Pre-training Methods in Information Retrieval

Other titles in Foundations and Trends® in Information Retrieval

Fairness in Information Access Systems
Michael D. Ekstrand, Anubrata Das, Robin Burke and Fernando Diaz
ISBN: 978-1-63828-040-8

Deep Learning for Dialogue Systems: Chit-Chat and Beyond
Rui Yan, Juntao Li and Zhou Yu
ISBN: 978-1-63828-022-4

Search Interface Design and Evaluation
Chang Liu, Ying-Hsang Liu, Jingjing Liu and Ralf Bierig
ISBN: 978-1-68083-922-7

Psychology-informed Recommender Systems
Elisabeth Lex, Dominik Kowald, Paul Seitlinger, Thi Ngoc Trang Tran, Alexander Felfernig and Markus Schedl
ISBN: 978-1-68083-844-2

Search and Discovery in Personal Email Collections
Michael Bendersky, Xuanhui Wang, Marc Najork and Donald Metzler
ISBN: 978-1-68083-838-1

Pre-training Methods in Information Retrieval

Yixing Fan
ICT, CAS
fanyixing@ict.ac.cn

Xiaohui Xie
Tsinghua University
xiexiaohui@mail.tsinghua.edu.cn

Yinqiong Cai
ICT, CAS
caiyinqiong18s@ict.ac.cn

Jia Chen
Tsinghua University
chenjia0831@gmail.com

Xinyu Ma
ICT, CAS
maxinyu17g@ict.ac.cn

Xiangsheng Li
Tsinghua University
lixsh6@gmail.com

Ruqing Zhang
ICT, CAS
zhangruqing@ict.ac.cn

Jiafeng Guo
ICT, CAS
guojiafeng@ict.ac.cn

now
the essence of knowledge
Boston — Delft

Foundations and Trends® in Information Retrieval

Published, sold and distributed by:
now Publishers Inc.
PO Box 1024
Hanover, MA 02339
United States
Tel. +1-781-985-4510
www.nowpublishers.com
sales@nowpublishers.com

Outside North America:
now Publishers Inc.
PO Box 179
2600 AD Delft
The Netherlands
Tel. +31-6-51115274

The preferred citation for this publication is

Y. Fan *et al.*. *Pre-training Methods in Information Retrieval*. Foundations and Trends® in Information Retrieval, vol. 16, no. 3, pp. 178–317, 2022.

ISBN: 978-1-63828-063-7

Foundations and Trends® in Information Retrieval
Volume 16, Issue 3, 2022

Editorial Scope

Topics

Foundations and Trends® in Information Retrieval publishes survey and tutorial articles in the following topics:

- Applications of IR
- Architectures for IR
- Collaborative filtering and recommender systems
- Cross-lingual and multilingual IR
- Distributed IR and federated search
- Evaluation issues and test collections for IR
- Formal models and language models for IR
- IR on mobile platforms
- Indexing and retrieval of structured documents
- Information categorization and clustering
- Information extraction
- Information filtering and routing
- Metasearch, rank aggregation and data fusion
- Natural language processing for IR
- Performance issues for IR systems, including algorithms, data structures, optimization techniques, and scalability
- Question answering
- Summarization of single documents, multiple documents, and corpora
- Text mining
- Topic detection and tracking
- Usability, interactivity, and visualization issues in IR
- User modelling and user studies for IR
- Web search

Information for Librarians

Foundations and Trends® in Information Retrieval, 2022, Volume 16, 5 issues. ISSN paper version 1554-0669. ISSN online version 1554-0677. Also available as a combined paper and online subscription.

Contents

Pre-training Methods in Information Retrieval

Yixing Fan*[1], Xiaohui Xie*[2], Yinqiong Cai[1], Jia Chen[2], Xinyu Ma[1], Xiangsheng Li[2], Ruqing Zhang[1] and Jiafeng Guo[1]

[1] *ICT, CAS, China; fanyixing@ict.ac.cn, caiyinqiong18s@ict.ac.cn, maxinyu17g@ict.ac.cn, zhangruqing@ict.ac.cn, guojiafeng@ict.ac.cn*
[2] *Tsinghua University, China; xiexiaohui@mail.tsinghua.edu.cn, chenjia0831@gmail.com, lixsh6@gmail.com*

