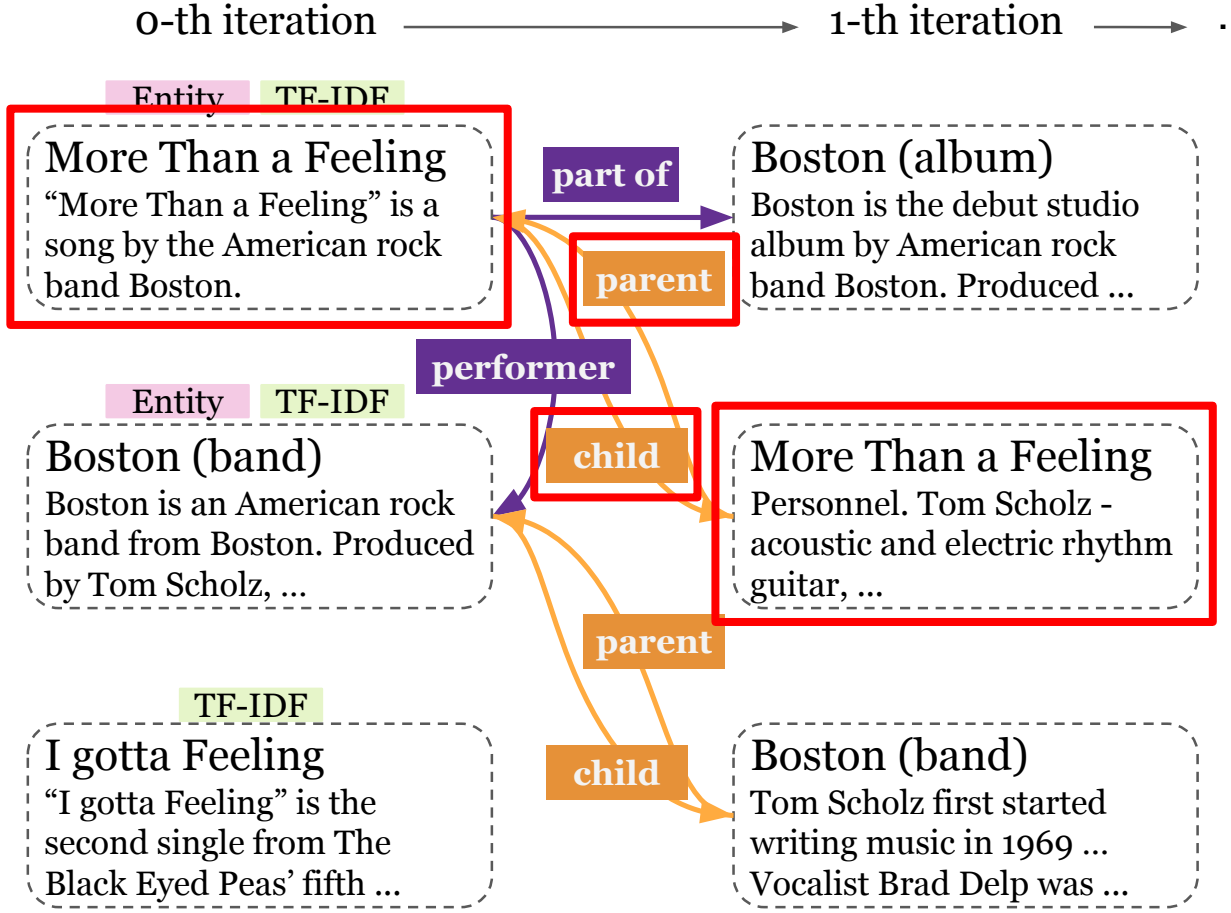
WIKIDATA

- Entities and relations from Wikidata

Min, Chen, Zettlemoyer, Hajishirzi, "Knowledge Guided Text Retrieval and Reading for Open Domain Question Answering"

Model – (1) GraphRetriever

Q: Who sang More than a Feeling by Boston?



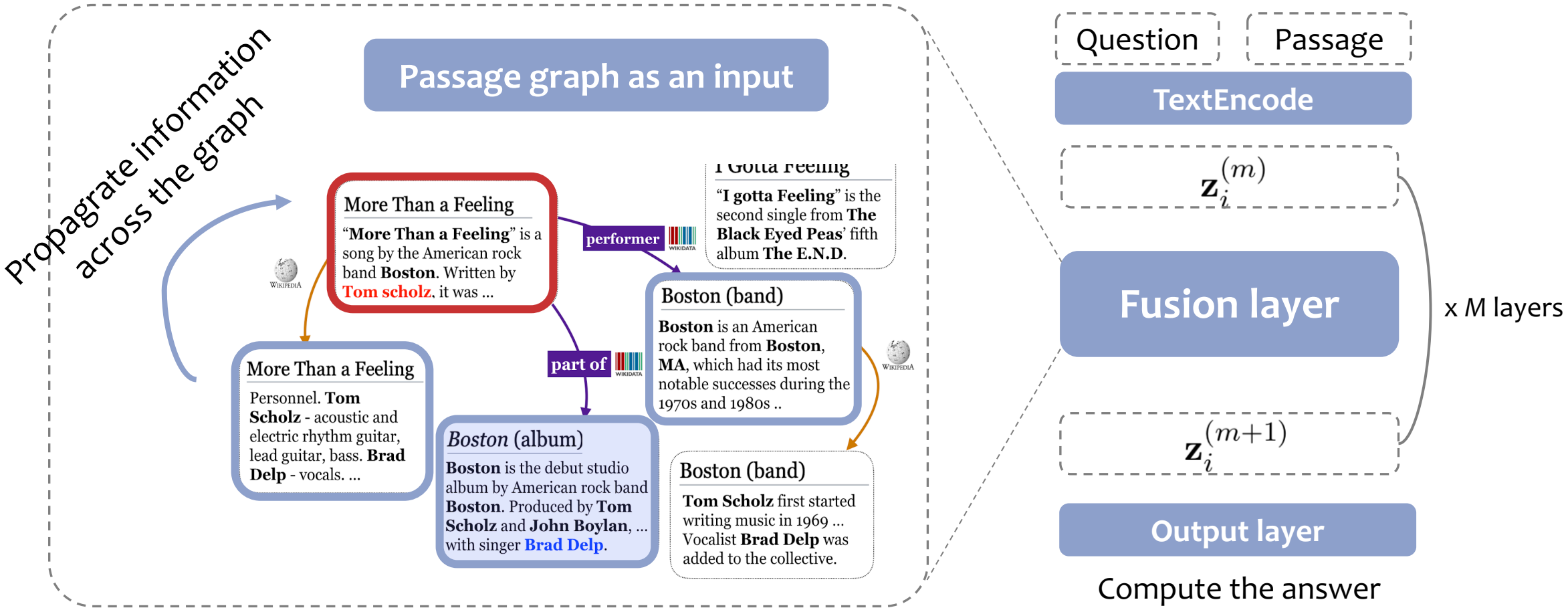
Init: Seed passages $\mathcal{P}^{(0)}$

Recursive: $\mathcal{P}^{(m-1)} \rightarrow \mathcal{P}^{(m)}$

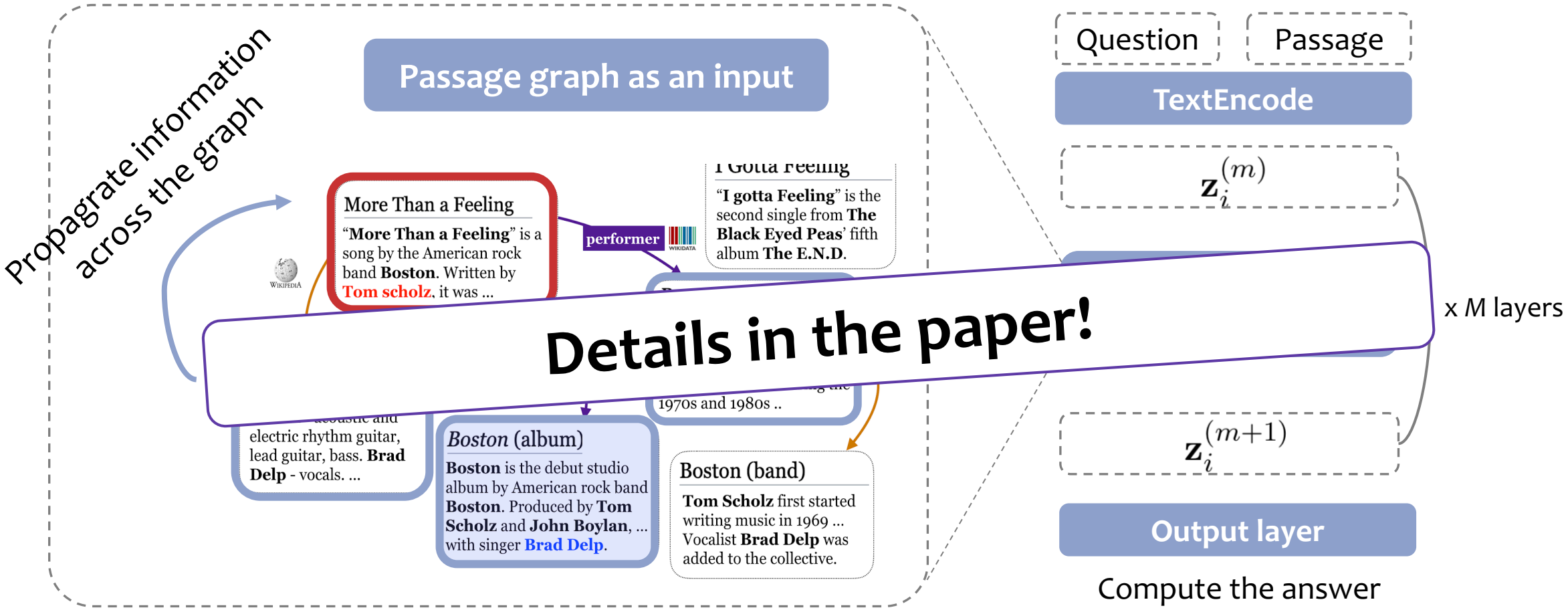
- Entities and relations from Wikidata
- Co-occurrence in the same Wikipedia article

Min, Chen, Zettlemoyer, Hajishirzi, "Knowledge Guided Text Retrieval and Reading for Open Domain Question Answering"

