Contribution Title*

First Author[1][0000−1111−2222−3333], Second Author[2,3][1111−2222−3333−4444], and Third Author[3][2222−−3333−4444−5555]

[1] Princeton University, Princeton NJ 08544, USA
[2] Springer Heidelberg, Tiergartenstr. 17, 69121 Heidelberg, Germany
lncs@springer.com
http://www.springer.com/gp/computer-science/lncs
[3] ABC Institute, Rupert-Karls-University Heidelberg, Heidelberg, Germany
{abc,lncs}@uni-heidelberg.de

Abstract. The abstract should briefly summarize the contents of the paper in 150–250 words.

Keywords: First keyword · Second keyword · Another keyword.

1 First Section

Table 1. Arabic AI Models Studies in Retrieval-Augmented Generation (RAG)

Studies on Retrieval-Only Models						
Paper	**year**	**Retrieval Component Studied**	**Arabic Model Name**	**Datasets Used**	**Evaluation Metrics**	**Addressed Challenges**
Semantic Embeddings for Arabic Retrieval Augmented Generation (ARAG) [1]	2023	**Retrieval** :Semantic Embeddings	Microsoft(e5s,e5b,e5l) DistillBert(hf1,hf2), Openai Ada embedding Cohere Multilingual Embedding Meta SONAR Google LaBSE mpnet-base-v2	Arabic Reading Comprehension Dataset (ARCD)	Recall@k.	-Embedding Size Constraints -Need for Language Specific Evaluation Metrics
Evaluation of Semantic Search and its Role in Retrieved for Arabic Language[2]	2024	**Retrieval** :Semantic search in Arabic	Encoder 1: MiniLM Encoder 2: CMLM Encoder 3: MPNet Encoder 4: DistilBERT Encoder 5 : XLM RoBERTa	FAQs: 816 questions with verifiable answers	NDCG@3 MRR@3 mAP@3	-Embedding size constraints. -Arabic complexity
Arabic RAG Leaderboard: A Comprehensive Framework for Evaluating Arabic Language Retrieval Systems[3]	2025	**Retrieval**: Semantic Embedding . **Reranking**:Refine retrieved documents	Retrieval :GATE-AraBERT-v1 Reranking : ARA-Reranker-V1	Retrieval:"Web Search Dataset" Reranking:sourced from TyDi QA and MKQA datasets	NDCG MRR mAP Recall@k	Arabic's morphological complexity Dialect diversity
Studies on Both Retrieval and Generation						
Paper	**Year**	**RAG Component Studied**	**Arbic Model Name**	**Datasets Used**	**Evaluation Metricsl**	**Addressed Chalenges**
Exploring Retrieval Augmented Generation in Arabic [4]	2024 .	**Retrieval**: Semantic Embedding **Generation** : generate Arabic response	**Retrieval**: AraVec , AraBERT ,OpenAI ,Cohere, Microsoft's E5, Ollama ,JAIS,BGE **Generator**:GPT3.5, urbo, Mistral 7B, Llama 3, Mixtral, and JAIS.	Ar-EduText dataset, ARCD dataset	**Retrival**:Recall@K (k=1,k=3,k=5) **Generator**: F1 Score Bleu Score Cosine SimilarityR	-Lack of Detailed Metrics -Dialect Diversity
Evaluating RAG Pipelines for Arabic Lexical Information Retrieval: A Comparative Study of Embedding and Generation Models[5]	2025	**Retrieval**: Word Embeddings , Sentence Embeddings **Generator** : generatte Arabic lexical information	**Retrival**:CAMeLBERT, AraBERT-v2, E5-large Arabic-NLI ,AraELECTRA **Generator**: GPT-4o, Gemini-1.5-flash, SILMA-9B-Instruct, Aya8B GPT-3.5, AceGPT13B	Riyadh dictionary (88,000+ words)	**Retrival** Top-k Recall (k = 1, 3, 5) and MRR **Generator**:F1-score(F1), Accuracy (Acc), and CosineSimilarity(Cos).	-handling of dialectal variations in queries and documents. -Disparity in performance between sentence embeddings and word

* Supported by organization x.

References

1. Abdelazim, H., Mohamed, A., Tharwat, M.: Semantic Embeddings for Arabic Retrieval Augmented Generation (ARAG). International Journal of Advanced Computer Science and Applications **14**(11), xx–yy (2023)
2. Mahboub, A., Za'ter, M. E., Al-Rfooh, B., Estaitia, Y., Jaljuli, A., Hakouz, A.: Evaluation of Semantic Search and its Role in Retrieved-Augmented-Generation (RAG) for Arabic Language. In: arXiv preprint, arXiv:2403.18350v2, May 2024. Available at: https://arxiv.org/abs/2403.18350
3. Rashad, M.A., Shahid, H.: The Arabic RAG Leaderboard. *Navid-AI* (2025). Available online at: https://huggingface.co/spaces/Navid-AI/The-Arabic-Rag-Leaderboard.
4. El-Beltagy, S.R., Abdallah, M.A.: Exploring Retrieval Augmented Generation in Arabic. Procedia Computer Science, **00**, 000–000 (2024). Available online at: www.sciencedirect.com.
5. Al-Rasheed, R., Al Muaddi, A., Aljasim, H., Al-Matham, R., Alhoshan, M., Al Wazrah, A., AlOsaimy, A.: Evaluating RAG Pipelines for Arabic Lexical Information Retrieval: A Comparative Study of Embedding and Generation Models. In: El-Haj, M. (ed.), Proc. 1st Workshop on NLP for Languages Using Arabic Script, pp. 155–164. Association for Computational Linguistics, Abu Dhabi, UAE (2025). Available: https://aclanthology.org/2025.abjadnlp-1.16/
6. Author, A.-B.: Contribution title. In: 9th International Proceedings on Proceedings, pp. 1–2. Publisher, Location (2010)
7. LNCS Homepage, http://www.springer.com/lncs. Last accessed 4 Oct 2017