ABSTRACT

The core of information retrieval (IR) is to identify relevant information from large-scale resources and return it as a ranked list to respond to user's information need. In recent years, the resurgence of deep learning has greatly advanced this field and leads to a hot topic named NeuIR (i.e., neural information retrieval), especially the paradigm of pre-training methods (PTMs). Owing to sophisticated pre-training objectives and huge model size, pre-trained models can learn universal language representations from massive textual data, which are beneficial to the ranking task of IR. Recently, a large number of works, which are dedicated to the application of PTMs in IR, have been introduced to promote the retrieval performance. Considering the rapid progress of this direction, this survey aims to provide

*Yixing Fan and Xiaohui Xie contributed equally to this survey. Jiafeng Guo is the corresponding author.

Yixing Fan, Xiaohui Xie, Yinqiong Cai, Jia Chen, Xinyu Ma, Xiangsheng Li, Ruqing Zhang and Jiafeng Guo (2022), "Pre-training Methods in Information Retrieval", Foundations and Trends® in Information Retrieval: Vol. 16, No. 3, pp 178–317. DOI: 10.1561/1500000100.

a systematic review of pre-training methods in IR. To be specific, we present an overview of PTMs applied in different components of an IR system, including the retrieval component, the re-ranking component, and other components. In addition, we also introduce PTMs specifically designed for IR, and summarize available datasets as well as benchmark leaderboards. Moreover, we discuss some open challenges and highlight several promising directions, with the hope of inspiring and facilitating more works on these topics for future research.

1

Introduction

Information retrieval (IR) is a fundamental task in many real-world applications, such as Web search, question answering systems, digital libraries, and so on. The core of IR is to identify information resources relevant to user's information need (e.g., query or question) from a large collection. Since there might be more than one relevant resource, the returned result is often organized as a ranked list of documents (e.g., Web pages, answers, or responses) according to their relevance degree against the information need. Such ranking property of IR makes it different from other tasks, and researchers have devoted substantial efforts to develop a variety of ranking models in IR.

Over the past decades, many different ranking models have been introduced and studied, including vector space models (Salton *et al.*, 1975), probabilistic ranking models (Robertson and Jones, 1976), and learning to rank (LTR) models (Li, 2014). These methods have been successfully applied in many different IR applications, such as Web search engines like Google, news recommender systems like Toutiao, community question answering platforms like Quora, to name a few. More recently, a large variety of neural ranking models have been proposed, leading to a hot topic named NeuIR (Craswell *et al.*, 2017) (i.e., neural information retrieval). Different from previous non-neural ranking models that rely on elaborately-designed features and manually-designed functions, neural ranking models can automatically learn low-level dense representations from data as ranking features. Despite the success of neural models in IR, a major performance bottleneck lies in the availability of large scale, high-quality and labeled datasets as deep neural models often have a large number of parameters to learn (Dehghani *et al.*, 2017b).

In recent years, PTMs have brought a storm and fueled a paradigm shift in Nature Language Processing (NLP) (Qiu *et al.*, 2020). The idea is to firstly pre-train models with self-supervised language modeling, e.g., predicting the probability of a masked token, and then adapt the pre-trained model to downstream tasks by introducing a small number of additional parameters and fine-tuning them with some task-specific objectives. As is demonstrated in recent works (Peters *et al.*, 2018; Howard and Ruder, 2018), these pre-trained models are able to capture a decent amount of linguistic knowledge as well as factual knowledge, which are beneficial for downstream tasks and can avoid learning such knowledge from scratch. Moreover, with the increasing amount of computational power and the emergence of the Transformer architecture (Vaswani *et al.*, 2017), we can further improve the capacity of pre-trained models by updating the parameter scale, e.g., from million-level to billion-level (e.g., BERT (Devlin *et al.*, 2019) and GPT-3 (Brown *et al.*, 2020)) and even trillion-level (e.g., Switch-Transformers (Fedus *et al.*, 2021)). Both of these are desirable properties for modeling the relevance in IR.