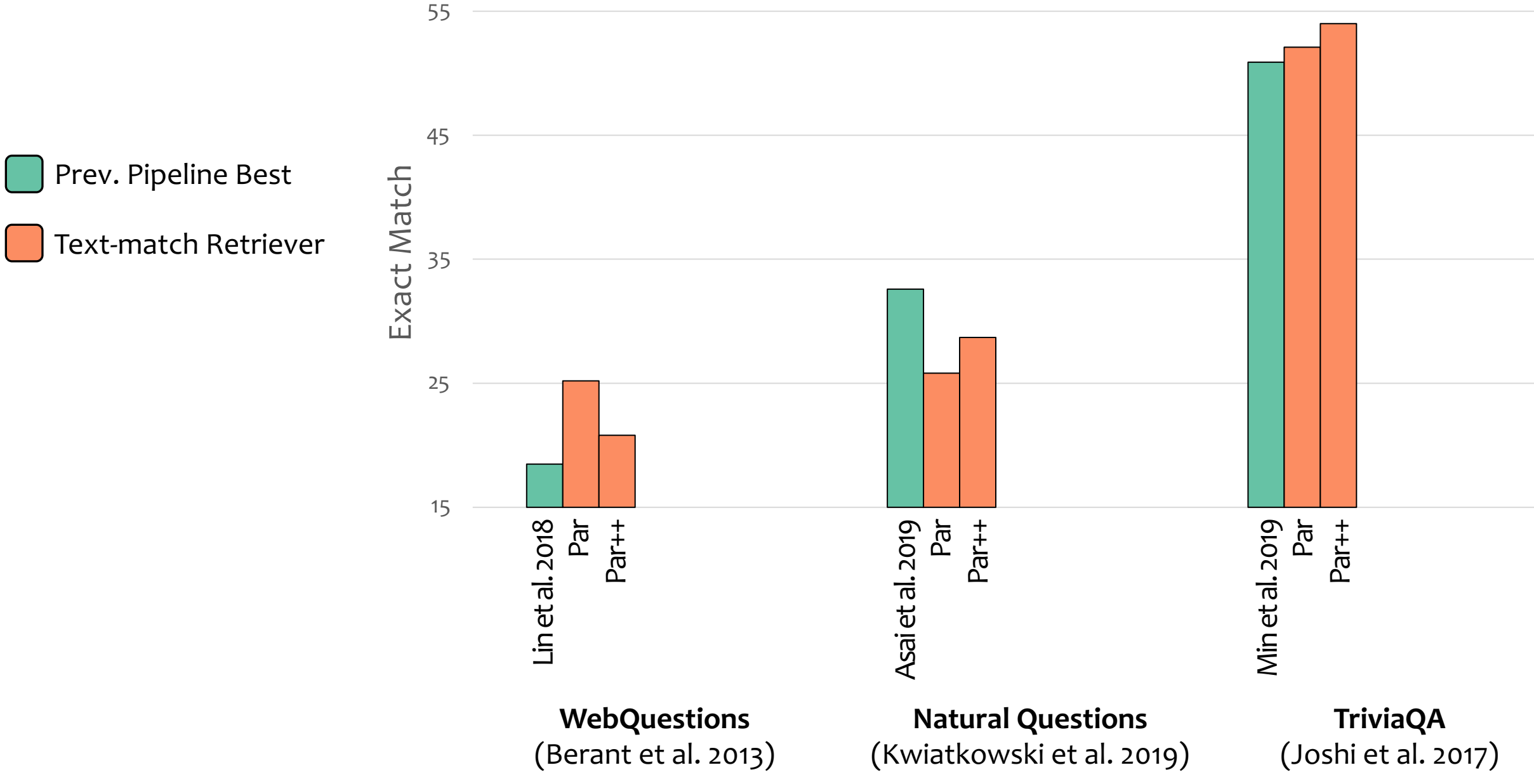
Model – (2) GraphReader



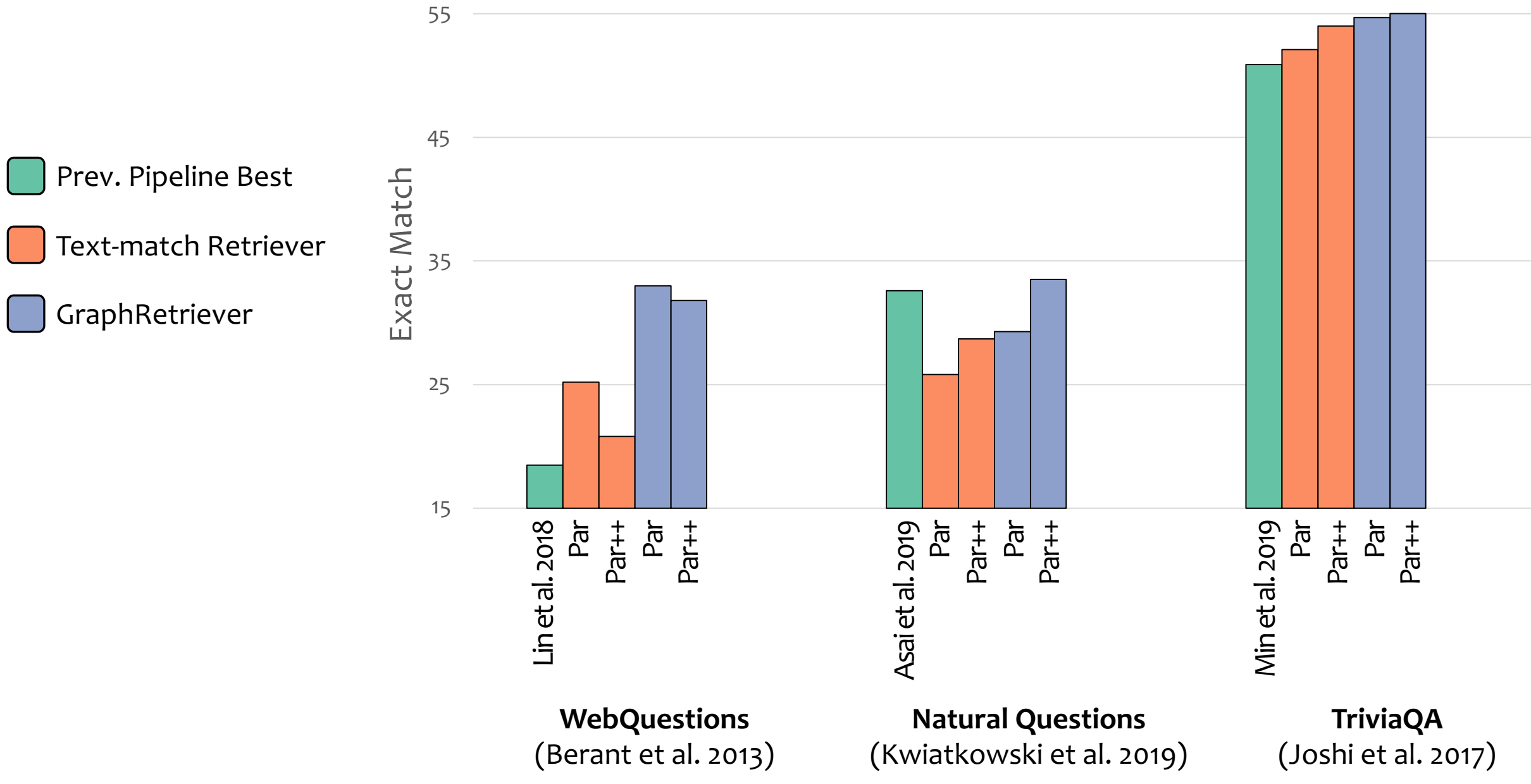
Model – (2) GraphReader



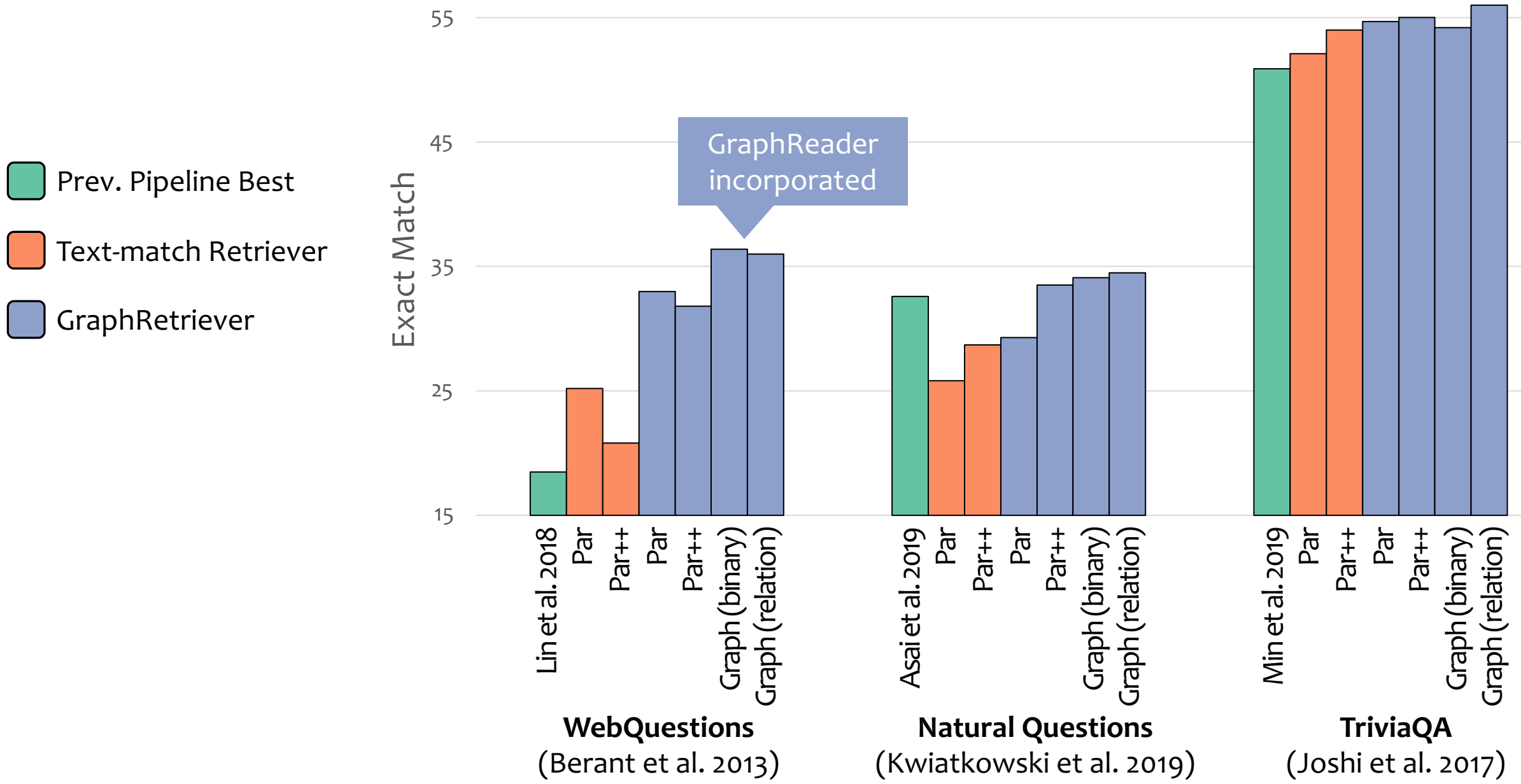
Results - Advance SOTA on three open-domain QA datasets



Results - Advance SOTA on three open-domain QA datasets



Results - Advance SOTA on three open-domain QA datasets



Follow-up work (Li et al EMNLP 2020)

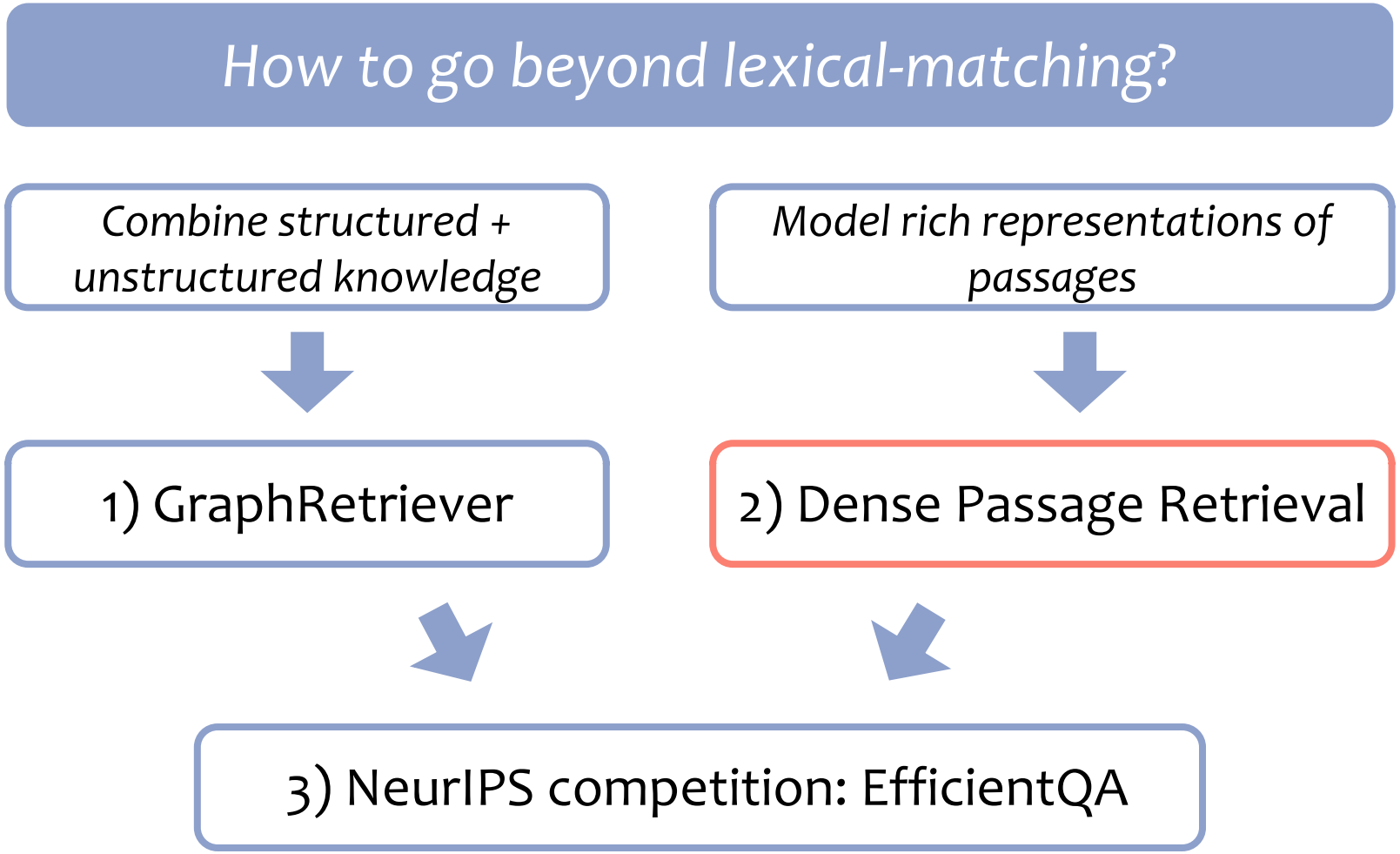
- Better entity linking further increases the performance

	WQ	NQ	TQA
TF-IDF†	20.8	28.7	54.0
TAGME + GRetriever†	31.8	33.5	55.0
ELQ_{Wiki} + GRetriever	37.4	**37.4**	**55.4**
ELQ_{QA} + GRetriever	**37.7**	37.0	54.7

Table 3: QA result (Exact Match) on the test set of WebQuestions (WQ), Natural Questions (NQ) and TriviaQA (TQA). ELQ_{Wiki} represents our model trained on Wikipedia data, while ELQ_{QA} represents our model trained on Wikipedia+$WebQSP_{EL}$ data.
†Result taken from (Min et al., 2019).

Li, **Min**, Iyer, Mehdad, Yih, "Efficient One-Pass End-to-End Entity Linking for Questions". EMNLP 2020

Our Work


How to go beyond lexical-matching?
Combine structured + unstructured knowledge
Model rich representations of passages
1) GraphRetriever
2) Dense Passage Retrieval
3) NeurIPS competition: EfficientQA

Dense Passage Retrieval (DPR)

- Bi-encoder approach (Dense retrieval)
  - $\boldsymbol{q} = Emb(\text{question}) \in \mathbb{R}^h$,
  - $\boldsymbol{x} = Emb(\text{passage}) \in \mathbb{R}^h$
  - $f_{sim}(\text{question}, \text{passage}) = \boldsymbol{q}^T \boldsymbol{x}$

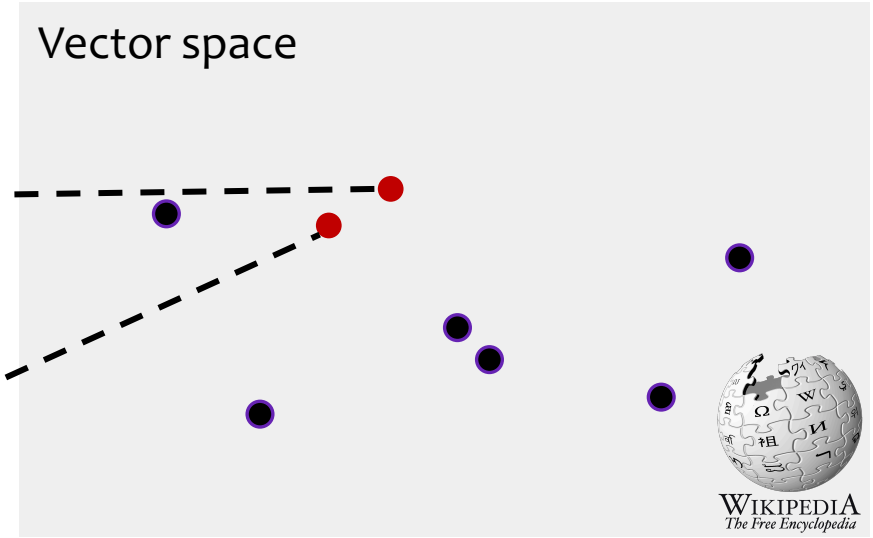
Karpukhin*, Oguz*, **Min**, Lewis, Wu, Edunov, Chen, Yih. “Dense Passage Retrieval for Open-domain Question Answering” EMNLP 2020

Dense Passage Retrieval (DPR)

- Bi-encoder approach (Dense retrieval)
  - $\boldsymbol{q} = Emb(\text{question}) \in \mathbb{R}^h$,
  - $\boldsymbol{x} = Emb(\text{passage}) \in \mathbb{R}^h$
  - $f_{sim}(\text{question}, \text{passage}) = \boldsymbol{q}^T \boldsymbol{x}$

The main story arc concerns Harry's struggle against a villain **Lord Voldemort**, a dark wizard who intends to become immortal, overthrow the wizard governing body

Q: Who is the bad guy in harry potter?



Karpukhin*, Oguz*, **Min**, Lewis, Wu, Edunov, Chen, Yih. "Dense Passage Retrieval for Open-domain Question Answering" EMNLP 2020

Related Work

- Also known as "Siamese" network in the literatures (Bromley et al. 1993, Chopra et al 2005, Yih et al 2011, Huang et al 2013)
- Why was it worse than TF-IDF in the past?

Karpukhin*, Oguz*, **Min**, Lewis, Wu, Edunov, Chen, Yih. "Dense Passage Retrieval for Open-domain Question Answering" EMNLP 2020

Related Work

- Also known as "Siamese" network in the literatures (Bromley et al. 1993, Chopra et al 2005, Yih et al 2011, Huang et al 2013)

- Why was it worse than TF-IDF in the past?

  - It is actually not easy to make bi-encoder model "work"
    - Needs large enough labeled data (e.g. 82M query-doc pairs from user clicks)
    - Not the strongest model for similarity (lack of cross-attention)

Related Work

- Also known as “Siamese” network in the literatures (Bromley et al. 1993, Chopra et al 2005, Yih et al 2011, Huang et al 2013)

- Why was it worse than TF-IDF in the past?

  - It is actually not easy to make bi-encoder model “work”
    - Needs large enough labeled data (e.g. 82M query-doc pairs from user clicks)
    - Not the strongest model for similarity (lack of cross-attention)

  - Efficient index/search on dense vector space is not easy either
    - In memory index, fancy data structure and tricks are required.

Related Work

- Also known as "Siamese" network in the literatures (Bromley et al. 1993, Chopra et al 2005, Yih et al 2011, Huang et al 2013)

- Why was it worse than TF-IDF in the past?

  - It is actually not easy to make bi-encoder model "work"
    - Needs large enough labeled data (e.g. 82M query-d… user clicks)
    - Not the strongest model for similarity (lack of cross-attention)

  - Efficient index/search on dense vector space is not easy either
    - In memory index, fancy data structure and tricks are required.

Pretraining (e.g. BERT) helps!

Advances in tools for Index/Search (e.g. FAISS, ScaNN)

Methodology

- Bi-encode model
  - $\boldsymbol{q} = \text{BERT}(\text{question})[\text{CLS}] \in \mathbb{R}^h$ (h=768)
  - $\boldsymbol{x} = \text{BERT}(\text{passage})[\text{CLS}] \in \mathbb{R}^h$
  - $f_{sim}(\text{question}, \text{passage}) = \boldsymbol{q}^T \boldsymbol{x}$

Karpukhin*, Oguz*, **Min**, Lewis, Wu, Edunov, Chen, Yih. “Dense Passage Retrieval for Open-domain Question Answering” EMNLP 2020

Methodology

- Bi-encode model
  - $\boldsymbol{q} = \text{BERT}(\text{question})[\text{CLS}] \in \mathbb{R}^h$ (h=768)
  - $\boldsymbol{x} = \text{BERT}(\text{passage})[\text{CLS}] \in \mathbb{R}^h$
  - $f_{sim}(\text{question}, \text{passage}) = \boldsymbol{q}^T \boldsymbol{x}$

- Model training
  - Training data: $\{< q, p^+, p^-_1, p^-_2, \dots, p^-_{m-1} >\}$

Karpukhin*, Oguz*, **Min**, Lewis, Wu, Edunov, Chen, Yih. “Dense Passage Retrieval for Open-domain Question Answering” EMNLP 2020

Methodology

- Bi-encode model
  - $\boldsymbol{q} = \text{BERT}(\text{question})[\text{CLS}] \in \mathbb{R}^h$ (h=768)
  - $\boldsymbol{x} = \text{BERT}(\text{passage})[\text{CLS}] \in \mathbb{R}^h$
  - $f_{sim}(\text{question}, \text{passage}) = \boldsymbol{q}^T \boldsymbol{x}$

- Model training
  - Training data: $\{< q, p^+, p^-_1, p^-_2, \dots, p^-_{m-1} >\}$
  - Negative log likelihood loss
    - $L\left(q, p^+, p^-_1, p^-_2, \dots, p^-_{m-1}\right) = -\log \frac{\exp\left(f_{sim}(q,p^+)\right)}{\exp\left(f_{sim}(q,p^+)\right)+\sum_{i=1}^{m-1} \exp\left(f_{sim}\left(q,p^-_i\right)\right)}$

Karpukhin*, Oguz*, **Min**, Lewis, Wu, Edunov, Chen, Yih. “Dense Passage Retrieval for Open-domain Question Answering” EMNLP 2020

Methodology

- Bi-encode model
  - $\boldsymbol{q} = \text{BERT}(\text{question})[\text{CLS}] \in \mathbb{R}^h$ (h=768)
  - $\boldsymbol{x} = \text{BERT}(\text{passage})[\text{CLS}] \in \mathbb{R}^h$
  - $f_{sim}(\text{question}, \text{passage}) = \boldsymbol{q}^T \boldsymbol{x}$

- Model training
  - Training data: $\{< q, p^+, p^-_1, p^-_2, \dots, p^-_{m-1} >\}$
  - Negative log likelihood loss
    - $L\left(q, p^+, p^-_1, p^-_2, \dots, p^-_{m-1}\right) = -\log \frac{\exp\left(f_{sim}(q,p^+)\right)}{\exp\left(f_{sim}(q,p^+)\right)+\sum_{i=1}^{m-1} \exp($

Still doesn't work!

