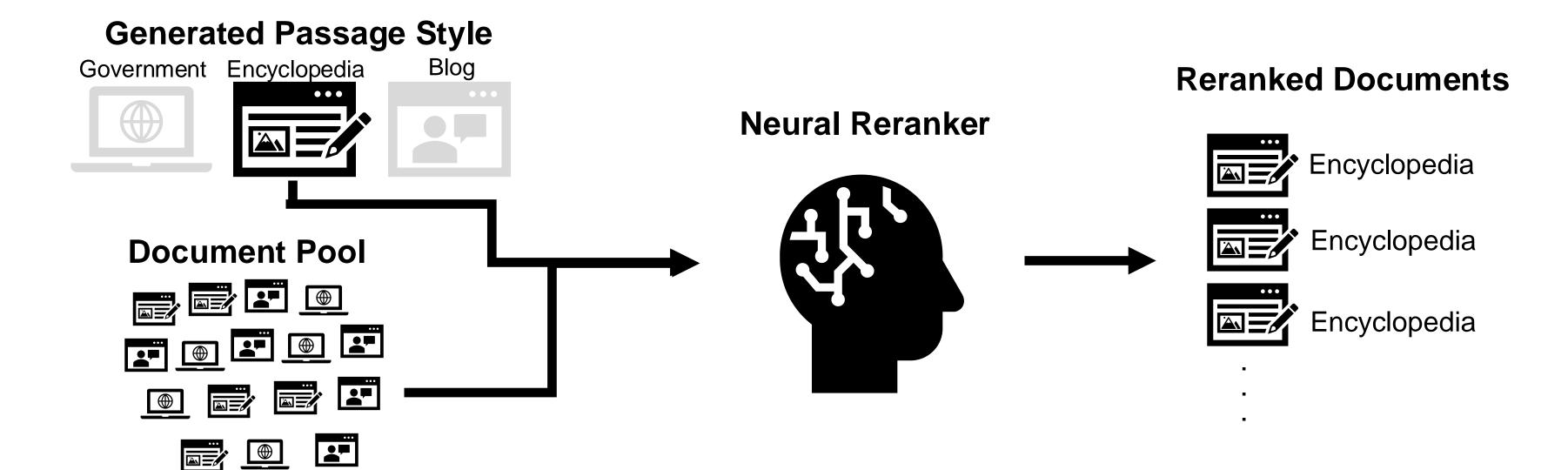# Does Query Tone and Structure Impact Neural Reranking Performance?

InfoSense Lab
Department of Computer Science
Victor De Lima (vad49@georgetown.edu)
Grace Hui Yang (grace.yang@georgetown.edu)

GEORGETOWN UNIVERSITY

Generated Passage Style

Government Encyclopedia Blog

Document Pool

Neural Reranker

Reranked Documents

Encyclopedia

Encyclopedia

Encyclopedia

Motivation and Background

- Neural reranking literature focuses on query keywords and exploring individual / ensemble neural models.
  - E.g. embeddings, scoring, fine-tuning.
- Query **tone** and **structure** impact still underexplored.
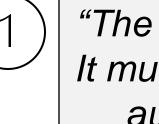- Analysis conducted within the TREC iKAT [1] context.

Experimental Setup

- Our system uses a Generate-then-Retrieve (GtR) Pipeline with Meta-Llama-3.1-8B-Instruct [2].
- Inputs are a processed conversation history and relevant user background information.
- Outputs are used for:
  - initial retrieval (BM25)
  - Neural reranking (all-MiniLM-L12-v2 [3], msmarco-distilbert-base-v4 [4]).
- Final responses generated from top-ranked documents.
- As part of the prompt to generate the query passages, we include one tone and style command.

Illustrative Example

- Consider turn 9-1_5 from the iKAT 2023: a user asks about incorporating fish into several diet options.

Llama Prompt Tone and Style Commands

① *"The writing style must be **Formal and Informative**. It must present information in a factual, concise, and authoritative manner, referencing sources and providing specific details. It uses precise language, making it suitable for educational or professional contexts."*

② *"The writing style must be **Informal and Conversational**. It must use a relaxed, approachable tone with casual language, directly addressing the reader. It aims to engage through rhetorical questions, practical advice, and a friendly, accessible approach."*

③ *"The writing style must be **Academic and Explanatory**. It must be detailed and structured, focusing on explaining the scientific rationale or principles behind a topic. It uses technical terms, provides context, and maintains a formal, educational tone, making it ideal for in-depth analysis or academic discussions."*

- The highest-ranked **target** documents have a **blog-like, conversational tone**.
- Their sources include:
  - *nutrition.org.uk, tofubud.com, keepingthepeas.com, thewholeportion.com*

Performance of turn 9-1_5 based on prompt added

	Formal and Informative	Informal and Conversational	Academic and Explanatory
nDCG@3	0.0000	**0.6257**	0.0000
nDCG@5	0.0000	**0.4522**	0.0000
nDCGP@20	0.1044	**0.3026**	0.1201
P@20	0.1500	**0.2000**	0.1500
Recall@20	0.0625	**0.0833**	0.0625
mAP	0.0334	**0.0874**	0.0446

Performance of conversation 9-1 based on prompt added

	Formal and Informative	Informal and Conversational	Academic and Explanatory
nDCG@3	0.1620	**0.3177**	0.1728
nDCG@5	0.1536	**0.2662**	0.1249
nDCGP@20	0.1359	**0.1867**	0.1317
P@20	0.1250	**0.1625**	0.1125
Recall@20	0.0416	**0.0467**	0.0330
mAP	0.0410	**0.0592**	0.0422

Results

- The **Informal and Conversational** command significantly improves reranking performance.
- This effect is widely observed across the dataset.

Challenges

- Difficulty in predicting question types.
- Non-determinism in outputs.
- Correctly understanding the user's query.
- Significant resource demands during experimentation.

References

1. Aliannejadi, M. et al. 2024. TREC iKAT 2023: The Interactive Knowledge Assistance Track Overview. (2024). DOI:https://doi.org/10.48550/ARXIV.2401.01330.
2. Meta Platforms, Inc. 2024. Meta-Llama-3.1-70B-Instruct. Hugging Face.
3. Sentence Transformers 2021. All-MiniLM-L12-v2. Hugging Face.
4. Sentence Transformers 2021. Msmarco-distilbert-base-v4. Hugging Face.