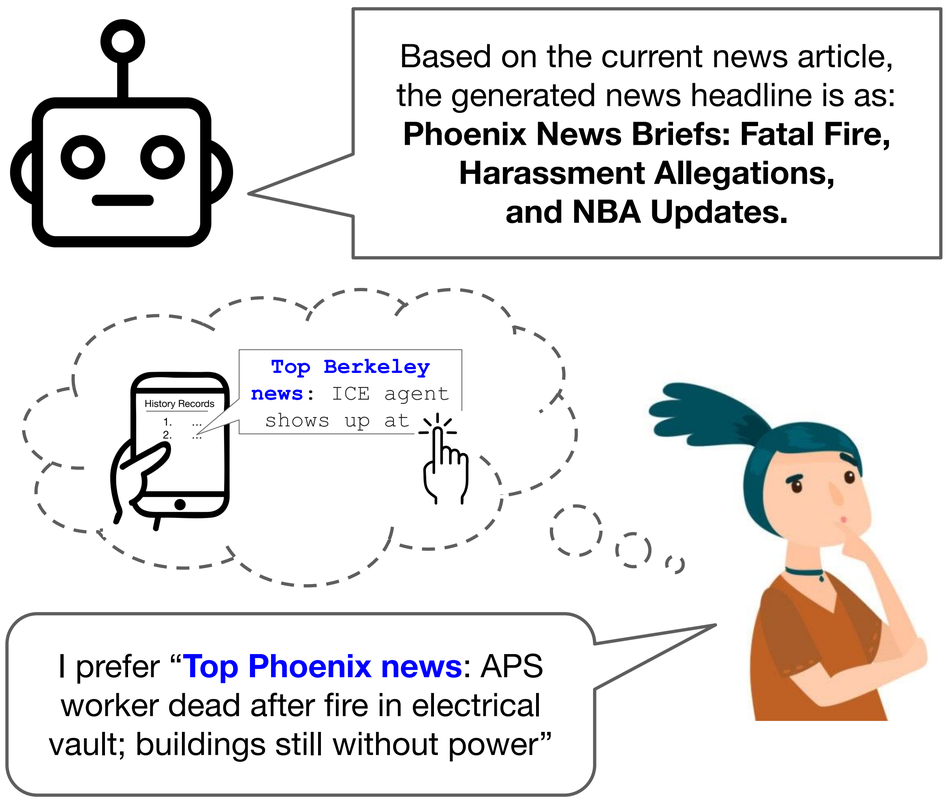
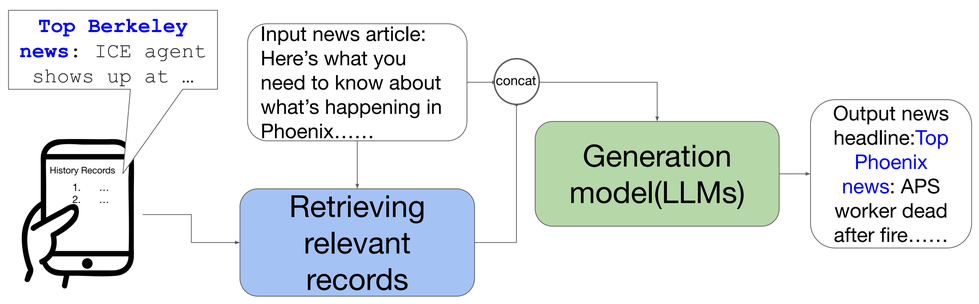
Personalizing News Headlines with Retrieval-Augmented Generation

Jiajing Wan, Samia Touileb, Lubos Steskal, Lilja Øvrelid, WP5.
UiB&TV2&UiO, jiajing.wan@uib.no.

Media Futures



Methology



We compare two types of retrievers in our RAG framework: a BM25 retriever, which performs lexical matching between the candidate article and the user's records, and a dense retriever, which captures semantic similarity based on embeddings. After retrieval, the records are re-ranked by relevance score, and the top-k results (k = 1~11) are concatenated with the current input to form the prompt for generation.

