

REFERENCES

- [1] A. Szeto and A. N. Das, “Classification of notices to airmen using natural language processing,” in *AIAA SCITECH 2024 Forum*, 2024, p. 2585.
- [2] K. K. Patel, G. Desaulniers, A. Lodi, and F. L. and, “Explainable prediction of qcodes for notams using column generation,” *Journal of the Operational Research Society*, vol. 75, no. 1, pp. 173–183, 2024.
- [3] A. Arnold, F. Ernez, C. Kobus, and M.-C. Martin, “Knowledge extraction from aeronautical messages (NOTAMs) with self-supervised language models for aircraft pilots,” in *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies: Industry Track*. Association for Computational Linguistics, 2022, pp. 188–196.
- [4] L. Wang, J. Chou, A. Tien, X. Zhou, and D. Baumgartner, “Aviationgpt: A large language model for the aviation domain,” in *AIAA AVIATION FORUM AND ASCEND 2024*, 2024, p. 4250.
- [5] R. M. Hoefl, J. A. Kochan, and F. J. and, “A human factors analysis of the current u.s. notices to airmen (notam) system,” *The International Journal of Aviation Psychology*, vol. 15, no. 1, pp. 91–109, 2005.
- [6] A. Prokhorov, “Impact of notam on security and efficiency performance of flights (overview),” *Civil Aviation High Technologies*, vol. 25, no. 1, pp. 21–34, 2022.
- [7] R. Gu, Y. Qu, D. Chen, Y. Tang, and A. Zhou, “Mitigating information overload in aviation safety: Ai-driven hierarchical tagging and summarization of notam for pre-flight information bulletin,” *International Journal of Aviation, Aeronautics, and Aerospace*, vol. 12, no. 2, p. 5, 2025.
- [8] B. Neumayr, C. G. Schuetz, E. Gringinger, C. Fabianek, A. Vennesland, M. Schrefl, and S. Wilson, “Providing packages of relevant atm information: An ontology-based approach,” *Journal of Air Transport Management*, vol. 90, p. 101937, 2021.

- [9] P. Lewis, E. Perez, A. Piktus, F. Petroni, V. Karpukhin, N. Goyal, H. Küttler, M. Lewis, W.-t. Yih, T. Rocktäschel, S. Riedel, and D. Kiela, “Retrieval-augmented generation for knowledge-intensive nlp tasks,” in *Advances in Neural Information Processing Systems*, vol. 33. Curran Associates, Inc., 2020, pp. 9459–9474.
- [10] J. Wu, J. Zhu, Y. Qi, J. Chen, M. Xu, F. Menolascina, and V. Grau, “Medical graph rag: Towards safe medical large language model via graph retrieval-augmented generation,” *arXiv preprint arXiv:2408.04187*, 2024.
- [11] J. Wang, Z. Yang, Z. Yao, and H. Yu, “Jmlr: Joint medical llm and retrieval training for enhancing reasoning and professional question answering capability,” *arXiv preprint arXiv:2402.17887*, 2024.
- [12] J. He, B. Zhang, H. Rouhizadeh, Y. Chen, R. Yang, J. Lu, X. Chen, N. Liu, I. Li, and D. Teodoro, “Retrieval-augmented generation in biomedicine: A survey of technologies, datasets, and clinical applications,” *arXiv preprint arXiv:2505.01146*, 2025.
- [13] C. Choi, J. Kwon, J. Ha, H. Choi, C. Kim, Y. Lee, J.-y. Sohn, and A. Lopez-Lira, “Finder: Financial dataset for question answering and evaluating retrieval-augmented generation,” *arXiv preprint arXiv:2504.15800*, 2025.
- [14] X. Wang, J. Chi, Z. Tai, T. S. T. Kwok, M. Li, Z. Li, H. He, Y. Hua, P. Lu, S. Wang *et al.*, “Finsage: A multi-aspect rag system for financial filings question answering,” *arXiv preprint arXiv:2504.14493*, 2025.
- [15] J. Choe, J. Kim, and W. Jung, “Hierarchical retrieval with evidence curation for open-domain financial question answering on standardized documents,” *arXiv preprint arXiv:2505.20368*, 2025.
- [16] C. Gondhalekar, U. Patel, and F.-C. Yeh, “Multifinrag: An optimized multimodal retrieval-augmented generation (rag) framework for financial question answering,” 2025.
- [17] X. Xie, X. Tang, S. Gu, and L. Cui, “An intelligent guided troubleshooting method for aircraft based on hybridrag,” *Scientific Reports*, vol. 15, no. 1, pp. 1–16, 2025.
- [18] J. Ni, C. Qu, J. Lu, Z. Dai, G. Hernandez Abrego, J. Ma, V. Zhao, Y. Luan, K. Hall, M.-W. Chang, and Y. Yang, “Large dual encoders are generalizable retrievers,” in *Proceedings of the 2022 Conference on Empirical Methods in*

Natural Language Processing. Abu Dhabi, United Arab Emirates: Association for Computational Linguistics, 2022, pp. 9844–9855.

