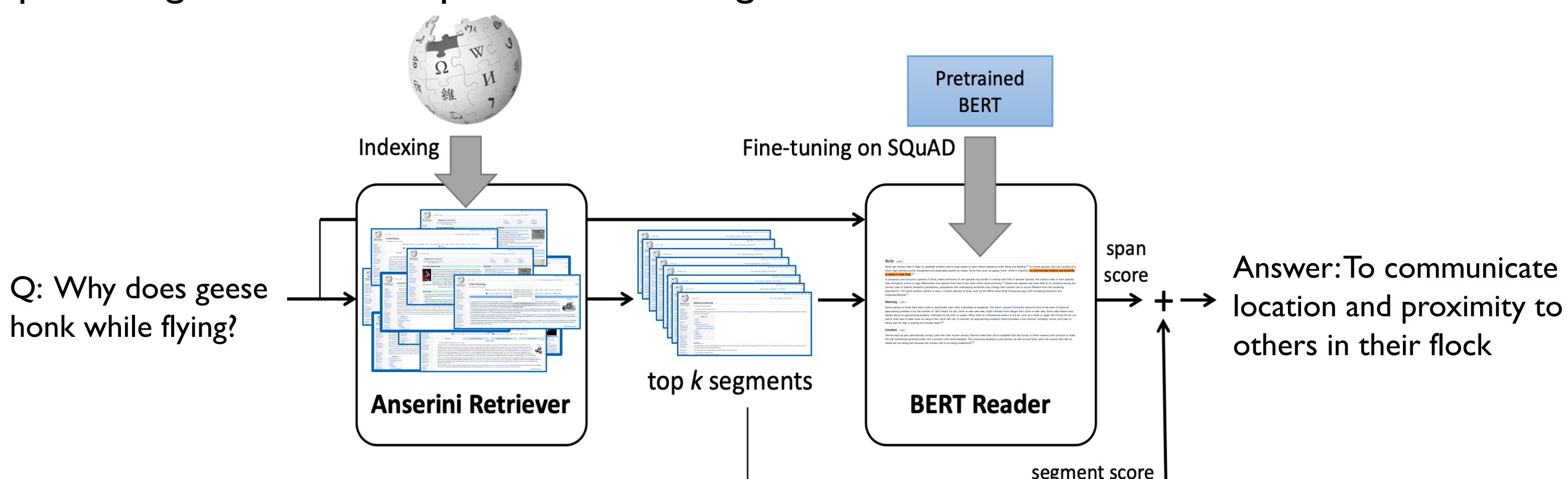


End-to-End Open-Domain Question Answering with BERTserini

Wei Yang,^{1,2} Yuqing Xie,^{1,2} Aileen Lin,² Xingyu Li,²
Luchen Tan,² Kun Xiong,² Ming Li,^{1,2} Jimmy Lin^{1,2}

What is BERTserini?

BERTserini answers questions from Wikipedia by integrating passage retrieval techniques using the Anserini open-source IR engine with a BERT reader.



How well does it work?

BERTserini achieves state of the art in exact match and F1 identifying answers from Wikipedia.

Model	EM	F1	R
Dr.QA (Chen et al., 2017)	27.1	-	77.8
Dr.QA + Fine-tune	28.4	-	-
Dr.QA + Multitask	29.8	-	-
R ³ (Wang et al., 2017)	29.1	37.5	-
Kratzwald and Feuerriegel (2018)	29.8	-	-
Par. R. (Lee et al., 2018)	28.5	-	83.1
Par. R. + Answer Agg.	28.9	-	-
Par. R. + Full Agg.	30.2	-	-
MINIMAL (Min et al., 2018)	34.7	42.5	64.0
BERTserini (Article, $k = 5$)	19.1	25.9	63.1
BERTserini (Paragraph, $k = 29$)	36.6	44.0	75.0
BERTserini (Sentence, $k = 78$)	34.0	41.0	67.5
BERTserini (Paragraph, $k = 100$)	38.6	46.1	85.8

Table 1: Results on SQuAD development questions.