Abstract

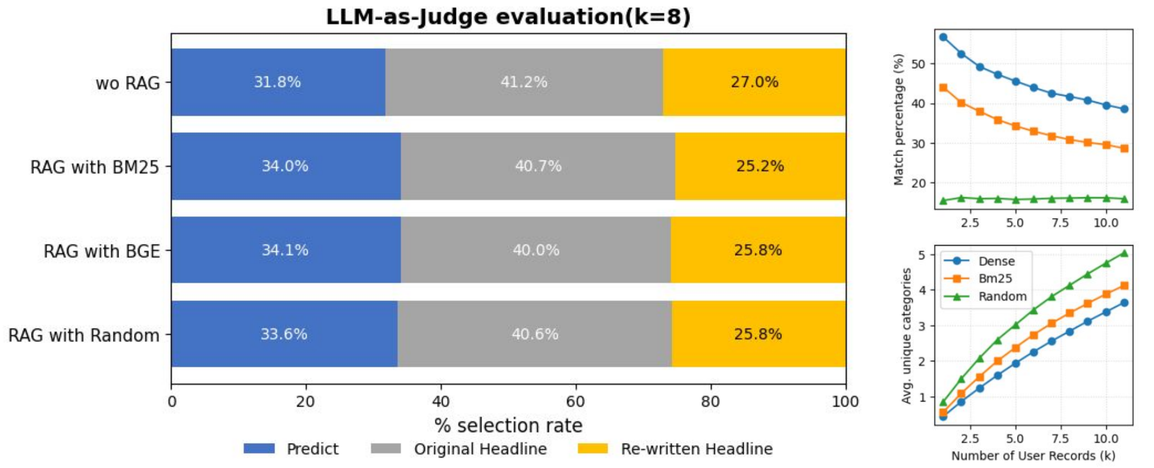
We propose a RAG-based system that takes into account the user history to improve news headline personalization. Our experiments reveal that to generate headlines that better align with user preferences, the retrieved user records that served as generation references should satisfy three conditions simultaneously: be of a certain quantity, have high relevance to the candidate news, and have adequate topical diversity.

Background

1. **How did we do Personalized news headline generation?**
   Generating a unified rhetorical style, or incorporating user records to make **customized personalization** of news headline generation.
2. **How could we do Personalized news headline generation with Large language models?**
   Retrieving the most relevant user records to the candidate news added it to the prompt as **a flexible user representation** while facing different needs of the candidate news.
   —> Retrieval-Augmented Generation!

Experiments and Results

	ROUGE-1			ROUGE-2			ROUGE-L			BLEU		
	bm25	dense	random	bm25	dense	random	bm25	dense	random	bm25	dense	random
w/o User records	31.18			12.63			26.00			34.48		
1 User records	30.15	30.74	30.66	11.90	12.14	12.12	25.39	25.84	25.88	33.55	32.27	34.04
2 User records	31.01	31.05	30.87	12.47	12.42	12.24	26.20	26.21	26.08	34.36	34.44	34.29
3 User records	31.29	31.13	30.99	12.68	12.53	12.36	26.43	26.32	26.19	34.58	34.52	34.32
4 User records	31.35	31.30	31.10	12.74	12.62	12.46	26.46	26.45	26.35	34.68	34.58	34.41
5 User records	31.38	31.41	31.11	12.74	12.73	12.46	26.48	26.54	26.33	34.72	34.73	34.50
6 User records	31.49	31.39	31.19	12.86	12.74	12.54	26.56	26.50	26.39	34.74	34.74	34.52
7 User records	31.50	31.46	31.23	12.90	12.77	12.56	26.52	26.56	26.42	34.80	34.76	34.57
8 User records	31.55	31.47	31.36	12.93	12.74	12.67	26.56	26.53	26.52	34.80	34.81	34.67
9 User records	**31.61**	31.49	31.32	12.99	12.81	12.71	**26.61**	26.57	26.47	**34.88**	34.81	34.69
10 User records	31.54	31.52	31.38	**13.00**	12.86	12.69	26.53	26.59	26.52	34.82	34.83	34.68
11 User records	31.45	31.50	31.46	12.91	12.84	12.81	26.47	26.56	26.59	34.78	34.80	34.75



As the number of user records increases, the performance of all RAG models improves. The BM25 retriever achieves a better balance between the relevant results and the diverse records.

PARTNERS

amedia Schibsted 2 VIZRT
NRK Bergens Tidende Faktisk.
nb.no Factiverse WOLFTECH
NORCE dn media group

HOST

UNIVERSITY OF BERGEN

FUNDED BY

This research is funded by SFI MediaFutures partners and the Research Council of Norway (grant number 309339).

Forskningsrådet
Norwegian Centre for Research-based Innovation