Karpukhin*, Oguz*, **Min**, Lewis, Wu, Edunov, Chen, Yih. "Dense Passage Retrieval for Open-domain Question Answering" EMNLP 2020

Tricks to make DPR work

Trick #1: Hard negatives

- Use negatives that models will be **confused** instead of random negatives
- Choose negatives that have high BM25 scores

Karpukhin*, Oguz*, **Min**, Lewis, Wu, Edunov, Chen, Yih. “Dense Passage Retrieval for Open-domain Question Answering” EMNLP 2020

Tricks to make DPR work

Trick #1: Hard negatives

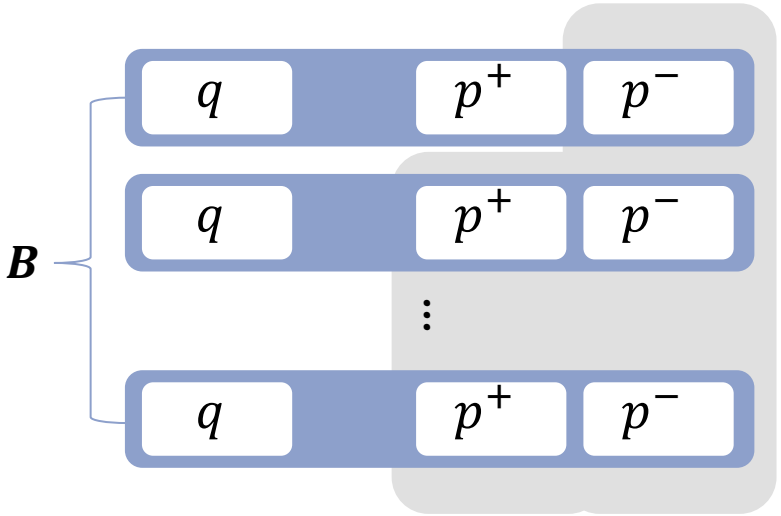
- Use negatives that models will be **confused** instead of random negatives
- Choose negatives that have high BM25 scores
- Much better than the model with random negatives

Karpukhin*, Oguz*, **Min**, Lewis, Wu, Edunov, Chen, Yih. "Dense Passage Retrieval for Open-domain Question Answering" EMNLP 2020

Tricks to make DPR work

Trick #2: In-batch negatives (Yih et al 2011, Henderson et al 2017, Gillick et al 2019)

A batch with $\boldsymbol{B}$ questions (assuming M=2 for simplicity)



Vanilla training

Each q receives $\boldsymbol{M-1}$ signals

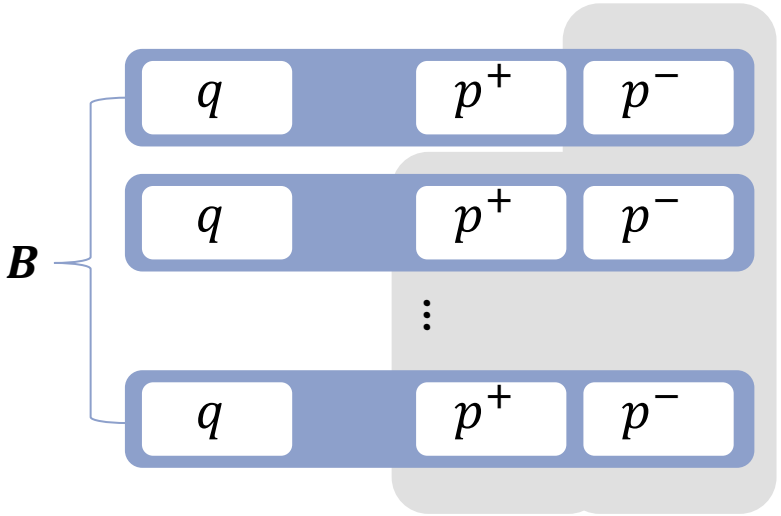
With in-batch negatives

Each q receives $\boldsymbol{BM-1}$ signals

Tricks to make DPR work

Trick #2: In-batch negatives (Yih et al 2011, Henderson et al 2017, Gillick et al 2019)

A batch with $\boldsymbol{B}$ questions (assuming M=2 for simplicity)



Vanilla training

Each q receives $\boldsymbol{M-1}$ signals

With in-batch negatives

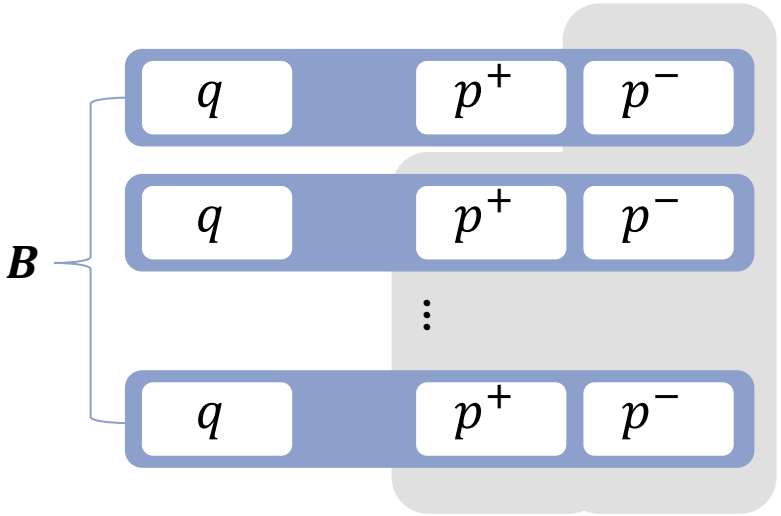
Each q receives $\boldsymbol{BM-1}$ signals

In-batch negatives have worse quality than original negatives

Tricks to make DPR work

Trick #2: In-batch negatives (Yih et al 2011, Henderson et al 2017, Gillick et al 2019)

A batch with $\boldsymbol{B}$ questions (assuming M=2 for simplicity)