On one hand, pre-trained embeddings, which are learned on huge textual corpus with self-supervised modeling objectives, are able to capture intrinsic semantics inside queries and documents. On the other hand, large-scale pre-trained models with deeply stacked Transformers have sufficient modeling capacities to learn complicated relevance patterns between queries and documents. Owing to these potential benefits, we have witnessed explosive growth of research interest in exploiting PTMs in IR (Onal *et al.*, 2017; Lin *et al.*, 2021a). Note that in this survey, we focus on PTMs in text retrieval, which is central to IR. Readers who are interested in PTMs in content-based image retrieval or multi-modal retrieval could refer to Dubey (2020) and Fei *et al.* (2021).

Up to now, numerous studies have been devoted to the application of PTMs in IR. In academia, researchers have carried out a variety of innovation and initiative in the usage of PTMs in IR. For example, earlier attempts tried to leverage pre-trained word embeddings to promote ranking models, and have achieved some notable results (Onal *et al.*, 2017). More recent works proposed improving existing pre-trained models by either reforming the model architecture (MacAvaney *et al.*, 2020; Khattab and Zaharia, 2020; Gao and Callan, 2021a) or considering novel pre-training objectives (Chang *et al.*, 2020; Ma *et al.*, 2021b; Ma *et al.*, 2021c), which can better meet the requirements of IR. Meanwhile, in industry, Google's October 2019 blog post[1] and Bing's November 2019 blog post[2] both showed that pre-trained ranking models (e.g., BERT-based models) can better understand the query intent and deliver a more useful result in practical search systems. Besides, looking at the ranking leaderboard[3] today, we can see that most top-ranked methods are built on PTMs, just by looking at the names of these submissions. Considering the increasing number of studies on PTMs in IR, we believe that it is the right time to survey the current literature, highlight advantages and limitations of existing methods, and gain some insights for future development.

In this survey, we aim to provide a systematic and comprehensive review of works about PTMs in IR. It covers PTMs published in major conferences (e.g., SIGIR, TheWebConf, ICLR, WSDM, CIKM, AAAI, ACL, and ECIR) and journals (e.g., TOIS, TKDE, TIST, IP&M, and TACL) in the fields of deep learning, natural language processing, and information retrieval from the year 2016 to 2021. There exists some previous works discussing related topics. For example, both Onal *et al.* (2017) and Guo *et al.* (2020) reviewed the landscape of neural retrieval models used in three major IR tasks. They also discussed early usage of pre-trained embeddings in neural ranking models, but did not cover

[1] https://www.blog.google/products/search/search-language-understanding-bert/ Date accessed: 23 June 2022.

[2] https://azure.microsoft.com/en-us/blog/bing-delivers-its-largest-improvement-\in-search-experience-using-azure-gpus/ Date accessed: 23 June 2022.

[3] https://microsoft.github.io/msmarco/#docranking Date accessed: 23 June 2022.

every aspect of PTMs in IR. Guo *et al.* (2022) reviewed semantic models for the first-stage retrieval, including early semantic retrieval models, neural retrieval models, and retrieval models based on PTMs. More recently, Lin *et al.* (2021a) provided a thorough survey of transformer-based models for IR, which reviews existing literature on the application of pre-trained contextual embedding in text ranking. Different from these works, we make a comprehensive overview of PTMs applied in IR, including the usage of pre-trained word embeddings as well as the application of pre-trained transformers. More specifically, we review the application of PTMs in different components of an IR system, including the first-stage retrieval component, the re-ranking component, and other components. We also describe PTMs specifically designed for IR tasks, as well as resources for pre-training or fine-tuning ranking models. In addition to the model discussion, we also introduce some open challenges and suggest potential research directions for future works.

The structure of this survey is organized as follows. We will firstly provide a systematic overview of IR in Section 2. Following this, we then review works about PTMs applied in the retrieval component, the re-ranking component, and other components in Sections 3 to 5, respectively. In Section 6, we present works in designing novel PTMs tailored for IR. We also summarize available large-scale datasets as well as popular benchmark leaderboards in Section 7. Finally, we conclude this survey in Section 8 and raise some promising directions for future research.

References

Agirre, E., X. Arregi, and A. Otegi. (2010). "Document Expansion Based on WordNet for Robust IR". In: *COLING 2010, 23rd International Conference on Computational Linguistics, Posters Volume, 23-27 August 2010, Beijing, China*. Ed. by C. Huang and D. Jurafsky. Chinese Information Processing Society of China. 9–17. URL: https://aclanthology.org/C10-2002/.