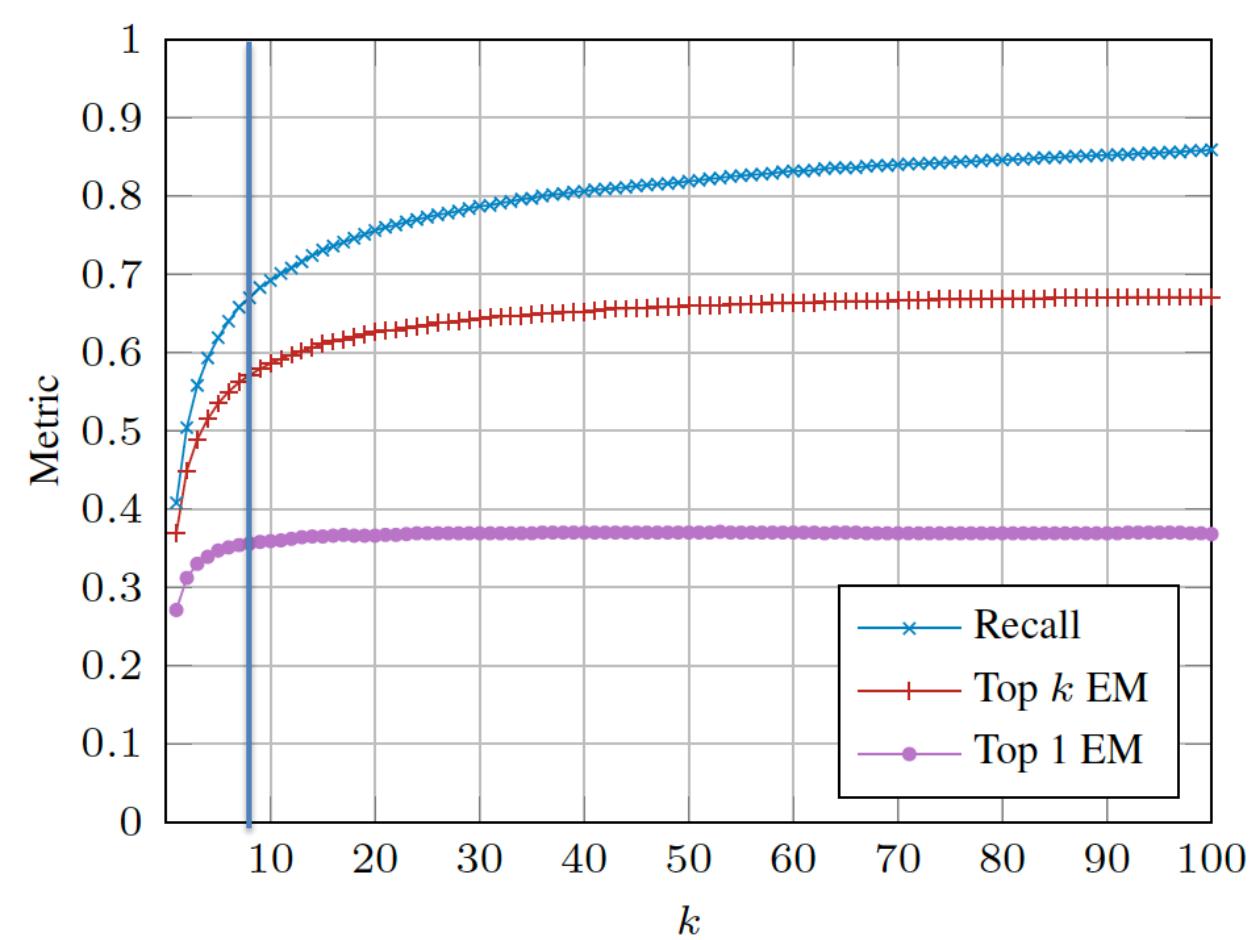


Figure 2: Model effectiveness with different numbers of retrieved paragraphs.

Where can I try it out?

Try BERTserini on our bot platform!



The screenshot shows a conversation with the BERTserini bot. The bot asks "How are you?", and the user replies "I am doing well.". When asked "Where was Van Gogh born?", the bot provides the answer: "Theo van Gogh was born on 23 July 1857 in The Hague, Netherlands, to Anneke and Johan van Gogh. (Wikipedia: Theo van Gogh (film director))". The bot then asks "By what name was Mongol leader Temujin better known?", and the user replies "Mongol chief Temujin (later to be known as Genghis Khan) falls for Bortai, the daughter of the Tartar leader, and steals her away, precipitating war. (Wikipedia: The Conqueror (1956 film))". The bot asks "Why do people wear Camouflage?", and the user replies "The camouflage pattern in clothing was developed to help military personnel be less visible to enemy forces. (Wikipedia: Fashion)". The bot concludes with a "Reset" button.

Adding @DoDo to your Slack workspace!

The screenshot shows a conversation in a Slack workspace. The bot (@DoDo) asks "Hi", and the user replies "Hello". The bot asks "Who is Prof Ming Li?", and the user replies "Ming Li is a CRC Chair Professor in Bioinformatics, of Computer Science at the University of Waterloo. (Wikipedia: Ming Li)". The bot asks "What's the weather in Canada?", and the user replies "The majority of the country has a cold or severely cold winter climate, but southerly areas are warm in summer. (Wikipedia: Canada)". The bot asks "Why do geese honk while flying?", and the user replies "Some fowl, such as geese, 'honk' while in migration to communicate location and proximity to others in their flock. (Wikipedia: Contact call)". The bot asks "What is the best algorithm for linear programming?", and the user replies "However, it takes only a moment to find the optimum solution by posing the problem as a linear program and applying the Simplex algorithm. (Wikipedia: George Dantzig)". The bot ends with a "Message DoDo" button.