Vanilla training

Each q receives $\boldsymbol{M-1}$ signals

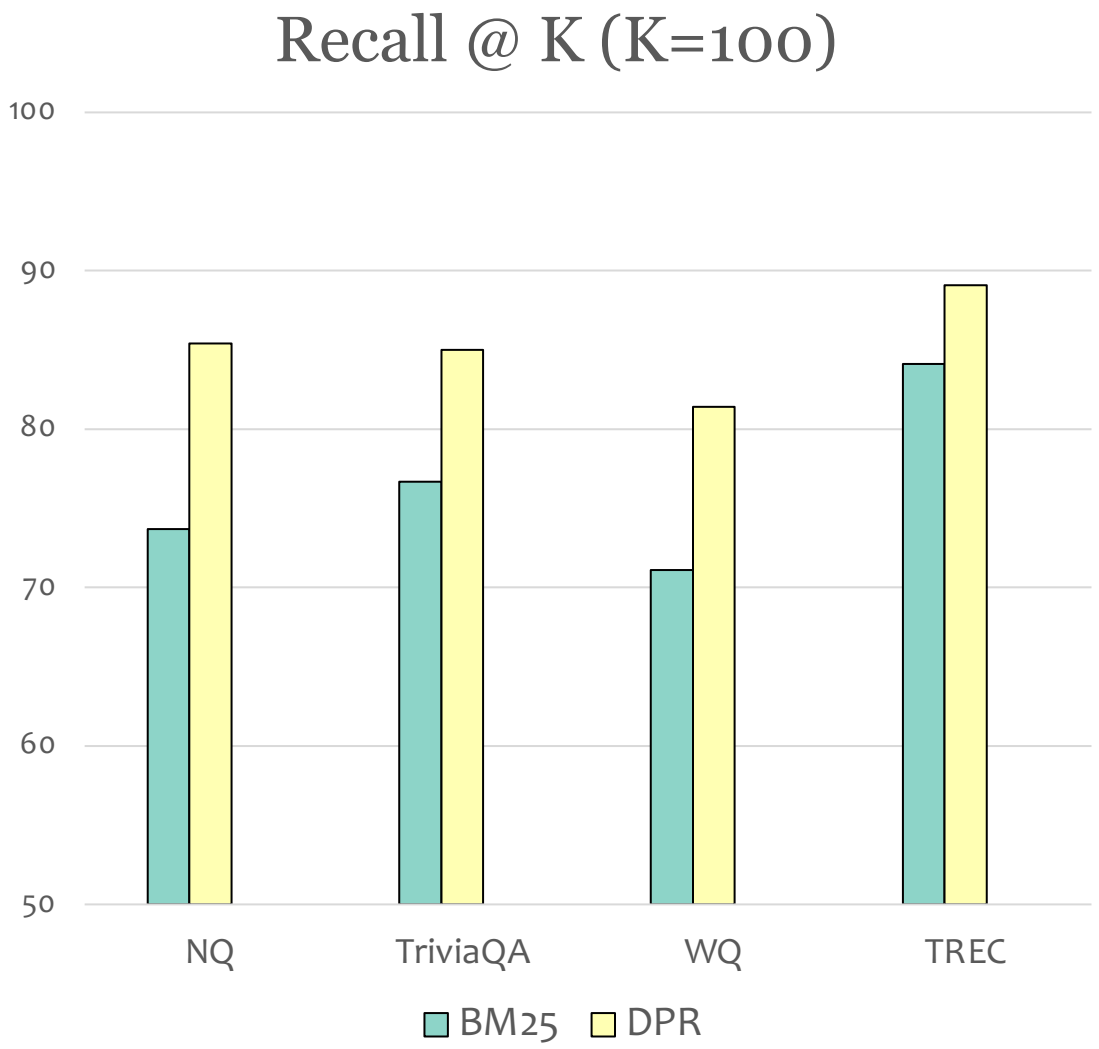
With in-batch negatives

Each q receives $\boldsymbol{BM-1}$ signals

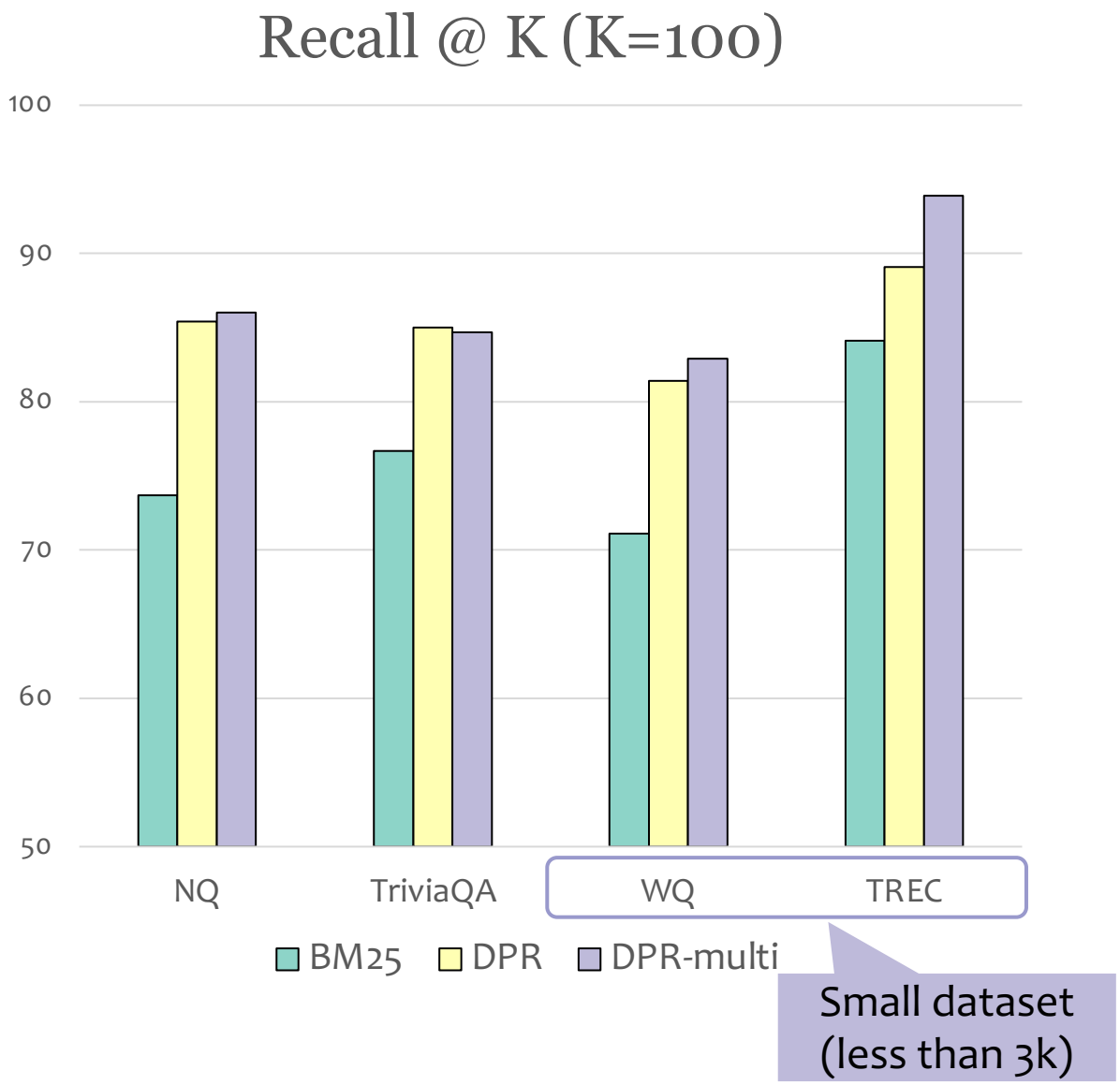
In-batch negatives

More signals == Better signals

Retrieval Results

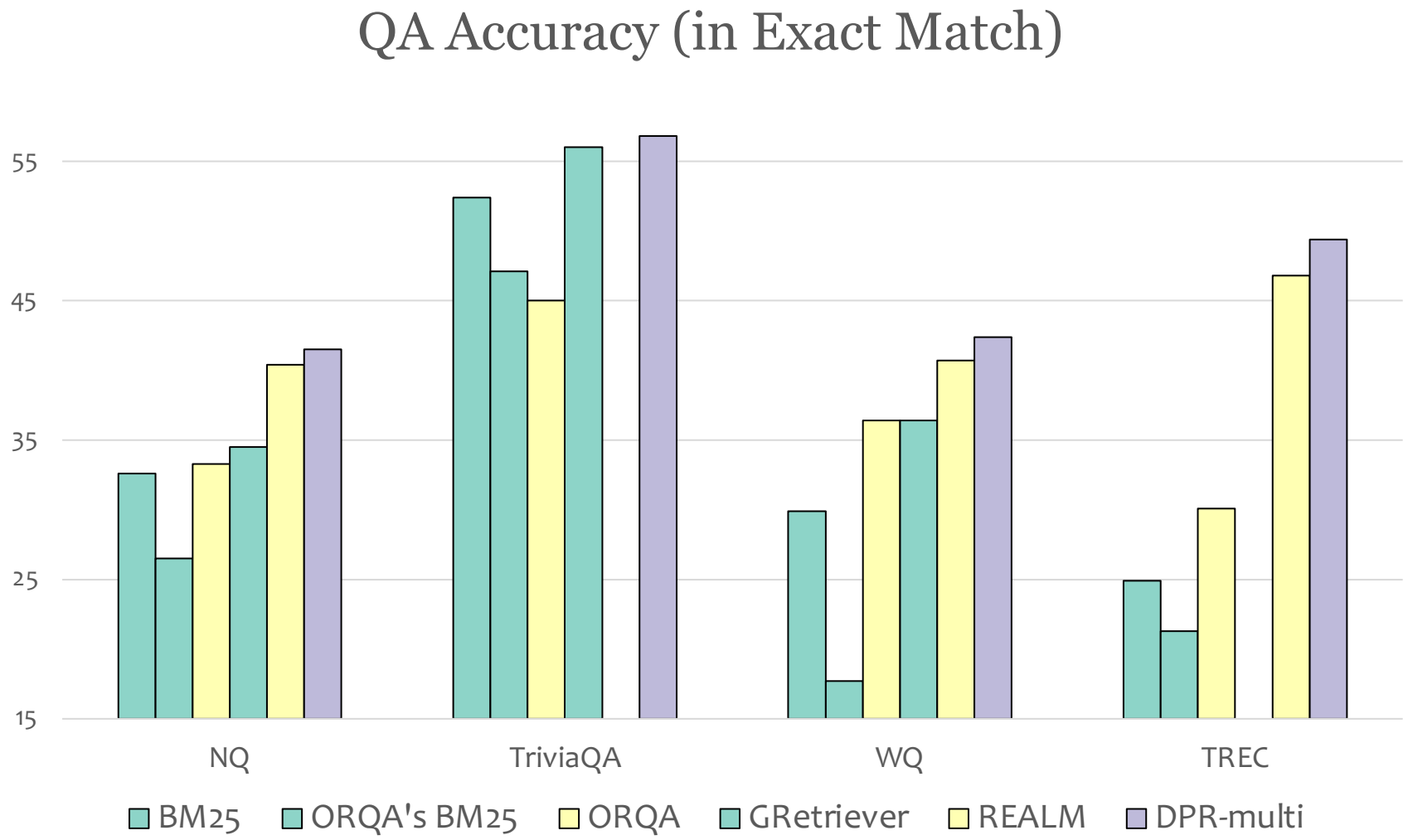


Retrieval Results



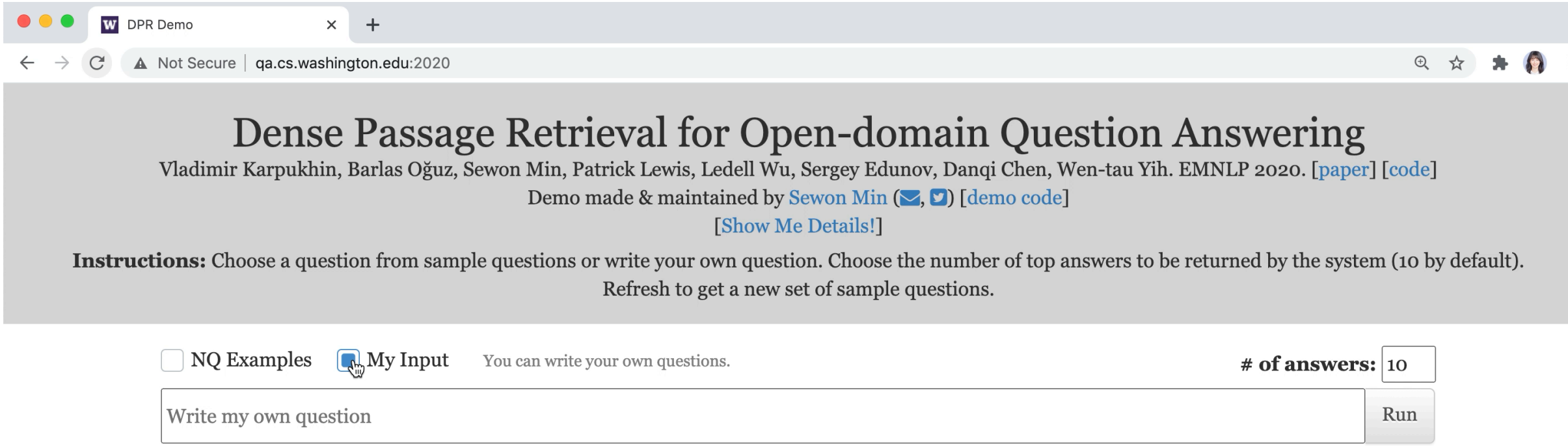
End QA Results

- *Pipeline*
- *End2end w/ pretraining*
- *DPR*

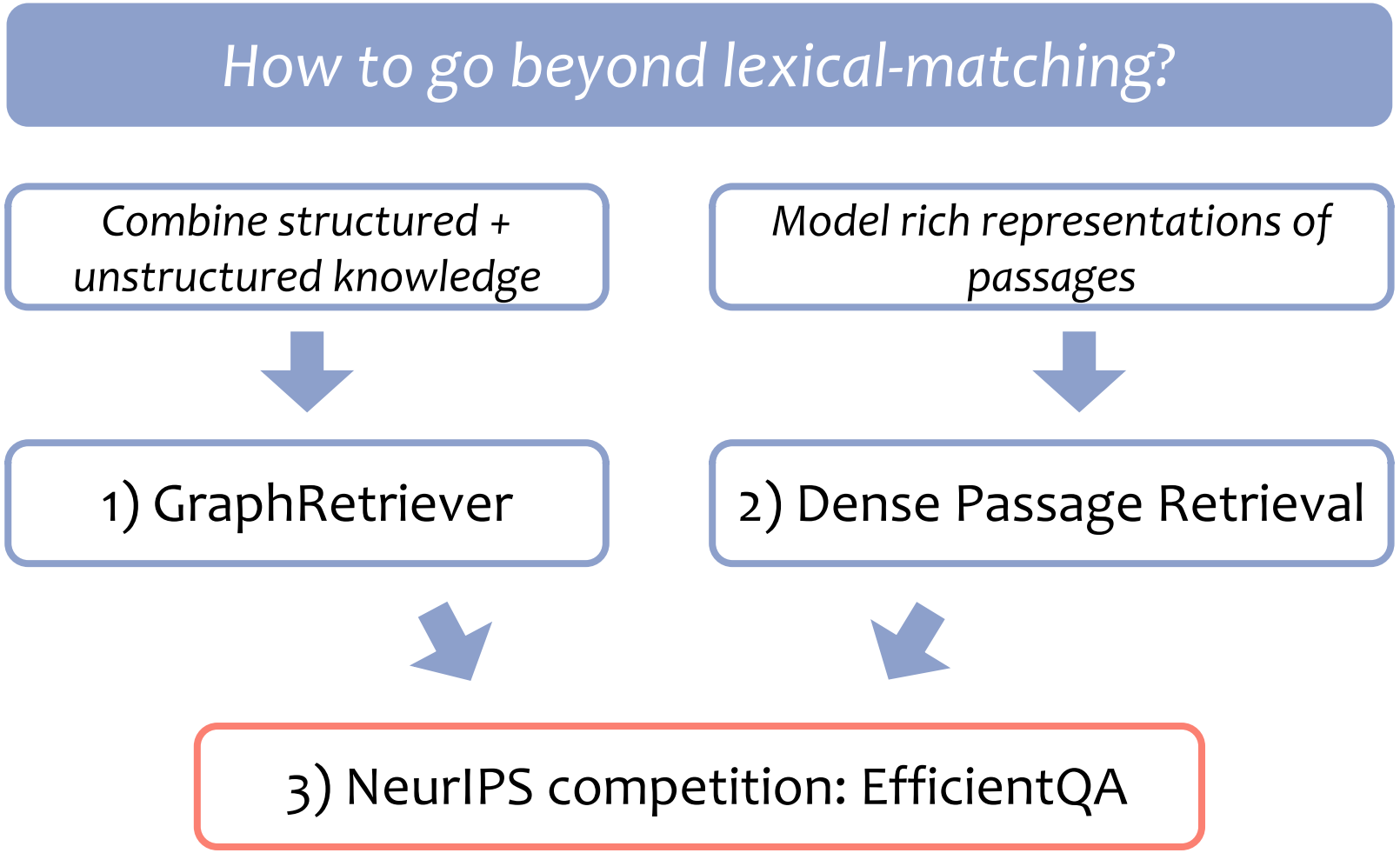


DPR Demo:

qa.cs.washington.edu:2020



Our Work


How to go beyond lexical-matching?
Combine structured + unstructured knowledge
Model rich representations of passages
1) GraphRetriever
2) Dense Passage Retrieval
3) NeurIPS competition: EfficientQA

EfficientQA competition @ NeurIPS 2020

- Focuses on naturally-occurring, open-domain QA with memory constraint
- 39 submissions from 18 unique teams within 2 months
- Significant advances in SOTA

All the top 8 submissions used ***DPR***

Top 1 submission combines DPR & structured + unstructured knowledge

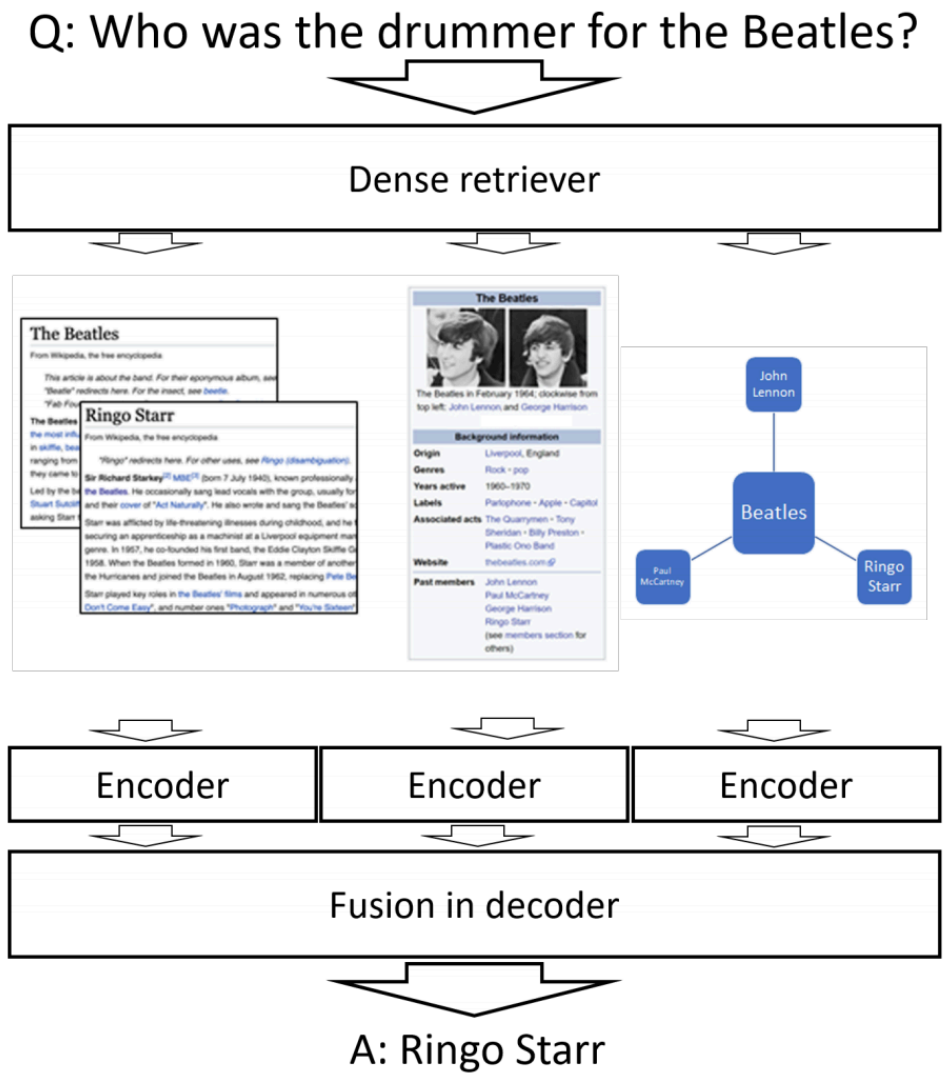
Min, Boyd-Graber, Alberti, Chen, Choi, Collins, Guu, Hajishirzi, Lee, Palomaki, Raffel, Roberts, Kwiatkowski. "NeurIPS 2020 EfficientQA Competition: Systems, Analyses and Lessons Learned" PMLR 2021

EfficientQA competition @ NeurIPS 2020

Top 1 submission from facebook Artificial Intelligence

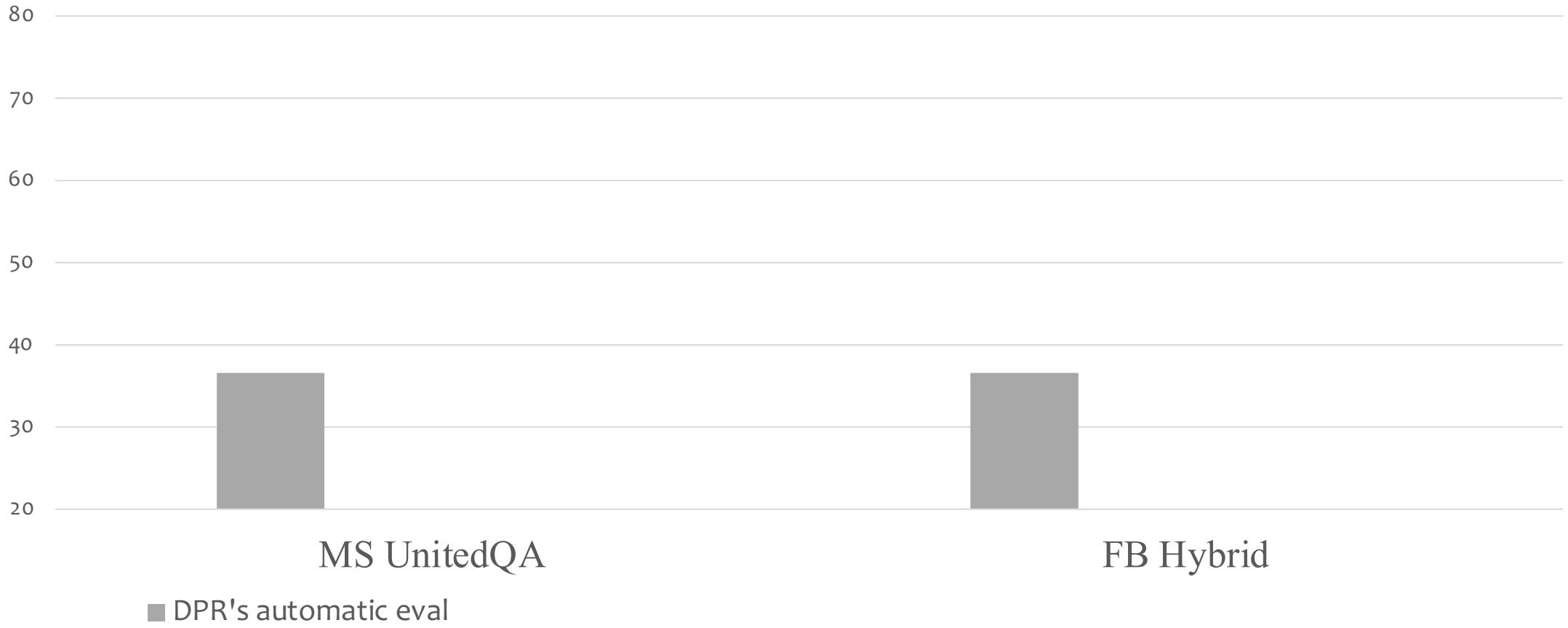
- Get dense representations of
  - Unstructured knowledge from Wikipedia (passages)
  - Structured knowledge from Wikipedia (tables, lists)
  - Structured knowledge from knowledge bases

(figure from their paper)