Agosti, M., S. Marchesin, and G. Silvello. (2020). "Learning Unsupervised Knowledge-Enhanced Representations to Reduce the Semantic Gap in Information Retrieval". *ACM Transactions on Information Systems*. 38(4): 1–48. DOI: 10.1145/3417996.

Ai, Q., L. Yang, J. Guo, and W. B. Croft. (2016a). "Analysis of the Paragraph Vector Model for Information Retrieval". In: *Proceedings of the 2016 ACM International Conference on the Theory of Information Retrieval*. ACM. DOI: 10.1145/2970398.2970409.

Ai, Q., L. Yang, J. Guo, and W. B. Croft. (2016b). "Improving Language Estimation with the Paragraph Vector Model for Ad-hoc Retrieval". In: *Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval*. ACM. DOI: 10.1145/2911451.2914688.

Alberti, C., J. Ling, M. Collins, and D. Reitter. (2019). "Fusion of Detected Objects in Text for Visual Question Answering". 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. DOI: 10.18653/v1/d19-1219.

Aliannejadi, M., H. Zamani, F. Crestani, and W. B. Croft. (2019). "Asking Clarifying Questions in Open-Domain Information-Seeking Conversations". In: *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM. DOI: 10.1145/3331184.3331265.

Amati, G. and C. J. V. Rijsbergen. (2002). "Probabilistic models of information retrieval based on measuring the divergence from randomness". *ACM Transactions on Information Systems*. 20(4): 357–389. DOI: 10.1145/582415.582416.

Amer, N. O., P. Mulhem, and M. Géry. (2016). "Toward Word Embedding for Personalized Information Retrieval". *CoRR*. abs/1606.06991. arXiv: 1606.06991.

Antol, S., A. Agrawal, J. Lu, M. Mitchell, D. Batra, C. L. Zitnick, and D. Parikh. (2015). "VQA: Visual Question Answering". In: *2015 IEEE International Conference on Computer Vision (ICCV)*. IEEE. DOI: 10.1109/iccv.2015.279.

Aumüller, M., E. Bernhardsson, and A. Faithfull. (2020). "ANN-Benchmarks: A benchmarking tool for approximate nearest neighbor algorithms". *Information Systems*. 87: 101374. DOI: 10.1016/j.is.2019.02.006.

Bae, S., T. Kim, J. Kim, and S. Lee. (2019). "Summary Level Training of Sentence Rewriting for Abstractive Summarization". *CoRR*. abs/1909.08752. arXiv: 1909.08752.

Bai, Y., X. Li, G. Wang, C. Zhang, L. Shang, J. Xu, Z. Wang, F. Wang, and Q. Liu. (2020). "SparTerm: Learning Term-based Sparse Representation for Fast Text Retrieval". *CoRR*. abs/2010.00768. arXiv: 2010.00768.

Bao, H., L. Dong, F. Wei, W. Wang, N. Yang, X. Liu, Y. Wang, J. Gao, S. Piao, M. Zhou, and H. Hon. (2020). "UniLMv2: Pseudo-Masked Language Models for Unified Language Model Pre-Training". In: *Proceedings of the 37th International Conference on Machine Learning, ICML 2020, 13-18 July 2020, Virtual Event*. Vol. 119. *Proceedings of Machine Learning Research*. PMLR. 642–652. URL: http://proceedings.mlr.press/v119/bao20a.html.

Baumel, T., M. Eyal, and M. Elhadad. (2018). "Query Focused Abstractive Summarization: Incorporating Query Relevance, Multi-Document Coverage, and Summary Length Constraints into seq2seq Models". *CoRR*. abs/1801.07704. arXiv: 1801.07704.

Beltagy, I., K. Lo, and A. Cohan. (2019). "SciBERT: A Pretrained Language Model for Scientific Text". 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. DOI: 10.18653/v1/d19-1371.

Beltagy, I., M. E. Peters, and A. Cohan. (2020). "Longformer: The Long-Document Transformer". *CoRR*. abs/2004.05150. arXiv: 2004.05150.

Bengio, Y., A. Courville, and P. Vincent. (2013). "Representation Learning: A Review and New Perspectives". *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 35(8): 1798–1828. DOI: 10.1109/tpami.2013.50.