- [19] N. Reimers and I. Gurevych, “Sentence-BERT: Sentence embeddings using Siamese BERT-networks,” in *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*. Association for Computational Linguistics, 2019, pp. 3982–3992.
- [20] N. Thakur, N. Reimers, A. Ruckl’e, A. Srivastava, and I. Gurevych, “Beir: A heterogenous benchmark for zero-shot evaluation of information retrieval models,” *ArXiv*, vol. abs/2104.08663, 2021.
- [21] N. Muennighoff, N. Tazi, L. Magne, and N. Reimers, “MTEB: Massive text embedding benchmark,” in *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics*. Association for Computational Linguistics, 2023, pp. 2014–2037.
- [22] Y. Ji, Z. Xu, Z. Liu, Y. Yan, S. Yu, Y. Li, Z. Liu, Y. Gu, G. Yu, and M. Sun, “Learning more effective representations for dense retrieval through deliberate thinking before search,” *arXiv preprint arXiv:2502.12974*, 2025.
- [23] A. Conneau and G. Lample, *Cross-lingual language model pretraining*. Curran Associates Inc., 2019.
- [24] J. Chen, S. Xiao, P. Zhang, K. Luo, D. Lian, and Z. Liu, “M3-embedding: Multi-linguality, multi-functionality, multi-granularity text embeddings through self-knowledge distillation,” in *Findings of the Association for Computational Linguistics: ACL 2024*, L.-W. Ku, A. Martins, and V. Srikumar, Eds. Association for Computational Linguistics, 2024, pp. 2318–2335.
- [25] X. Zhang, Y. Zhang, D. Long, W. Xie, Z. Dai, J. Tang, H. Lin, B. Yang, P. Xie, F. Huang, M. Zhang, W. Li, and M. Zhang, “mGTE: Generalized long-context text representation and reranking models for multilingual text retrieval,” in *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing: Industry Track*. Association for Computational Linguistics, 2024, pp. 1393–1412.
- [26] P. Yu, L. Merrick, G. Nuti, and D. Campos, “Arctic-embed 2.0: Multilingual retrieval without compromise,” *arXiv preprint arXiv:2412.04506*, 2024.

- [27] J. S. Rayo Mosquera, C. R. De La Rosa Peredo, and M. Garrido Cordoba, “A hybrid approach to information retrieval and answer generation for regulatory texts,” in *Proceedings of the 1st Regulatory NLP Workshop (RegNLP 2025)*. Abu Dhabi, UAE: Association for Computational Linguistics, 2025, pp. 31–35.
- [28] K. Song, X. Tan, T. Qin, J. Lu, and T.-Y. Liu, “Mpnet: masked and permuted pre-training for language understanding,” in *Proceedings of the 34th International Conference on Neural Information Processing Systems*, ser. NIPS ’20. Curran Associates Inc., 2020.
- [29] T. Gao, X. Yao, and D. Chen, “SimCSE: Simple contrastive learning of sentence embeddings,” in *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*. Online and Punta Cana, Dominican Republic: Association for Computational Linguistics, 2021, pp. 6894–6910.
- [30] G. Izacard, M. Caron, L. Hosseini, S. Riedel, P. Bojanowski, A. Joulin, and E. Grave, “Unsupervised dense information retrieval with contrastive learning,” *arXiv preprint arXiv:2112.09118*, 2021.
- [31] N. Muennighoff, “Sgpt: Gpt sentence embeddings for semantic search,” *arXiv preprint arXiv:2202.08904*, 2022.
- [32] A. Kusupati, G. Bhatt, A. Rege, M. Wallingford, A. Sinha, V. Ramanujan, W. Howard-Snyder, K. Chen, S. Kakade, P. Jain *et al.*, “Matryoshka representation learning,” *Advances in Neural Information Processing Systems*, vol. 35, pp. 30 233–30 249, 2022.
- [33] A. Conneau, K. Khandelwal, N. Goyal, V. Chaudhary, G. Wenzek, F. Guzmán, E. Grave, M. Ott, L. Zettlemoyer, and V. Stoyanov, “Unsupervised cross-lingual representation learning at scale,” in *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics, 2020, pp. 8440–8451.
- [34] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, “Attention is all you need,” *Advances in neural information processing systems*, vol. 30, 2017.
- [35] A. Radford and K. Narasimhan, “Improving language understanding by generative pre-training,” 2018.
- [36] A. Radford, J. Wu, R. Child, D. Luan, D. Amodei, and I. Sutskever, “Language models are unsupervised multitask learners,” 2019.

- [37] T. Brown, B. Mann, N. Ryder, M. Subbiah, J. D. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell *et al.*, “Language models are few-shot learners,” in *Advances in Neural Information Processing Systems*, H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan, and H. Lin, Eds., vol. 33. Curran Associates, Inc., 2020, pp. 1877–1901.
- [38] J. Achiam, S. Adler, S. Agarwal, L. Ahmad, I. Akkaya, F. L. Aleman, D. Almeida, J. Altenschmidt, S. Altman, S. Anadkat *et al.*, “Gpt-4 technical report,” *arXiv preprint arXiv:2303.08774*, 2023.
- [39] W. X. Zhao, J. Liu, R. Ren, and J.-R. Wen, “Dense text retrieval based on pretrained language models: A survey,” *ACM Trans. Inf. Syst.*, vol. 42, no. 4, 2024.
- [40] Y. Wang, L. Wang, Y. Li, D. He, and T.-Y. Liu, “A theoretical analysis of ndcg type ranking measures,” in *Proceedings of the 26th Annual Conference on Learning Theory*, ser. Proceedings of Machine Learning Research, S. Shalev-Shwartz and I. Steinwart, Eds., vol. 30. Princeton, NJ, USA: PMLR, 12–14 Jun 2013, pp. 25–54.