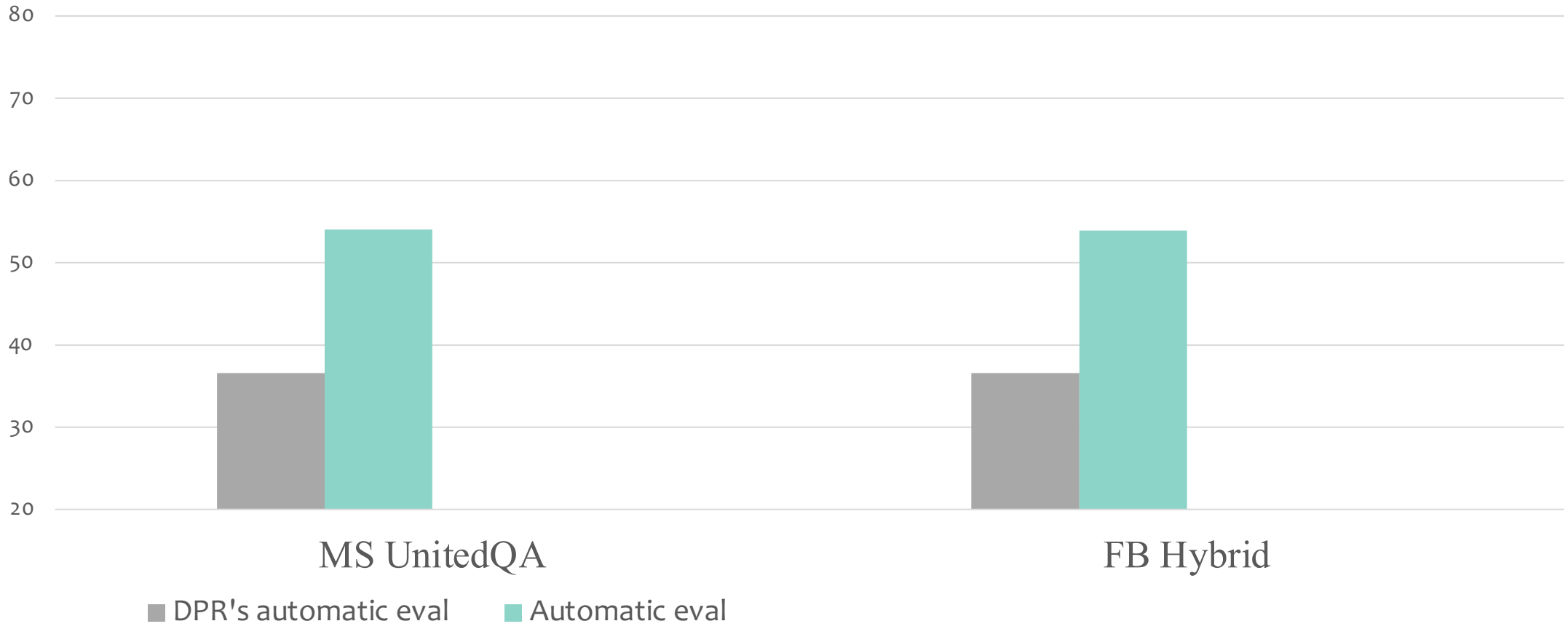
Oguz et al. 2020. "Unified Open-Domain Question Answering with Structured and Unstructured Knowledge"

EfficientQA competition @ NeurIPS 2020



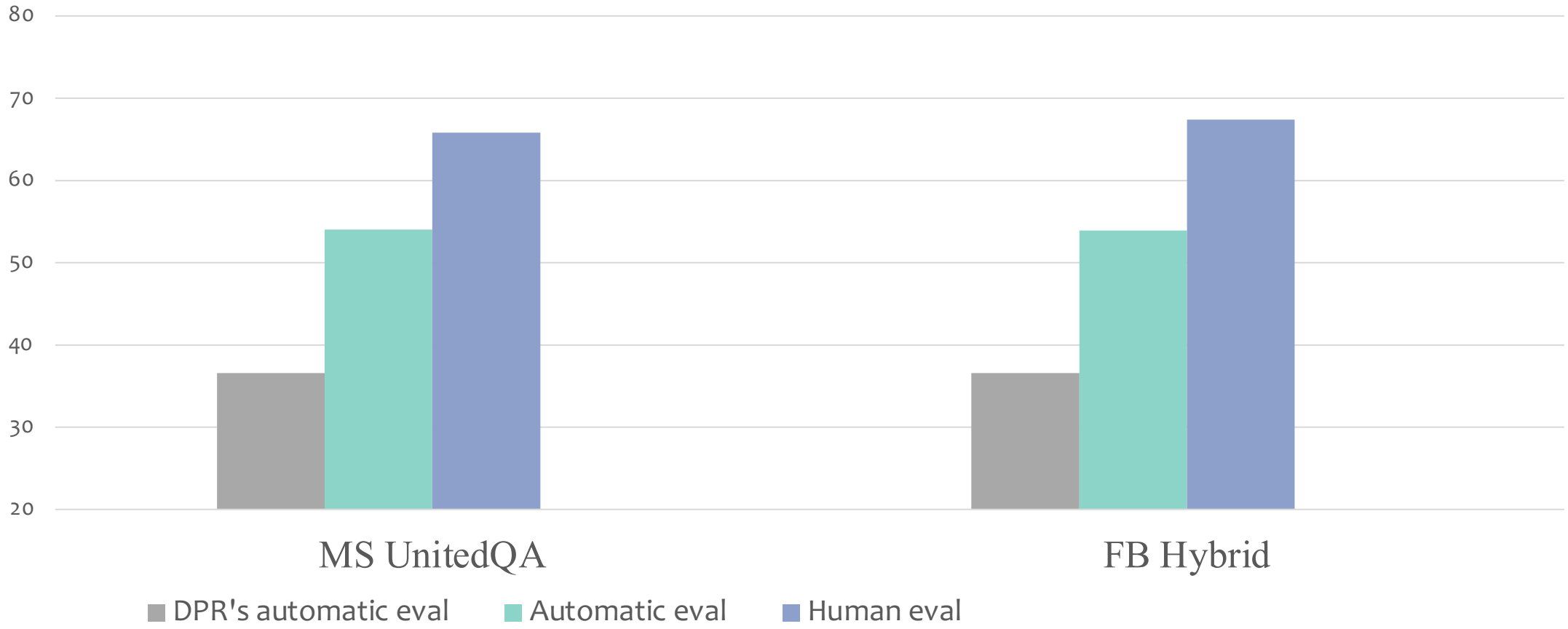
Min, Boyd-Graber, Alberti, Chen, Choi, Collins, Guu, Hajishirzi, Lee, Palomaki, Raffel, Roberts, Kwiatkowski. "NeurIPS 2020 EfficientQA Competition: Systems, Analyses and Lessons Learned" PMLR 2021

EfficientQA competition @ NeurIPS 2020



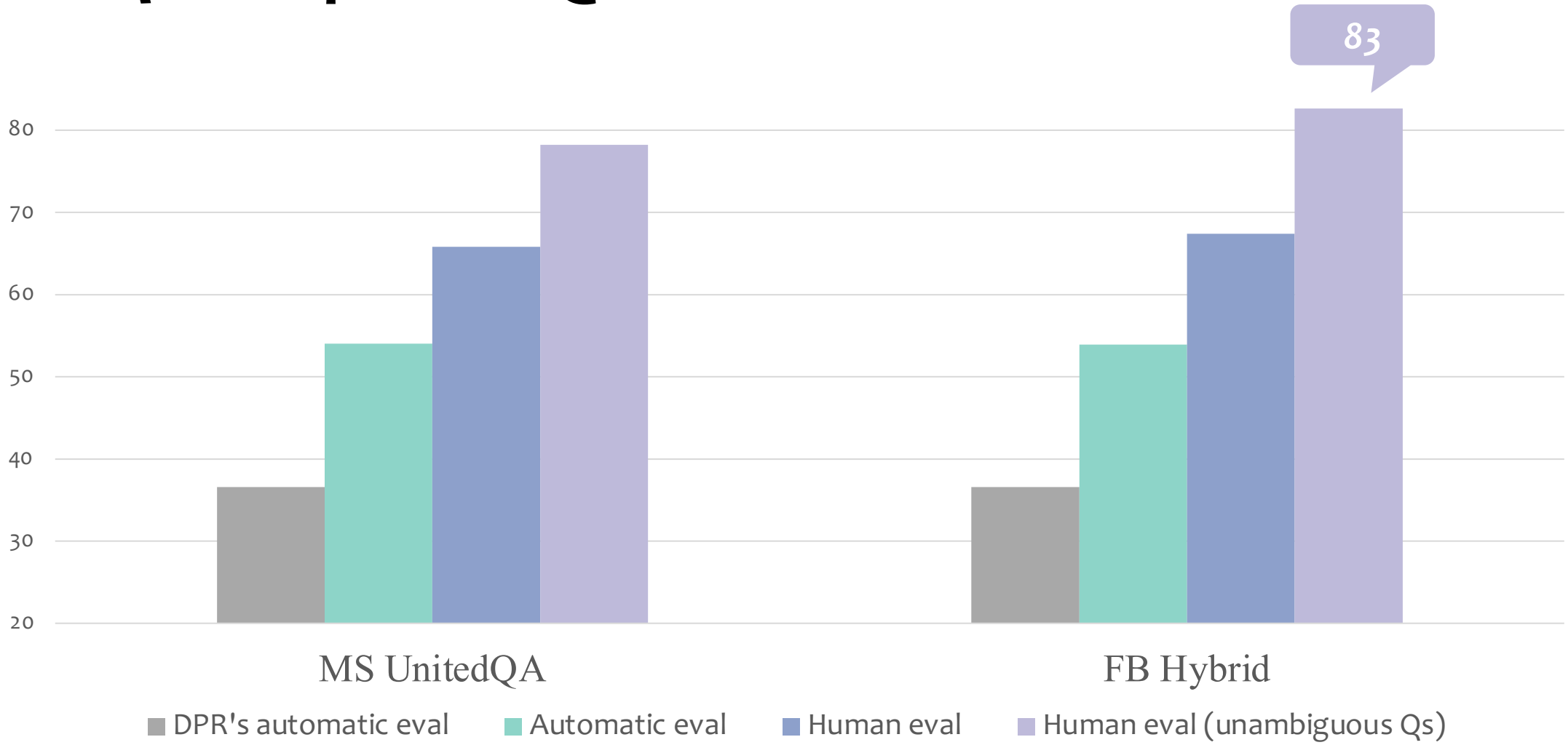
Min, Boyd-Graber, Alberti, Chen, Choi, Collins, Guu, Hajishirzi, Lee, Palomaki, Raffel, Roberts, Kwiatkowski. "NeurIPS 2020 EfficientQA Competition: Systems, Analyses and Lessons Learned" PMLR 2021

EfficientQA competition @ NeurIPS 2020



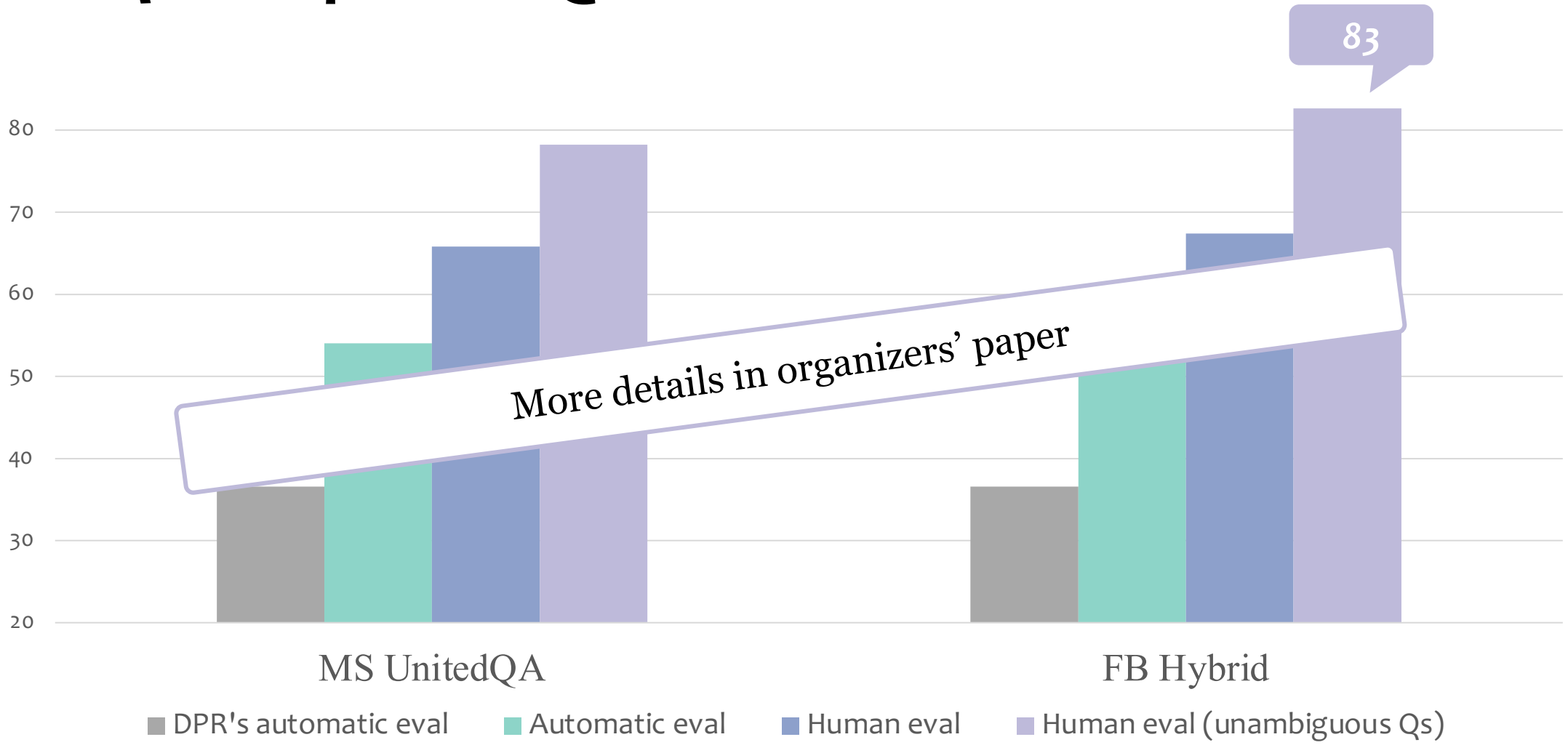
Min, Boyd-Graber, Alberti, Chen, Choi, Collins, Guu, Hajishirzi, Lee, Palomaki, Raffel, Roberts, Kwiatkowski. "NeurIPS 2020 EfficientQA Competition: Systems, Analyses and Lessons Learned" PMLR 2021

EfficientQA competition @ NeurIPS 2020



Min, Boyd-Graber, Alberti, Chen, Choi, Collins, Guu, Hajishirzi, Lee, Palomaki, Raffel, Roberts, Kwiatkowski. "NeurIPS 2020 EfficientQA Competition: Systems, Analyses and Lessons Learned" PMLR 2021

EfficientQA competition @ NeurIPS 2020



Min, Boyd-Graber, Alberti, Chen, Choi, Collins, Guu, Hajishirzi, Lee, Palomaki, Raffel, Roberts, Kwiatkowski. "NeurIPS 2020 EfficientQA Competition: Systems, Analyses and Lessons Learned" PMLR 2021

Overview

I. Beyond questions with lexical cues

Advanced SOTA: 26 → 83

GraphRetriever (Min et al. 2020, Li et al. EMNLP 2020)

DPR (Karpukhin et al. EMNLP 2020)

EfficientQA competition (Min et al. 2021, Submitted to PMLR 2021)

II. Beyond unambiguous questions

Identified problem for the first time

AmbigQA (Min et al. EMNLP 2020)

Joint Passage Retrieval (Min et al. 2021)

III. Future directions

Motivation

When did Harry Potter and the Sorcerer's stone movie come out?

What season do Meredith and Derek get married in Grey's Anatomy?

*from NQ, a benchmark in naturally-occurring open-domain QA

Motivation

When did Harry Potter and the Sorcerer's stone movie come out?

Harry Potter and the Philosopher's Stone (film)

... had its world premiere at the Odeon Leicester Square in London on 4 November 2001 (...) released to cinemas in the United Kingdom and United States on 16 November 2001.

What season do Meredith and Derek get married in Grey's Anatomy?

Now or Never (Grey's Anatomy)

... Season 5 ... Meredith and Derek have decided not to wait any longer to get married and just go to City Hall that evening ... writes their vows down on a post-it note that they both sign ...

Grey's Anatomy (Season 7)

... She and Derek decide to adopt Zola, an orphaned baby, and make their marriage legal.

Min, Michael, Hajishirzi, Zettlemoyer. "AmbigQA: Answering Ambiguous Open-domain Questions". EMNLP 2020

Motivation

When did Harry Potter and the Sorcerer's stone movie come out?

Harry Potter and the Philosopher's Stone (film)

... had its world premiere at the Odeon Leicester Square in London on 4 November 2001 (...) released to cinemas in the United Kingdom and United States ... 01.

What season d... marri...

... Never (Grey's Anatomy)

... Season 5 ... Meredith and Derek have decided not to wait any longer to get married and just go to City Hall that evening ... writes their vows down on a post-it note that they both sign ...

Grey's Anatomy (Season 7)

... She and Derek decide to adopt Zola, an orphaned baby, and make their marriage legal.

