

ReSCORE: Label-free Iterative Retriever Training for Multi-hop Question Answering with Relevance-Consistency Supervision

Dosung Lee, Wonjun Oh, Boyoung Kim, Minyoung Kim,
Joonsuk Park, Paul Hongsuck Seo

Korea University, NAVER AI Lab, NAVER Cloud, University of Richmond
{dslee1219, owj0421, bykimby, omniverse186, phseo}@korea.ac.kr, park@joonsuk.org

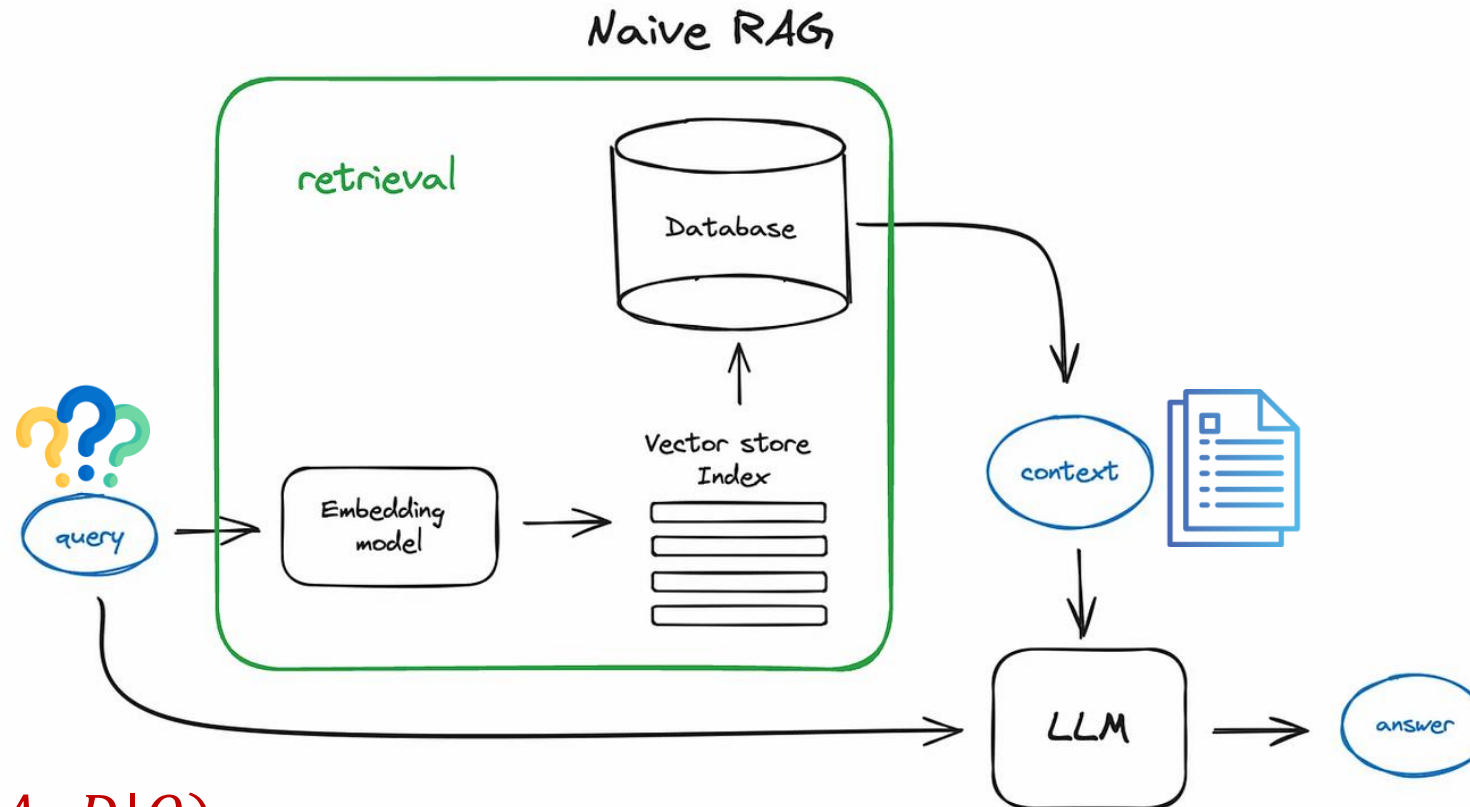


KOREA
UNIVERSITY



Multimodal Interactive
Intelligence Laboratory

Retrieval Augmented Generation

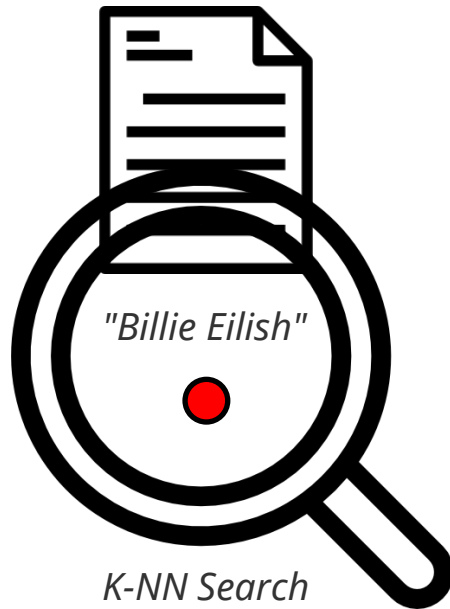


$$P_{QA}(A|Q) = \sum_D P(A, D|Q)$$

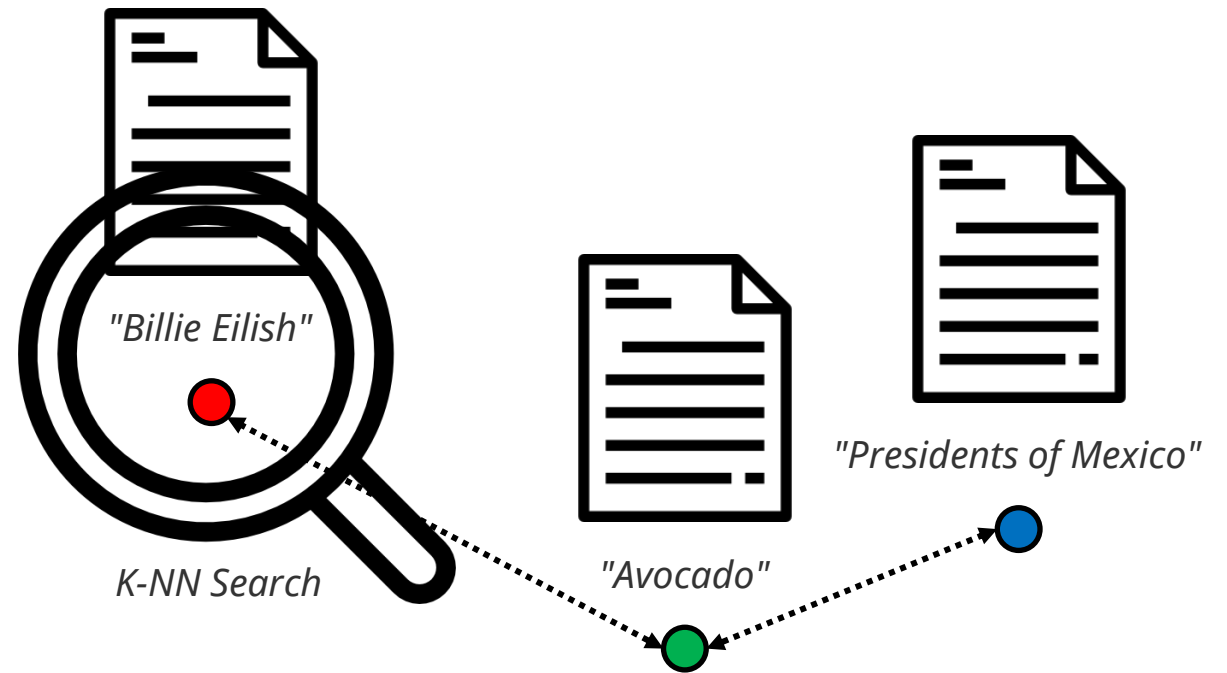
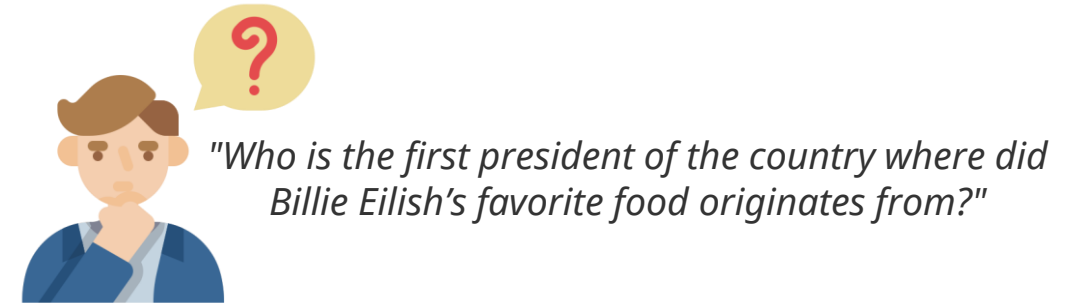
$$P_{RAG}(A, D|Q) = P_{generate}(A|D, Q) P_{retrieve}(D|Q)$$

