AACL 2025 Presenter: Yu Xia yux078@ucsd.edu

Knowledge-Aware Query Expansion with Large Language Models for Textual and Relational Retrieval

Yu Xia¹ Junda Wu¹ Sungchul Kim² Tong Yu² Ryan A. Rossi² Haoliang Wang² Julian McAuley¹

¹University of California San Diego ²Adobe Research

Adobe

Motivation

LLMs have been used to generate query expansions augmenting information retrieval.

Existing methods:

- o focus on textual similarities between queries and documents
- overlook relations between documents.
- often fail to handle complex queries with both textual and relational requirements

Our method:

- augment LLMs with structured document relations from Knowdelge Graph (KG)
- o use both textual and relational information for query expansion

Search Query Find me a highly rated camera for wildlife photography compatible with my Nikon F-Mount lenses. \$ **Query Expansion Query Expansion Query Expansion** Title: Nikon Coolpix P1000 Nikon camera Nikon Z7 II Feature: Feature: Feature: 45.7 megapixels; autofocus... impressive 125x zoom for the camera is designed wildlife photography... Description: for wildlife photography... Description: Description: wildlife flexible lens adapter... compatible with Nikon built-in lens options inc. Reviews: F-Mount lenses 4.5/5 "autofocus fast animals..." Nikon F-Mount lenses... Relations: Reviews: Reviews: 5 stars "captured wildlife "This is a great camera for products also bought: shooting wildlife animals! Nikon F-Mount lenses, ... birds without a hitch!" - products also viewed: FTZ adapter, .. **Final Retrieval Final Retrieval Final Retrieval S HyDE RAR** KAR (Ours)