Over 50% of questions from NQ are ambiguous

Min, Michael, Hajishirzi, Zettlemoyer. "AmbigQA: Answering Ambiguous Open-domain Questions". EMNLP 2020

Motivation

When did Harry Potter and the Sorcerer's stone movie come out?

Harry Potter and the Philosopher's Stone (film)

... had its world premiere at the Odeon Leicester Square in London on 4 November 2001 (...) released to cinemas in the Kingdom and

Real questions are inherently ambiguous

When people ask questions in new / unfamiliar topics, questions cannot be guaranteed to have a single clear answer (even when people intend to do so)

Wha

cided

ust go to

City Hall that evening ... writes their vows down on a post-it note that they both sign ...

Grey's Anatomy (Season 7)

... She and Derek decide to adopt Zola, an orphaned baby, and make their marriage legal.

Min, Michael, Hajishirzi, Zettlemoyer. "AmbigQA: Answering Ambiguous Open-domain Questions". EMNLP 2020

AmbigQA task

When did Harry Potter and the Sorcerer's stone movie come out?

Q: When did harry potter and the sorcerer's stone movie come out at the Odeon Leicester Square?
A: 4 November 2001

Q: When did harry potter and the sorcerer's stone movie come out in cinemas?
A: 16 November 2001

What season do Meredith and Derek get married in Grey's Anatomy?

Q: What season do Meredith and Derek get informally married in Grey's Anatomy?
A: Season 5

Q: What season do Meredith and Derek get legally married in Grey's Anatomy?
A: Season 7

Explicit answers to the original question
+ disambiguation in a more well-defined way

Min, Michael, Hajishirzi, Zettlemoyer. "AmbigQA: Answering Ambiguous Open-domain Questions". EMNLP 2020

A new problem!

- Open-domain QA research has always assumed each question has a single clear answer
- Nonetheless, Kwiatkowski et al 2019 report that the answers are often debatable; an average pairwise agreement of the answers in NQ annotations is 49.2%

We **embrace ambiguity** as **inherent** to open-domain questions

Min, Michael, Hajishirzi, Zettlemoyer. "AmbigQA: Answering Ambiguous Open-domain Questions". EMNLP 2020

A new problem!

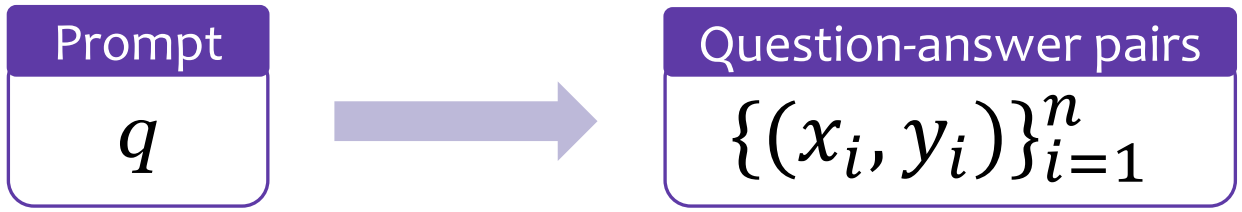
- Open-domain QA research has always assumed each question has a single clear answer
- Nonetheless, Kwiatkowski et al 2019 report that the answers are often debatable; an average pairwise agreement of the answers in NQ annotations is 49.2%

We **embrace ambiguity** as **inherent** to open-domain questions

- Other work in ambiguous QA studied
  - Annotated, intentional ambiguous questions (Xu et al 2019)
  - Simple and vague query, e.g. *"dinasour"* (Zhai et al 2003, Aliannejadi et al 2019)

We study **unintentional** & **subtle** ambiguity

Task Definition



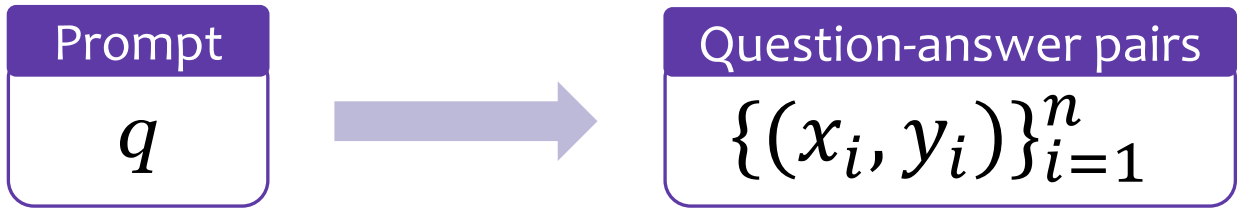
q: What season do … get married in Grey's Anatomy?

x_1: What season … get informally married in Grey's Anatomy?
y_1: Season 5

x_2: What season … get legally married in Grey's Anatomy?
y_2: Season 7

Min, Michael, Hajishirzi, Zettlemoyer. "AmbigQA: Answering Ambiguous Open-domain Questions". EMNLP 2020

Task Definition



q: What season do … get married in Grey's Anatomy?

x_1: What season … get informally married in Grey's Anatomy?

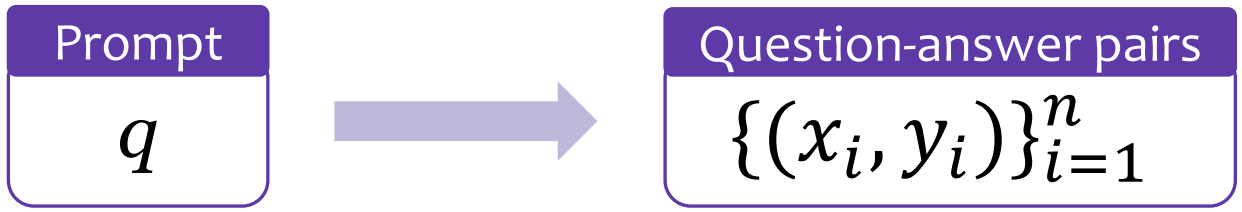
y_1: Season 5

x_2: What season … get legally married in Grey's Anatomy?

y_2: Season 7

y_i: an equally plausible answer

Task Definition



q: What season do … get married in Grey's Anatomy?

x_1: What season … get informally married in Grey's Anatomy?

y_1: Season 5

x_2: What season … get legally married in Grey's Anatomy?

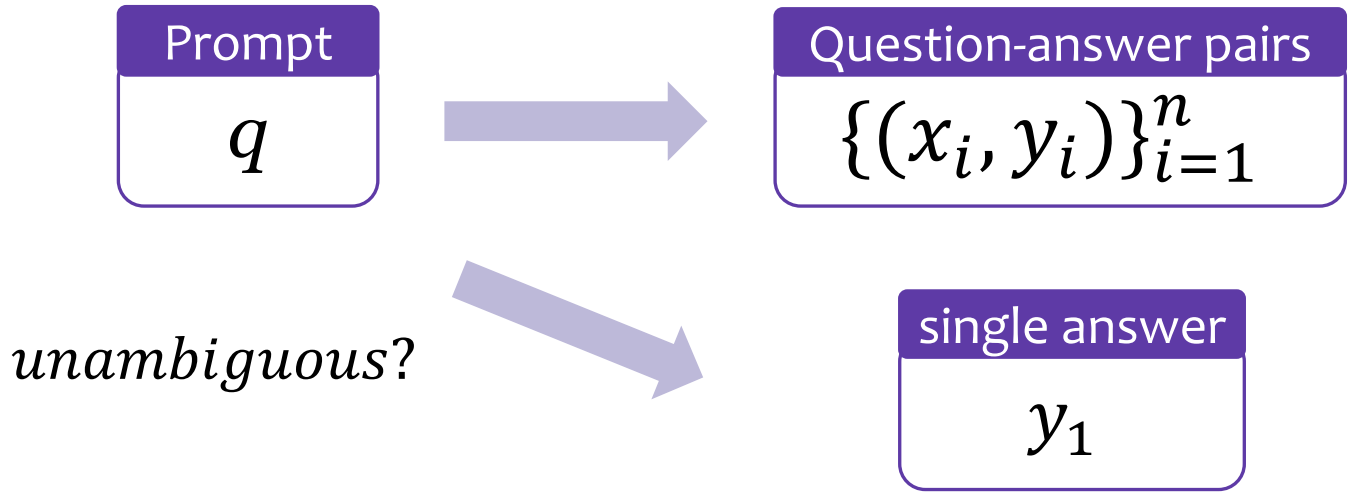
y_2: Season 7

y_i: an equally plausible answer

x_i: a *minimal* modification of q whose answer is unambiguously y_i

Min, Michael, Hajishirzi, Zettlemoyer. "AmbigQA: Answering Ambiguous Open-domain Questions". EMNLP 2020

Task Definition



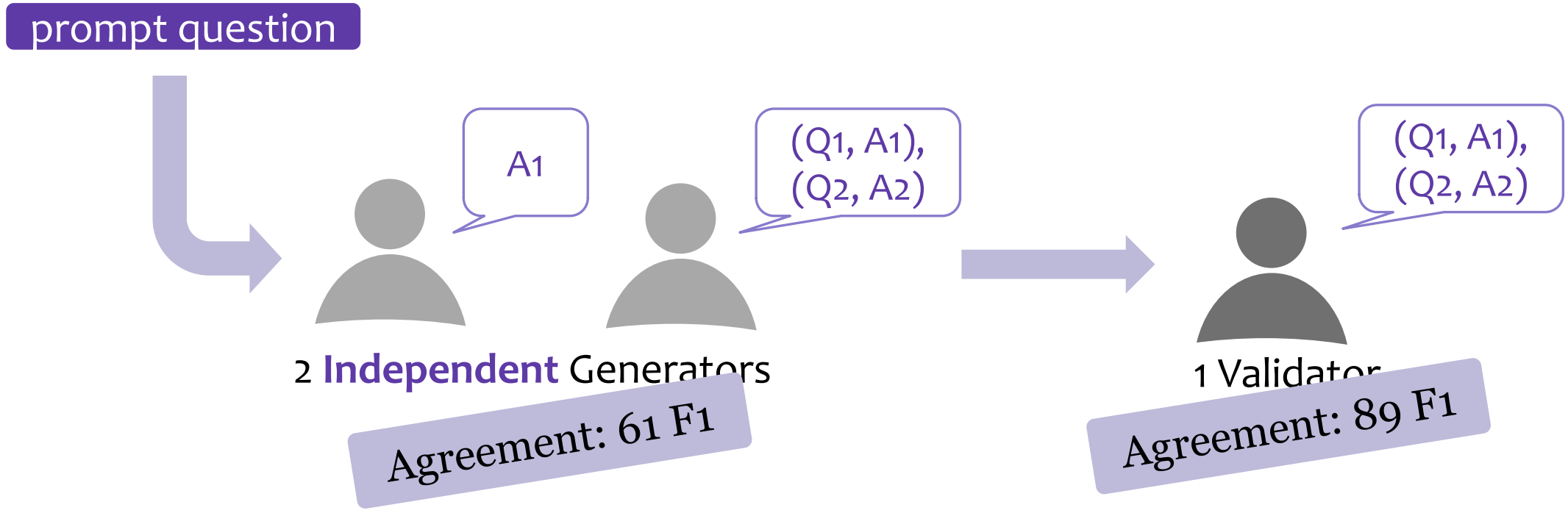
Data Collection

- Maximizing recall is difficult even for humans
- We were able to collect high quality data with high levels of ambiguity using **careful worker selection** and **an annotation pipeline: generation + validation**

Data Collection

- Maximizing recall is difficult even for humans
- We were able to collect high quality data with high levels of ambiguity using **careful worker selection** and **an annotation pipeline: generation + validation**

Naturally-occurring questions from NQ (Kwiatkowski et al 2019)



Data Analysis

- 14,042 questions
- Over 50% of questions are ambiguous
- Diverse types of ambiguity with fairly long-tailed edits

Type of Ambiguity	%
Event references	39
Properties	27
Entity references	23
Answer types	16
Time-dependency	13
Multiple sub-questions	3

(c) Word cloud of the edits made in questions; ■ and ■ indicate added and deleted unigrams, respectively.

Data Analysis

- 14,042 questions
- Over 50% of questions are ambiguous
- Diverse types of ambiguity with fairly long-tailed edits

Type of Ambiguity	%
Ever	
Prop	27
Entity references	23
Answer types	16
Time-dependency	13
Multiple sub-questions	3

More analyses in the paper

(c) Word cloud of the edits made in questions; ■ and ■ indicate added and deleted unigrams, respectively.

Models

Disambig-First

- Disambiguation before reading Wikipedia

DPR

- SOTA QA model for multi-answer prediction, and then Disambiguation
- Does not consider a set of answers jointly

SpanSeqGen

- Combine SOTA QA and seq2seq that generates a sequence of answers, separated by `[SEP]`
- Considers a set of answers jointly

Please see the paper for details!

Democratic Co-training (Zhou & Goldman, 2004)

- Prevalence of unlabeled ambiguity in NQ → Let's use ***a single known answer*** from NQ as ***weak supervision***

Please see the paper for details!

Democratic Co-training (Zhou & Goldman, 2004)

- Prevalence of unlabeled ambiguity in NQ → Let's use ***a single known answer*** from NQ as ***weak supervision***

1. Train C sequence-to-sequence QA models on ***AmbigQA data***

SpanSeqGen 1

SpanSeqGen 2

⋮

SpanSeqGen C

Please see the paper for details!

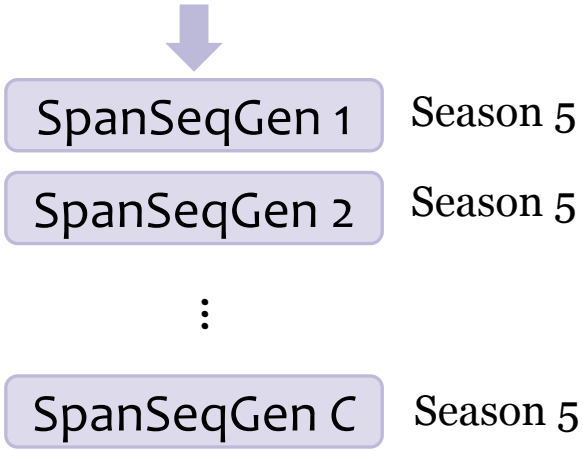
Democratic Co-training (Zhou & Goldman, 2004)

- Prevalence of unlabeled ambiguity in NQ → Let's use ***a single known answer*** from NQ as ***weak supervision***

1. Train C sequence-to-sequence QA models on ***AmbigQA data***

2. Make inference on ***NQ***, using a known answer as **prefix** during generation

What season does Meredith and Derek get married in Grey's Anatomy? (NQ answer: Season 5)



Please see the paper for details!

Democratic Co-training (Zhou & Goldman, 2004)

- Prevalence of unlabeled ambiguity in NQ → Let's use ***a single known answer*** from NQ as ***weak supervision***

1. Train C sequence-to-sequence QA models on ***AmbigQA data***

2. Make inference on ***NQ***, using a known answer as **prefix** during generation

What season does Meredith and Derek get married in Grey's Anatomy? (NQ answer: Season 5)

SpanSeqGen 1 — Season 5 [EOS]

SpanSeqGen 2 — Season 5 [SEP] Season 7 [EOS]

⋮

SpanSeqGen C — Season 5 [SEP] Season 7 [EOS]

Please see the paper for details!

Democratic Co-training (Zhou & Goldman, 2004)

- Prevalence of unlabeled ambiguity in NQ → Let's use ***a single known answer*** from NQ as ***weak supervision***

1. Train C sequence-to-sequence QA models on ***AmbigQA data***

2. Make inference on ***NQ***, using a known answer as **prefix** during generation

3. If there are **extra answers** generated by majority of C models, include them to ***(expanded) AmbigQA data***

What season does Meredith and Derek get married in Grey's Anatomy? (NQ answer: Season 5)

SpanSeqGen 1 — Season 5 [EOS]

SpanSeqGen 2 — Season 5 [SEP] Season 7 [EOS]

⋮

SpanSeqGen C — Season 5 [SEP] Season 7 [EOS]

Data = Data ∪ {"What season .. Anatomy?", [Season 5, Season 7]}

Please see the paper for details!