Bi, K., Q. Ai, and W. B. Croft. (2020). "A Transformer-based Embedding Model for Personalized Product Search". In: *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM. DOI: 10.1145/3397271.3401192.

Bi, K., Q. Ai, and W. B. Croft. (2021a). "Asking Clarifying Questions Based on Negative Feedback in Conversational Search". In: *Proceedings of the 2021 ACM SIGIR International Conference on Theory of Information Retrieval*. ACM. DOI: 10.1145/3471158.3472232.

Bi, K., Q. Ai, and W. B. Croft. (2021b). "Learning a Fine-Grained Review-based Transformer Model for Personalized Product Search". In: *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM. DOI: 10.1145/3404835.3462911.

Bi, K., P. Metrikov, C. Li, and B. Byun. (2021c). "Leveraging User Behavior History for Personalized Email Search". In: *Proceedings of the Web Conference 2021*. ACM. DOI: 10.1145/3442381.3450110.

Bojanowski, P., E. Grave, A. Joulin, and T. Mikolov. (2017). "Enriching Word Vectors with Subword Information". *Trans. Assoc. Comput. Linguistics*. 5: 135–146. URL: https://transacl.org/ojs/index.php/tacl/article/view/999.

Bromley, J., I. Guyon, Y. LeCun, E. Säckinger, and R. Shah. (1993). "Signature Verification Using a Siamese Time Delay Neural Network". In: *Advances in Neural Information Processing Systems 6, [7th NIPS Conference, Denver, Colorado, USA, 1993]*. Ed. by J. D. Cowan, G. Tesauro, and J. Alspector. Morgan Kaufmann. 737–744. URL: http://papers.nips.cc/paper/769-signature-verification-using-a-siamese-time-delay-neural-network.

Brown, T. B., B. Mann, N. Ryder, M. Subbiah, J. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell, S. Agarwal, A. Herbert-Voss, G. Krueger, T. Henighan, R. Child, A. Ramesh, D. M. Ziegler, J. Wu, C. Winter, C. Hesse, M. Chen, E. Sigler, M. Litwin, S. Gray, B. Chess, J. Clark, C. Berner, S. McCandlish, A. Radford, I. Sutskever, and D. Amodei. (2020). "Language Models are Few-Shot Learners". In: *Advances in Neural Information Processing Systems 33: Annual Conference on Neural Information Processing Systems 2020, NeurIPS 2020, December 6-12, 2020, virtual*. Ed. by H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan, and H. Lin. URL: https://proceedings.neurips.cc/paper/2020/hash/1457c0d6bfcb4967418bfb8ac142f64a-Abstract.html.

Burges, C., T. Shaked, E. Renshaw, A. Lazier, M. Deeds, N. Hamilton, and G. Hullender. (2005). "Learning to rank using gradient descent". In: *Proceedings of the 22nd international conference on Machine learning - ICML '05*. ACM Press. DOI: 10.1145/1102351.1102363.

Burges, C. J. C., R. Ragno, and Q. V. Le. (2006). "Learning to Rank with Nonsmooth Cost Functions". In: *Advances in Neural Information Processing Systems 19, Proceedings of the Twentieth Annual Conference on Neural Information Processing Systems, Vancouver, British Columbia, Canada, December 4-7, 2006*. Ed. by B. Schölkopf, J. C. Platt, and T. Hofmann. MIT Press. 193–200. URL: https://proceedings.neurips.cc/paper/2006/hash/af44c4c56f385c43f2529f9b1b018f6a-Abstract.html.

Cao, J., Z. Gan, Y. Cheng, L. Yu, Y.-C. Chen, and J. Liu. (2020). "Behind the Scene: Revealing the Secrets of Pre-trained Vision-and-Language Models". In: *Computer Vision – ECCV 2020*. Springer International Publishing. 565–580. DOI: 10.1007/978-3-030-58539-6_34.

Carbinell, J. and J. Goldstein. (2017). "The Use of MMR, Diversity-Based Reranking for Reordering Documents and Producing Summaries". *ACM SIGIR Forum*. 51(2): 209–210. DOI: 10.1145/3130348.3130369.