Methodology

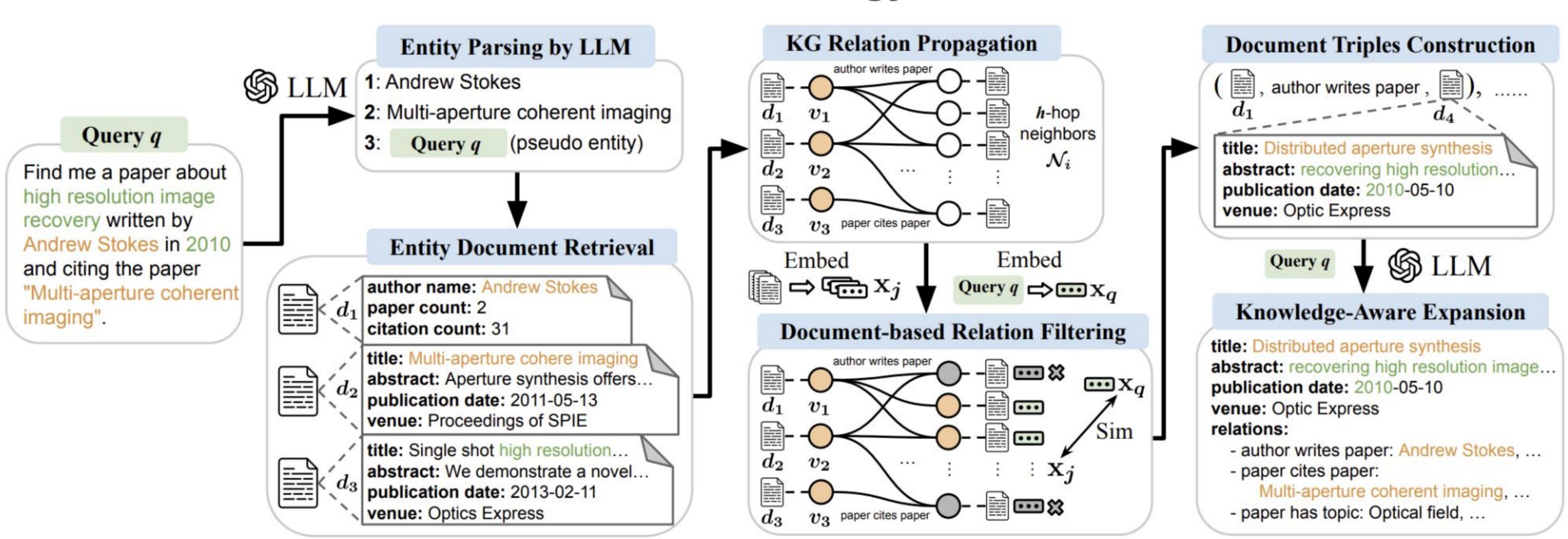


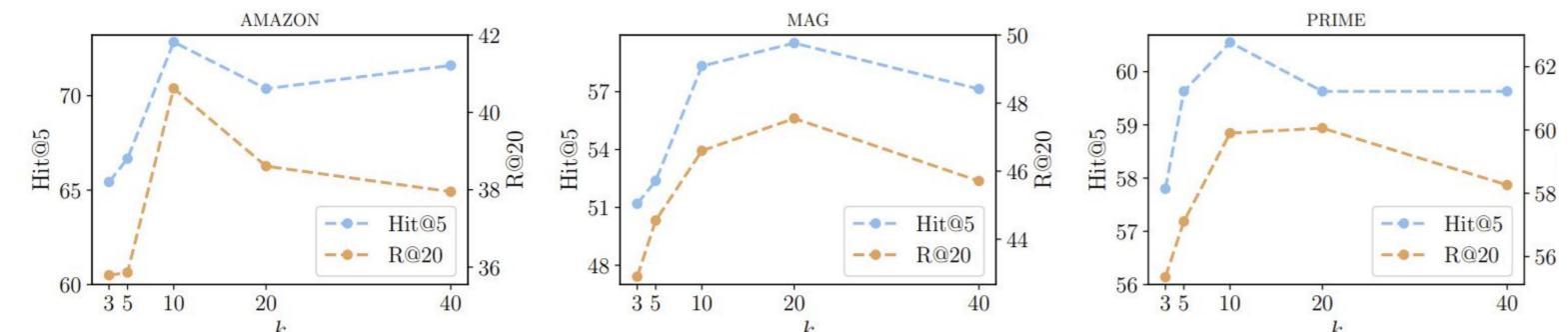
Figure 2: Overview of our knowledge-aware query expansion framework illustrated with an example academic paper search query with textual and relational requirements.

Experiments

- Three semi-structured retrieval datasets from STaRK benchmark.
- Table 2 is the results with
- GPT4o as LLM
- text-embedding-ada-002 as embedding.
- More evaluations and results with other LLMs and embeddings please refer to our paper.

	AMAZON				MAG				PRIME			
Method	Hit@1	Hit@5	R@20	MRR	Hit@1	Hit@5	R@20	MRR	Hit@1	Hit@5	R@20	MRR
Supervised Sea	ttings											
DPR	15.29	47.93	44.49	30.20	10.51	35.23	42.11	21.34	4.46	21.85	30.13	12.38
QAGNN	26.56	50.01	52.05	37.75	12.88	39.01	46.97	29.12	8.85	21.35	29.63	14.73
AvaTaR	49.87	69.16	60.57	58.70	44.36	59.66	50.63	51.15	18.44	36.73	39.31	26.73
Zero-Shot Sett	ings											
Base	39.16	62.73	53.29	50.35	29.08	49.61	48.36	38.62	12.63	31.49	36.00	21.41
PRF	40.07	60.66	51.24	49.79	29.04	47.65	46.69	37.90	12.46	28.63	33.04	20.06
HyDE	40.31	64.43	53.71	51.42	29.98	50.10	50.02	39.58	16.85	37.59	43.55	26.56
RAR	51.52	66.63	54.63	58.73	39.02	52.87	50.87	45.74	22.53	40.84	44.50	30.93
AGR	49.82	62.97	53.38	56.77	39.29	53.66	51.89	46.20	25.85	44.41	46.63	35.04
KAR _{w/o KG}	43.54	60.29	51.83	51.80	31.14	46.75	46.86	38.88	18.03	36.27	42.00	26.84
KAR _{w/o DRF}	47.99	67.54	56.91	57.14	45.44	63.83	58.67	53.85	25.85	46.52	48.10	35.52
KAR	54.20	68.70	57.24	61.29	50.47	65.37	60.28	57.51	30.35	49.30	50.81	39.22

Table 2: Retrieval results on test sets of synthetic search queries.



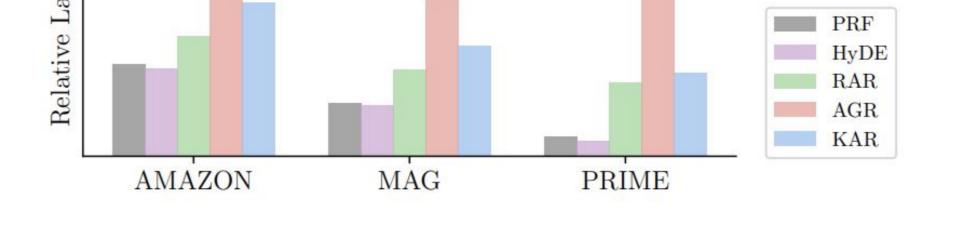


Figure 3: Influence of different values of k for filtered top-k neighbors in KAR.

Figure 5: Latency comparison of query expansions.