Democratic Co-training (Zhou & Goldman, 2004)

- Prevalence of unlabeled ambiguity in NQ → Let's use ***a single known answer*** from NQ as ***weak supervision***

1. Train C sequence-to-sequence QA models on ***AmbigQA data***
2. Make inference on ***NQ***, using a known answer as **prefix** during generation
3. If there are **extra answers** generated by majority of C models, include them to ***(expanded) AmbigQA data***

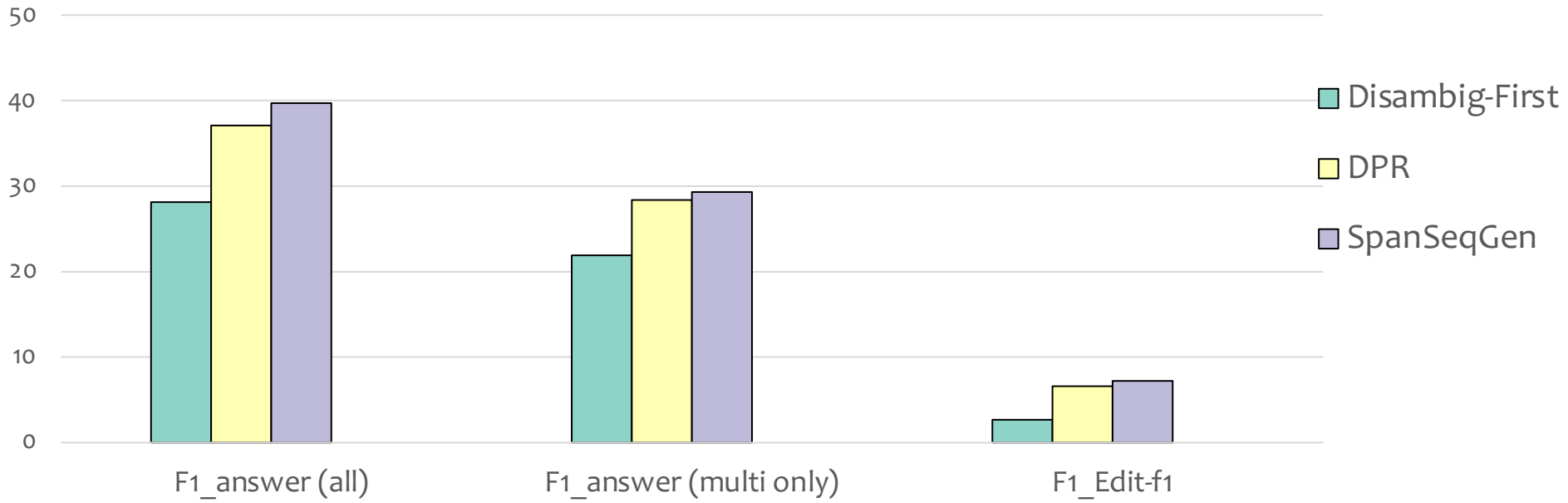
What season does Meredith and Derek get married in Grey's Anatomy? (NQ answer: Season 5)

SpanSeqGen 1	Season 5 [EOS]
SpanSeqGen 2	Season 5 [SEP] Season 7 [EOS]
⋮	
SpanSeqGen C	Season 5 [SEP] Season 7 [EOS]

Data = Data ∪ {"What season .. Anatomy?", [Season 5, Season 7]}

Please see the paper for details!

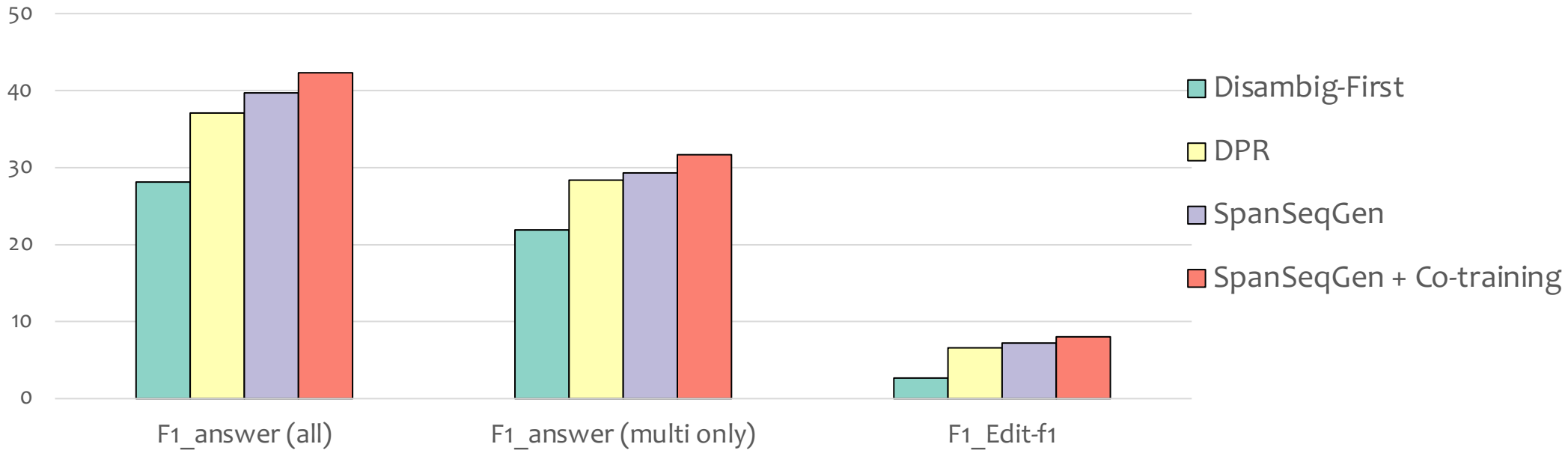
Results



SpanSeqGen outperforms other baselines.

(Esp. Disambig-First, which does disambiguation before reading passages)

Results



SpanSeqGen outperforms other baselines.

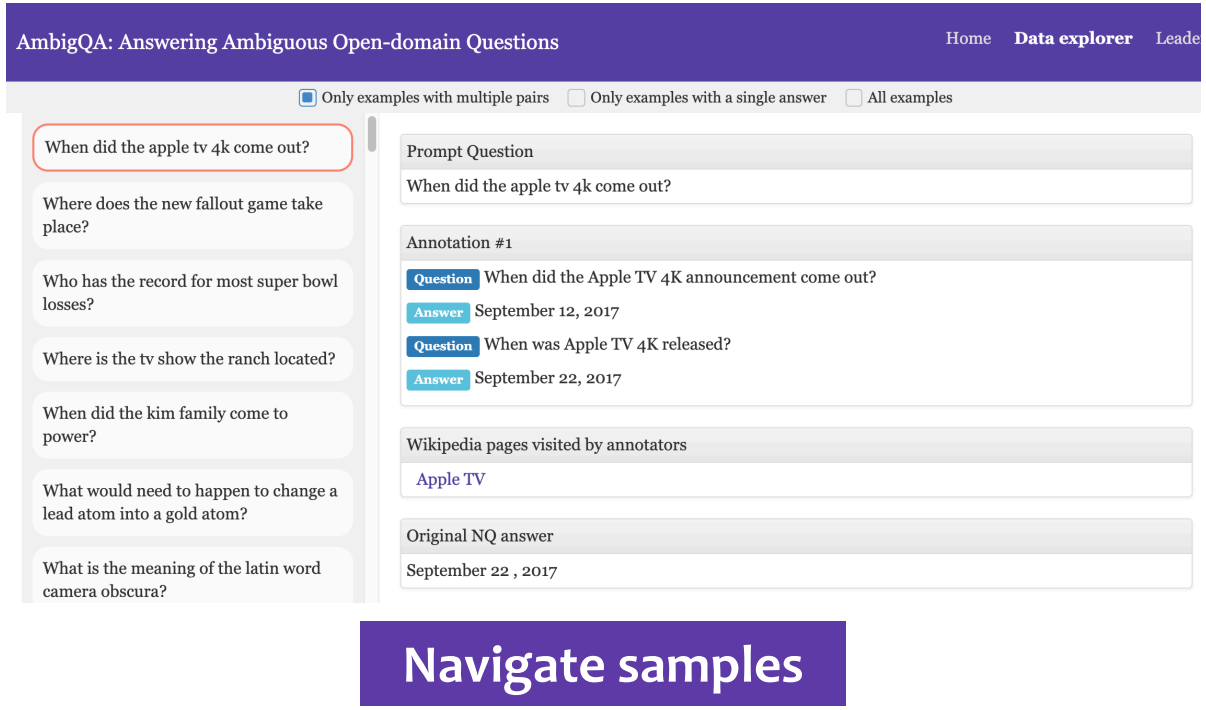
(Esp. Disambig-First, which does disambiguation before reading passages)

SpanSeqGen + Co-training further boosts the performance

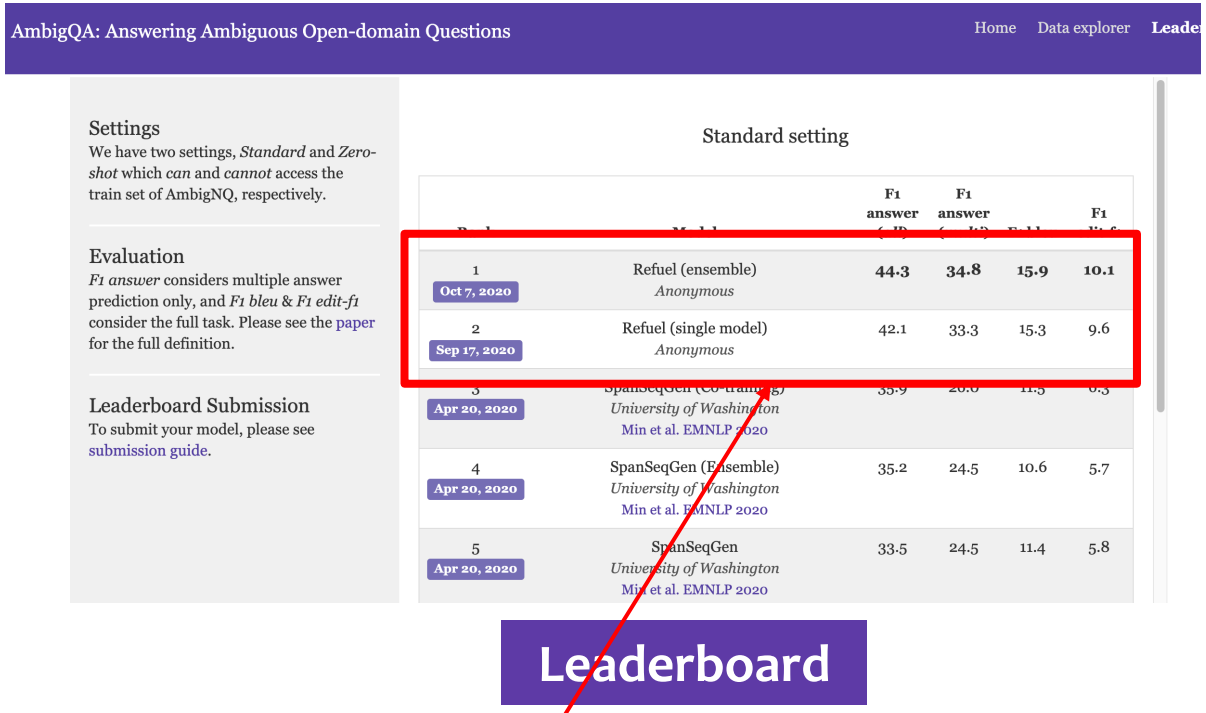
Still huge room for improvements

Website

https://nlp.cs.washington.edu/ambigqa/



Navigate samples



Leaderboard

New SOTA models from the community!

One step toward progress

One step toward progress: Multi-answer Retrieval

"Who played Mark on the TV show Roseanne?"

Glenn Quinn … his portrayal of Mark Healy on the popular '90s family sitcom Roseanne.

Roseanne (season 10) … Ames McNamara was announced to be cast as Mark Conner-Healy.

Min, Lee, Chang, Toutanova, Hajishirzi. "Joint Passage Ranking for Multi-Answer Retrieval". Preprint. 2021

One step toward progress: Multi-answer Retrieval

- Previous retrieval

"Who played Mark on the TV show Roseanne?"

Glenn Quinn … his portrayal of Mark Healy on the popular '90s family sitcom Roseanne.

Glenn Quinn, who played Becky's husband, Mark, …

Roseanne (season 10) … Ames McNamara was announced to be cast as Mark Conner-Healy.

on the hit ABC sitcom Roseanne as the younger brother of Mark Healy (Glenn Quinn) …

Becky begins dating Mark Healy (Glenn Quinn) …

Any of these passages is good

Min, Lee, Chang, Toutanova, Hajishirzi. "Joint Passage Ranking for Multi-Answer Retrieval". Preprint. 2021

One step toward progress: Multi-answer Retrieval

- Previous retrieval focuses on independent modeling of each passage: $P(p_i|q)$

"Who played Mark on the TV show Roseanne?"

Glenn Quinn … his portrayal of Mark Healy on the popular '90s family sitcom Roseanne.

Glenn Quinn, who played Becky's husband, Mark, …

Roseanne (season 10) … Ames McNamara was announced to be cast as Mark Conner-Healy.

on the hit ABC sitcom Roseanne as the younger brother of Mark Healy (Glenn Quinn) …

Becky begins dating Mark Healy (Glenn Quinn) …

Any of these passages is good

Min, Lee, Chang, Toutanova, Hajishirzi. "Joint Passage Ranking for Multi-Answer Retrieval". Preprint. 2021

One step toward progress: Multi-answer Retrieval

- Previous retrieval focuses on independent modeling of each passage: $P(p_i|q)$
- Multi-answer retrieval

"Who played Mark on the TV show Roseanne?"

Glenn Quinn … his portrayal of Mark Healy on the popular '90s family sitcom Roseanne.

Glenn Quinn, who played Becky's husband, Mark, …

Roseanne (season 10) … Ames McNamara was announced to be cast as Mark Conner-Healy.

Good

on the hit ABC sitcom Roseanne as the younger brother of Mark Healy (Glenn Quinn) …

Becky begins dating Mark Healy (Glenn Quinn) …

Bad

Min, Lee, Chang, Toutanova, Hajishirzi. "Joint Passage Ranking for Multi-Answer Retrieval". Preprint. 2021

One step toward progress: Multi-answer Retrieval

- Previous retrieval focuses on independent modeling of each passage: $P(p_i|q)$
- Multi-answer retrieval intrinsically requires ***joint*** modeling of passages: $P(p_1 \ldots p_k|q)$

"Who played Mark on the TV show Roseanne?"

Glenn Quinn … his portrayal of Mark Healy on the popular '90s family sitcom Roseanne.

Glenn Quinn, who played Becky's husband, Mark, …

Roseanne (season 10) … Ames McNamara was announced to be cast as Mark Conner-Healy.