Multi-hop Question Answering & Challenges

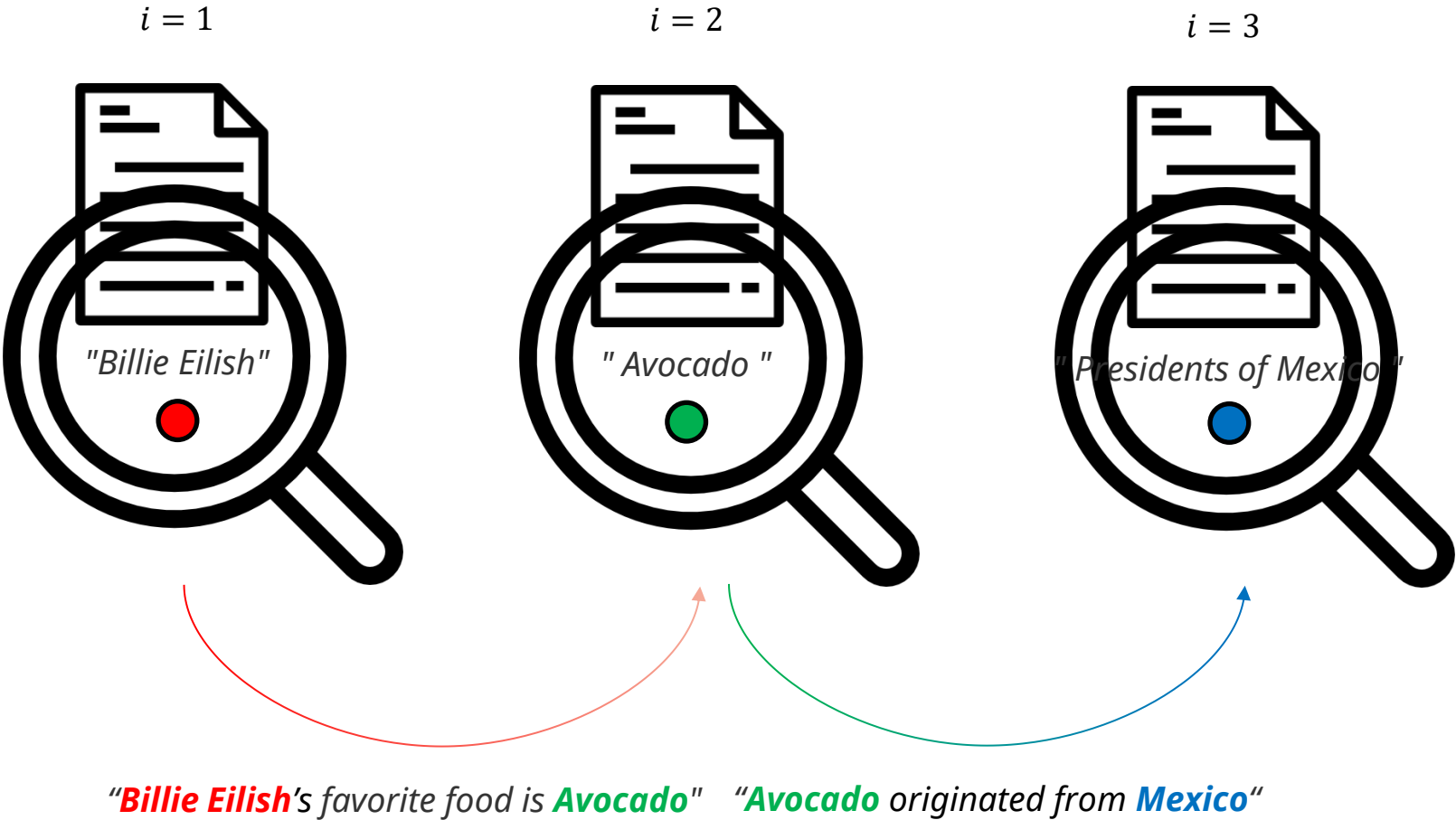
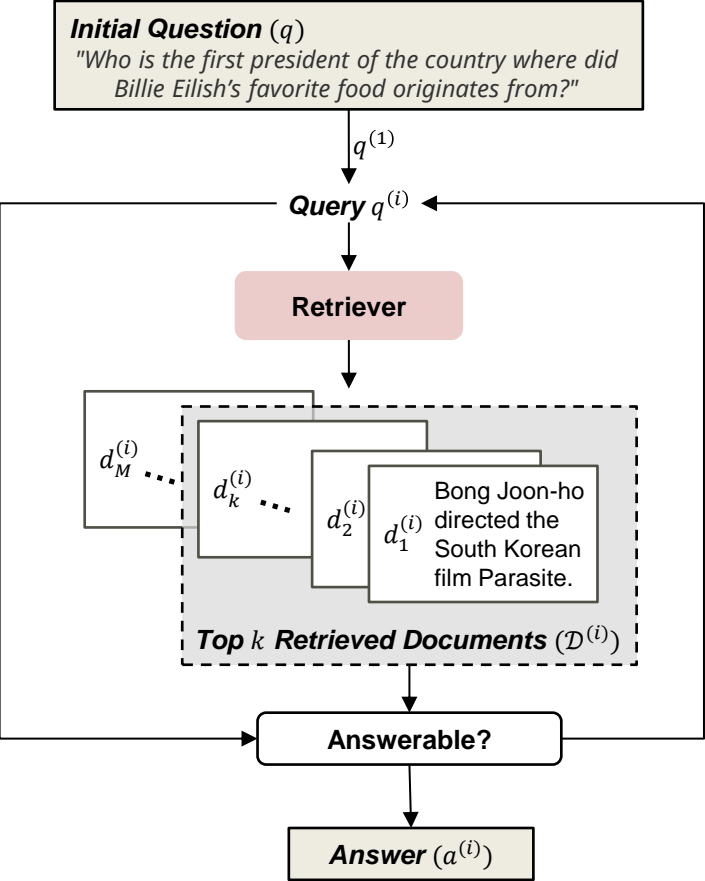
Single-hop