Chang, W., F. X. Yu, Y. Chang, Y. Yang, and S. Kumar. (2020). "Pre-training Tasks for Embedding-based Large-scale Retrieval". In: *8th International Conference on Learning Representations, ICLR 2020, Addis Ababa, Ethiopia, April 26-30, 2020*. OpenReview.net. URL: https://openreview.net/forum?id=rkg-mA4FDr.

Chen, J., Y. Liu, Y. Fang, J. Mao, H. Fang, S. Yang, X. Xie, M. Zhang, and S. Ma. (2022). "Axiomatically Regularized Pre-training for Ad hoc Search". In: *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*.

Chen, J., J. Mao, Y. Liu, Z. Ye, W. Ma, C. Wang, M. Zhang, and S. Ma. (2021a). "A Hybrid Framework for Session Context Modeling". *ACM Transactions on Information Systems*. 39(3): 1–35. DOI: 10.1145/3448127.

Chen, R.-C. and C.-J. Lee. (2020). "Incorporating Behavioral Hypotheses for Query Generation". In: *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Association for Computational Linguistics. DOI: 10.18653/v1/2020.emnlp-main.251.

Chen, W.-F., S. Syed, B. Stein, M. Hagen, and M. Potthast. (2020a). "Abstractive Snippet Generation". In: *Proceedings of The Web Conference 2020*. ACM. DOI: 10.1145/3366423.3380206.

Chen, X., N. Zhang, X. Xie, S. Deng, Y. Yao, C. Tan, F. Huang, L. Si, and H. Chen. (2021b). "KnowPrompt: Knowledge-aware Prompt-tuning with Synergistic Optimization for Relation Extraction". *CoRR*. abs/2104.07650. arXiv: 2104.07650.

Chen, X., B. He, K. Hui, L. Sun, and Y. Sun. (2021c). "Simplified TinyBERT: Knowledge Distillation for Document Retrieval". In: *Lecture Notes in Computer Science*. Springer International Publishing. 241–248. DOI: 10.1007/978-3-030-72240-1_21.

Chen, Z., X. Fan, and Y. Ling. (2020b). "Pre-Training for Query Rewriting in a Spoken Language Understanding System". In: *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE. DOI: 10.1109/icassp40776.2020.9053531.

Chiang, W.-T. M., M. Hagenbuchner, and A. C. Tsoi. (2005). "The WT10G dataset and the evolution of the web". In: *Special interest tracks and posters of the 14th international conference on World Wide Web - WWW '05*. ACM Press. DOI: 10.1145/1062745.1062807.

Choi, J., E. Jung, J. Suh, and W. Rhee. (2021). "Improving Bi-encoder Document Ranking Models with Two Rankers and Multi-teacher Distillation". In: *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM. DOI: 10.1145/3404835.3463076.

Chuang, Y.-S., C.-L. Liu, H.-y. Lee, and L.-s. Lee. (2020). "SpeechBERT: An Audio-and-Text Jointly Learned Language Model for End-to-End Spoken Question Answering". In: *Interspeech 2020*. ISCA. DOI: 10.21437/interspeech.2020-1570.

Clinchant, S. and F. Perronnin. (2013). "Aggregating Continuous Word Embeddings for Information Retrieval". In: *Proceedings of the Workshop on Continuous Vector Space Models and their Compositionality, CVSM@ACL 2013, Sofia, Bulgaria, August 9, 2013*. Ed. by A. Allauzen, H. Larochelle, C. D. Manning, and R. Socher. Association for Computational Linguistics. 100–109. URL: https://aclanthology.org/W13-3212/.

Conneau, A. and D. Kiela. (2018). "SentEval: An Evaluation Toolkit for Universal Sentence Representations". In: *Proceedings of the Eleventh International Conference on Language Resources and Evaluation, LREC 2018, Miyazaki, Japan, May 7-12, 2018.* Ed. by N. Calzolari, K. Choukri, C. Cieri, T. Declerck, S. Goggi, K. Hasida, H. Isahara, B. Maegaard, J. Mariani, H. Mazo, A. Moreno, J. Odijk, S. Piperidis, and T. Tokunaga. European Language Resources Association (ELRA). URL: http://www.lrec-conf.org/proceedings/lrec2018/summaries/757.html.