Good

on the hit ABC sitcom Roseanne as the younger brother of Mark Healy (Glenn Quinn) …

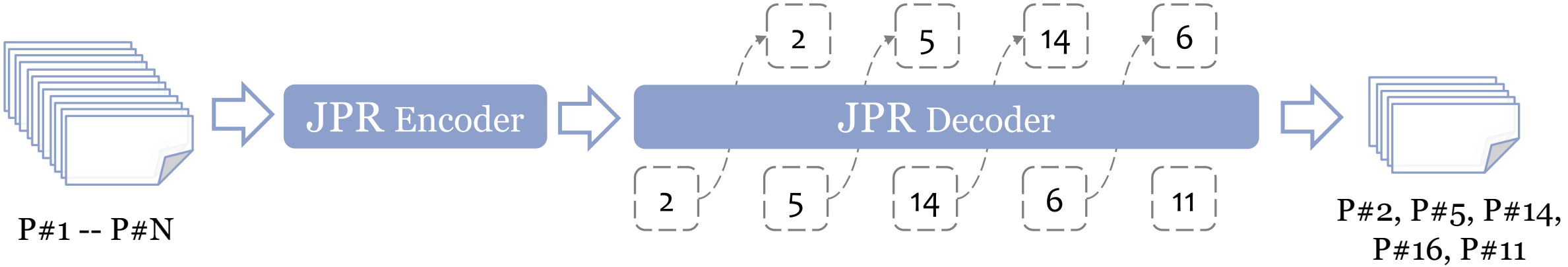
Becky begins dating Mark Healy (Glenn Quinn) …

Bad

Min, Lee, Chang, Toutanova, Hajishirzi. "Joint Passage Ranking for Multi-Answer Retrieval". Preprint. 2021

Joint Passage Retrieval (JPR)

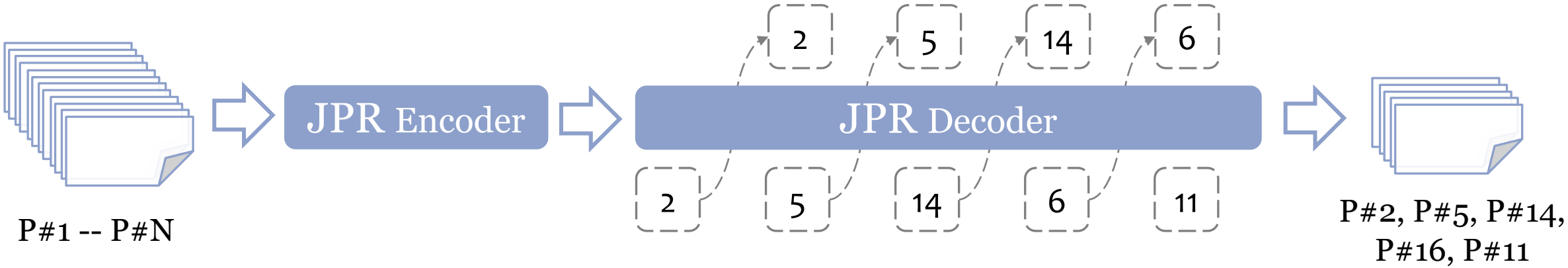
- Goal: maximize the set coverage



Min, Lee, Chang, Toutanova, Hajishirzi. "Joint Passage Ranking for Multi-Answer Retrieval". Preprint. 2021

Joint Passage Retrieval (JPR)

- Goal: maximize the set coverage



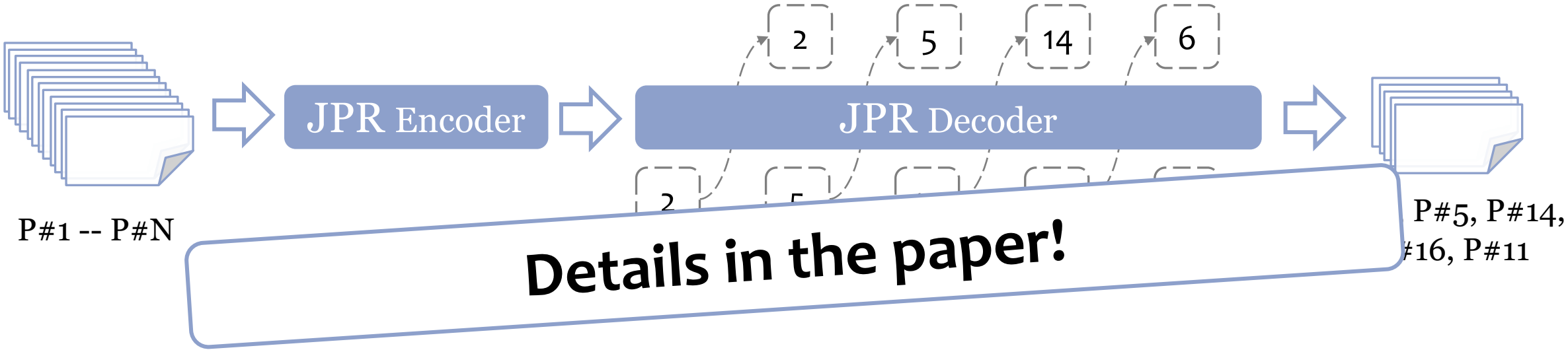
Training: encourages multiple ways to reach the multi-answer supervision

Inference: a new algorithm that decodes a tree

Min, Lee, Chang, Toutanova, Hajishirzi. "Joint Passage Ranking for Multi-Answer Retrieval". Preprint. 2021

Joint Passage Retrieval (JPR)

- Goal: maximize the set coverage



Training: encourages multiple ways to reach the multi-answer supervision

Inference: a new algorithm that decodes a tree

Min, Lee, Chang, Toutanova, Hajishirzi. "Joint Passage Ranking for Multi-Answer Retrieval". Preprint. 2021

Joint Passage Retrieval (JPR)

Q: Who play Mark on the TV show Roseanne?	
INDEPPR	JPR
#1 Glenn Quinn ... He was best known for his portrayal of Mark Healy on the popular '90s family sitcom Roseanne.	**#1** Glenn Quinn ... He was best known for his portrayal of Mark Healy on the popular '90s family sitcom Roseanne.
#2 Glenn Quinn, who played Becky's husband, Mark, died in December 2002 of a heroin overdose at the age of 32 ...	**#2** Becky begins dating Mark Healy (Glenn Quinn) ...
#3 Becky begins dating Mark Healy (Glenn Quinn) ...	**#3** Glenn Quinn, who played Becky's husband, Mark, died in December 2002 of a heroin overdose at the age of 32 ...
#4 Johnny Galecki ... on the hit ABC sitcom Roseanne as the younger brother of Mark Healy (Glenn Quinn) ...	**#4** Roseanne (season 10) ... In September 2017, Ames McNamara was announced to be cast as Mark Conner-Healy.

Baseline: $P(p_i|q)$

Repeatedly retrieve the same answer

JPR: $P(p_1 \ldots p_k|q)$

Successfully cover a novel answer

Follow-up work by other researchers

- Proposing better AmbigQA models (Gao et al 2020, Dhole 2020)
- Studying ambiguity as a failure of the presuppositions about "uniqueness" (Kim et al 2021)
- Suggesting explainable evidence to resolve ambiguity (Lamm et al 2020, Gonzalez et al 2020)
- Focusing on specific type of ambiguity, e.g. temporal ambiguity
- Analyze ambiguity in questions w.r.t. gender and nationality bias (Gor et al 2021)

Overview

I. Beyond questions with lexical cues

Advanced SOTA: 26 → 83

GraphRetriever (Min et al. 2020, Li et al. EMNLP 2020)

DPR (Karpukhin et al. EMNLP 2020)

EfficientQA competition (Min et al. 2021, Submitted to PMLR 2021)

II. Beyond unambiguous questions

Identified problem for the first time

AmbigQA (Min et al. EMNLP 2020)

Joint Passage Retrieval (Min et al. 2021)

III. Future directions

Future direction (1/3)

Q: Why is the winter solstice only the beginning of winter and not the middle?

Seasonal lag is the phenomenon whereby the date of maximum average air temperature at a geographical location on a planet is delayed until some time after the date of maximum insolation (i.e. the Summer Solstice).

Open-ended QA

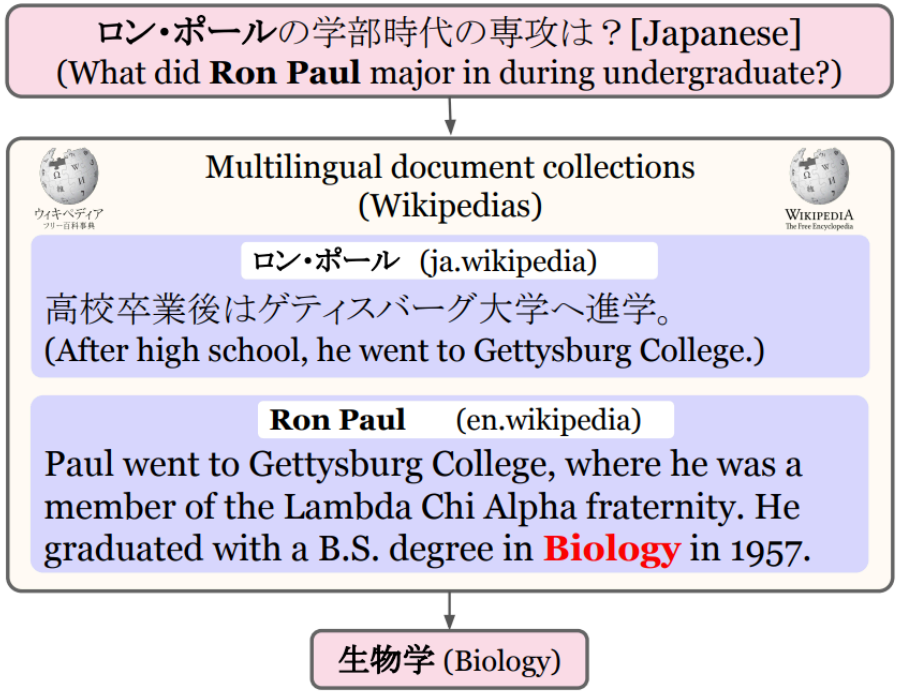
Q: If you can get frostbite within 10 minutes outside, how can people do cryogenic therapy that's -200 degrees?

Cryotherapy This treatment involves exposing individuals to extremely cold dry air (below −100 °C) for two to four minutes. (…) **Adverse effects** … are suspected of being underreported. (…) there is the risk of inert gas asphyxiation as well as frostbite.

Questions with false presuppositions

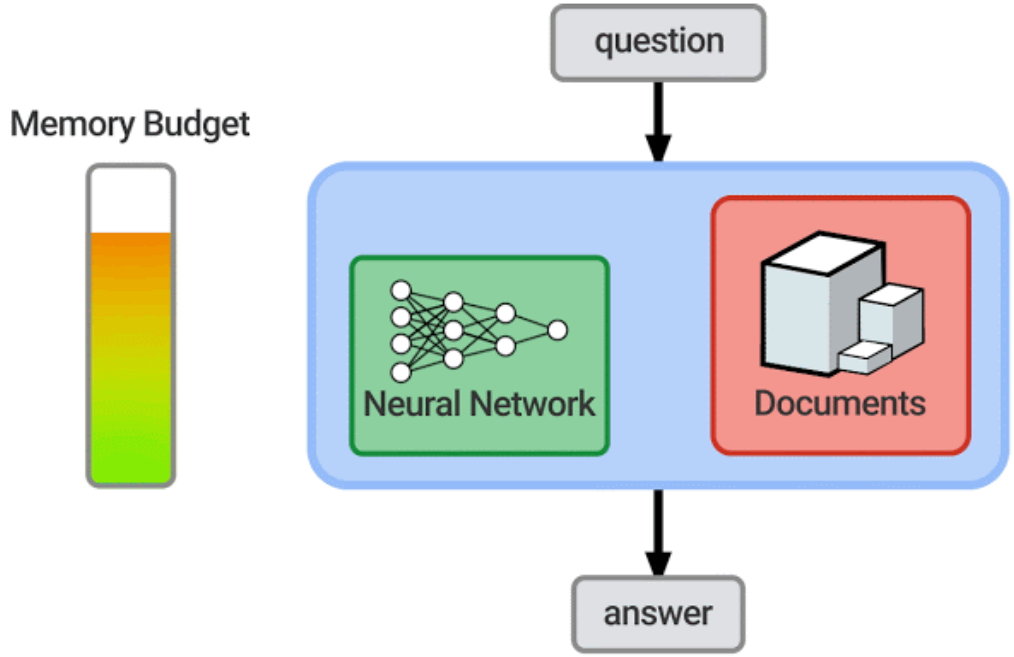
Questions from ELI5 (Fan et al 2019)

Future direction (2/3)



Multilingual / Cross-lingual NLP/QA

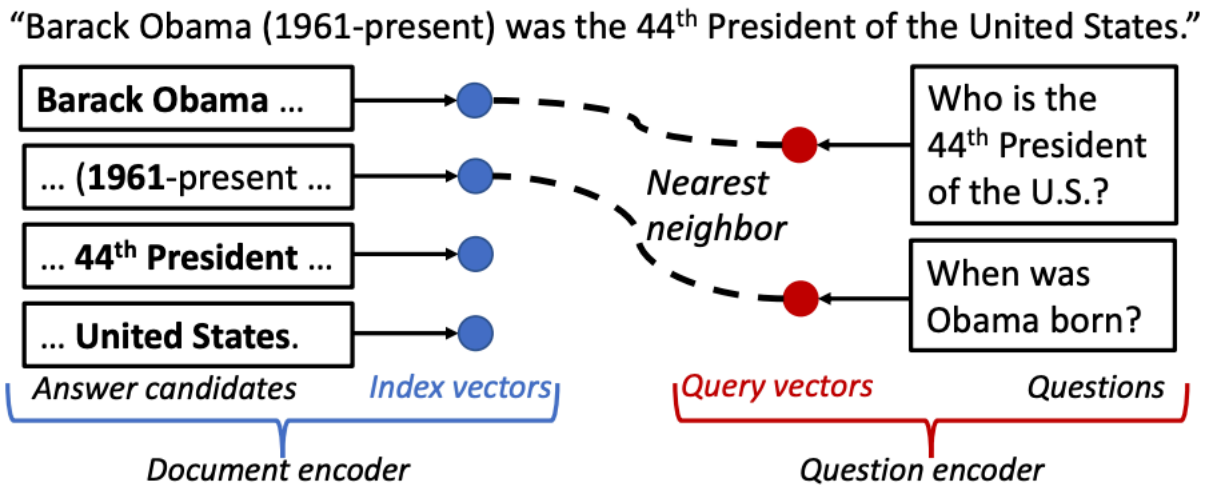
(Figure from Asai et al 2021)



Memory-efficient NLP/QA (e.g. on-device system)

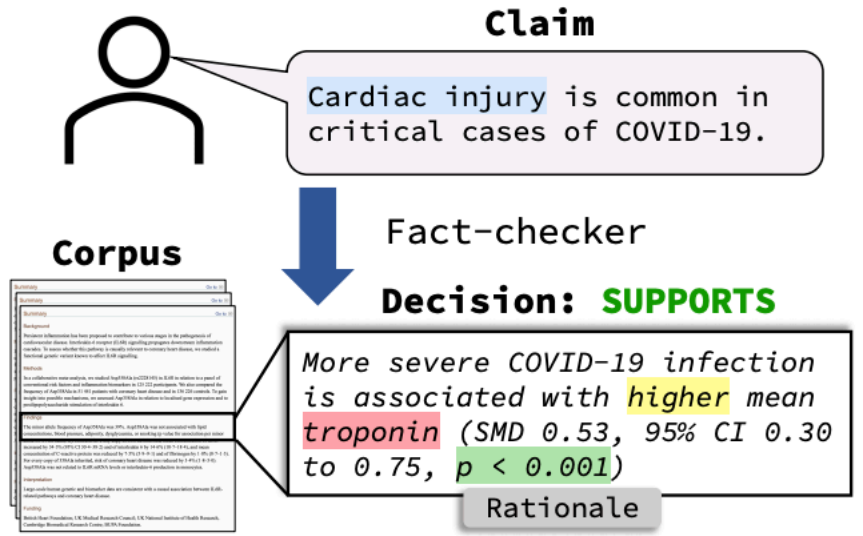
(Figure from EfficientQA blog)

Future direction (3/3)



Low-latency QA

(Figure from Seo et al 2018)



Domain adaptation / generalization to OOD

(Figure from Wadden et al 2020)

Summary

Prior work on annotated data suffer from answering naturally-occurring, open-domain questions.

We have made progress on questions with little lexical cues through advanced retrieval systems

We have introduced a new class of task, data and models for answering ambiguous questions

More problems to be explored on a wider type of questions, in various different languages and domains, with deployable systems in real-world.

Summary

Prior work on annotated data suffer from answering naturally-occurring, open-domain questions.

We have made progress on questions with little lexical cues through advanced retrieval systems

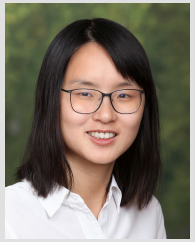
We have introduced a new class of task, data and models for answering ambiguous questions

More problems to be explored on a wider type of questions, in various different languages and domains, with deployable systems in real-world.

Altogether, we are in the progress toward answering **any kind of questions** in the world!

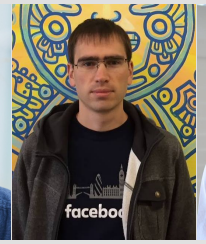


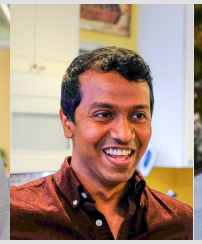


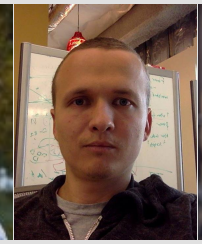



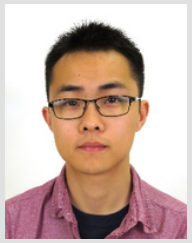
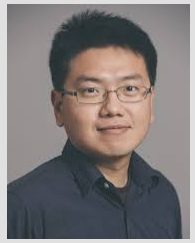




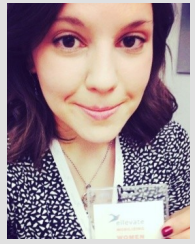



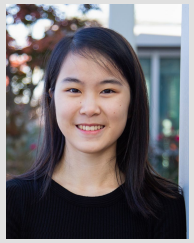


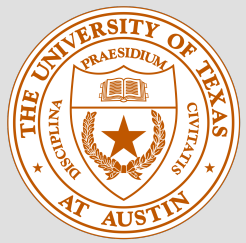










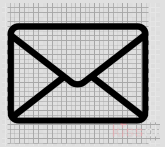




Thanks to my **collaborators** in academia + industry

Thank you for listening!

 shmsw25.github.io

 sewon@cs.washington.edu

 @sewon__min