Multi-hop



Iterative RAG Inference Framework



Challenges of Labeling MHQA & Label-free Training



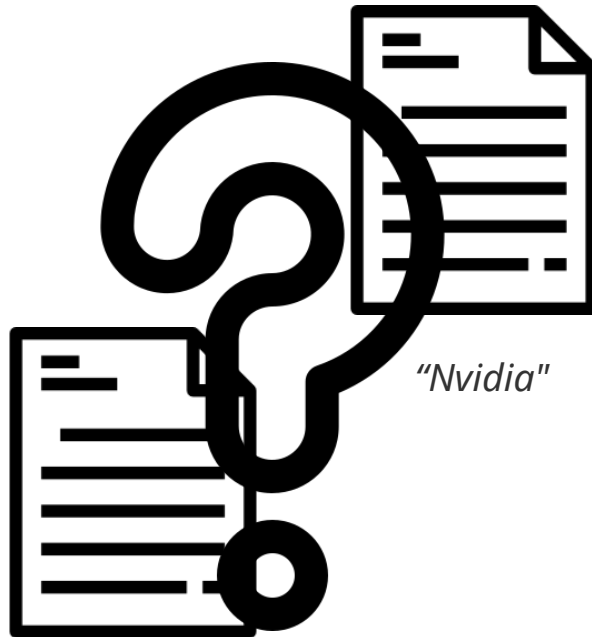
"Who is the first president of the country where did Billie Eilish's favorite food originates from?"



"Billie Eilish"



"Elon Musk"



"South Korea"



"Nvidia"



" Presidents of Mexico "



" Avocado"



"Paris"

Relevance-Consistency Supervision

$$P_{QA}(A|Q) = \sum_D P(A, D|Q)$$

$$P_{RAG}(A, D|Q) = P_{generate}(A|D, Q) P_{retrieve}(D|Q)$$

$$P_{Retriever}(D|Q) \propto P_{LLM}(A, Q|D)$$

$$P_{LLM}(A, Q|D) = P_{consistency}(A|D, Q) P_{relevance}(Q|D)$$



"Who is the first president of the country where did Billie Eilish's favorite food originates from?"



"Presidents of Mexico"

$P_{consistency}(A|D, Q)$



"Billie Eilish"

$P_{relevance}(Q|D)$

ReSCORE Training