Conneau, A. and G. Lample. (2019). "Cross-lingual Language Model Pretraining". In: *Advances in Neural Information Processing Systems 32: Annual Conference on Neural Information Processing Systems 2019, NeurIPS 2019, December 8-14, 2019, Vancouver, BC, Canada.* Ed. by H. M. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. B. Fox, and R. Garnett. 7057–7067. URL: https://proceedings.neurips.cc/paper/2019/hash/c04c19c2c2474dbf5f7ac4372c5b9af1-Abstract.html.

Crammer, K. and Y. Singer. (2001). "Pranking with Ranking". In: *Advances in Neural Information Processing Systems 14 [Neural Information Processing Systems: Natural and Synthetic, NIPS 2001, December 3-8, 2001, Vancouver, British Columbia, Canada].* Ed. by T. G. Dietterich, S. Becker, and Z. Ghahramani. MIT Press. 641–647. URL: https://proceedings.neurips.cc/paper/2001/hash/5531a5834816222280f20d1ef9e95f69-Abstract.html.

Craswell, N., W. B. Croft, J. Guo, B. Mitra, and M. de Rijke. (2017). "Report on the SIGIR 2016 Workshop on Neural Information Retrieval (Neu-IR)". *ACM SIGIR Forum.* 50(2): 96–103. DOI: 10.1145/3053408.3053425.

Craswell, N., B. Mitra, E. Yilmaz, D. Campos, and J. Lin. (2021). "MS MARCO: Benchmarking Ranking Models in the Large-Data Regime". In: *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval.* ACM. DOI: 10.1145/3404835.3462804.

Croft, W. B. and J. D. Lafferty. (2003). "Language Modeling for Information Retrieval". In: *The Springer International Series on Information Retrieval.*

Dai, Z. and J. Callan. (2019a). "Context-Aware Sentence/Passage Term Importance Estimation For First Stage Retrieval". *CoRR*. abs/1910.10687. arXiv: 1910.10687.

Dai, Z. and J. Callan. (2019b). "Deeper Text Understanding for IR with Contextual Neural Language Modeling". In: *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM. DOI: 10.1145/3331184.3331303.

Dai, Z. and J. Callan. (2020a). "Context-Aware Document Term Weighting for Ad-Hoc Search". In: *Proceedings of The Web Conference 2020*. ACM. DOI: 10.1145/3366423.3380258.

Dai, Z. and J. Callan. (2020b). "Context-Aware Term Weighting For First Stage Passage Retrieval". In: *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM. DOI: 10.1145/3397271.3401204.

Dang, H. T. (2005). "Overview of DUC 2005". In: *Proceedings of the document understanding conference*. Vol. 2005. 1–12.

Dehghani, M., S. Rothe, E. Alfonseca, and P. Fleury. (2017a). "Learning to Attend, Copy, and Generate for Session-Based Query Suggestion". In: *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*. ACM. DOI: 10.1145/3132847.3133010.

Dehghani, M., H. Zamani, A. Severyn, J. Kamps, and W. B. Croft. (2017b). "Neural Ranking Models with Weak Supervision". In: *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM. DOI: 10.1145/3077136.3080832.

Devlin, J., M.-W. Chang, K. Lee, and K. Toutanova. (2019). "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding". In: *Proceedings of the 2019 Conference of the North*. Association for Computational Linguistics. DOI: 10.18653/v1/n19-1423.

Diaz, F., B. Mitra, and N. Craswell. (2016). "Query Expansion with Locally-Trained Word Embeddings". In: *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Association for Computational Linguistics. DOI: 10.18653/v1/p16-1035.

Dietz, L., M. Verma, F. Radlinski, and N. Craswell. (2017). "TREC Complex Answer Retrieval Overview". In: *Proceedings of The Twenty-Sixth Text REtrieval Conference, TREC 2017, Gaithersburg, Maryland, USA, November 15-17, 2017*. Ed. by E. M. Voorhees and A. Ellis. Vol. 500-324. *NIST Special Publication*. National Institute of Standards and Technology (NIST). URL: https://trec.nist.gov/pubs/trec26/papers/Overview-CAR.pdf.

Dou, Z.-Y., P. Liu, H. Hayashi, Z. Jiang, and G. Neubig. (2021). "GSum: A General Framework for Guided Neural Abstractive Summarization". In: *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. Association for Computational Linguistics. DOI: 10.18653/v1/2021.naacl-main.384.

Dubey, S. R. (2020). "A Decade Survey of Content Based Image Retrieval using Deep Learning". *CoRR*. abs/2012.00641. arXiv: 2012.00641.

Echihabi, K., K. Zoumpatianos, T. Palpanas, and H. Benbrahim. (2019). "Return of the Lernaean Hydra". *Proceedings of the VLDB Endowment*. 13(3): 403–420. DOI: 10.14778/3368289.3368303.

Efron, M., P. Organisciak, and K. Fenlon. (2012). "Improving retrieval of short texts through document expansion". In: *Proceedings of the 35th international ACM SIGIR conference on Research and development in information retrieval - SIGIR '12*. ACM Press. DOI: 10.1145/2348283.2348405.

Erkan, G. and D. R. Radev. (2004). "LexRank: Graph-based Lexical Centrality as Salience in Text Summarization". *J. Artif. Intell. Res.* 22: 457–479. DOI: 10.1613/jair.1523.

Fan, Y., J. Guo, X. Ma, R. Zhang, Y. Lan, and X. Cheng. (2021). "A Linguistic Study on Relevance Modeling in Information Retrieval". In: *Proceedings of the Web Conference 2021*. ACM. DOI: 10.1145/3442381.3450009.

Fedus, W., B. Zoph, and N. Shazeer. (2021). "Switch Transformers: Scaling to Trillion Parameter Models with Simple and Efficient Sparsity". *CoRR*. abs/2101.03961. arXiv: 2101.03961.

Fei, H., T. Yu, and P. Li. (2021). "Cross-lingual Cross-modal Pretraining for Multimodal Retrieval". In: *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. Association for Computational Linguistics. DOI: 10.18653/v1/2021.naacl-main.285.

Feigenblat, G., H. Roitman, O. Boni, and D. Konopnicki. (2017). "Unsupervised Query-Focused Multi-Document Summarization using the Cross Entropy Method". In: *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM. DOI: 10.1145/3077136.3080690.

Fisher, S. and B. Roark. (2006). "Query-focused summarization by supervised sentence ranking and skewed word distributions". In: *Proceedings of the Document Understanding Conference, DUC-2006, New York, USA*. Citeseer.

Formal, T., B. Piwowarski, and S. Clinchant. (2021). "SPLADE: Sparse Lexical and Expansion Model for First Stage Ranking". In: *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM. DOI: 10.1145/3404835.3463098.

Frej, J., P. Mulhem, D. Schwab, and J.-P. Chevallet. (2020). "Learning Term Discrimination". In: *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM. DOI: 10.1145/3397271.3401211.

Frieder, O., D. A. Grossman, A. Chowdhury, and G. Frieder. (2000). "Efficiency Considerations for Scalable Information Retrieval Servers". *J. Digit. Inf.* 1(5). URL: http://journals.tdl.org/jodi/article/view/21.

Ganesh, P., Y. Chen, X. Lou, M. A. Khan, Y. Yang, D. Chen, M. Winslett, H. Sajjad, and P. Nakov. (2020). "Compressing Large-Scale Transformer-Based Models: A Case Study on BERT". *CoRR*. abs/2002.11985. arXiv: 2002.11985.

Ganguly, D., D. Roy, M. Mitra, and G. J. Jones. (2015). "Word Embedding based Generalized Language Model for Information Retrieval". In: *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM. DOI: 10.1145/2766462.2767780.

Gao, L. and J. Callan. (2021a). "Condenser: a Pre-training Architecture for Dense Retrieval". In: *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics. DOI: 10.18653/v1/2021.emnlp-main.75.

Gao, L. and J. Callan. (2021b). "Is Your Language Model Ready for Dense Representation Fine-tuning?" *CoRR*. abs/2104.08253. arXiv: 2104.08253.

Gao, L., Z. Dai, and J. Callan. (2020a). "Modularized Transfomer-based Ranking Framework". In: *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Association for Computational Linguistics. DOI: 10.18653/v1/2020.emnlp-main.342.

Gao, L., Z. Dai, and J. Callan. (2020b). "Understanding BERT Rankers Under Distillation". In: *Proceedings of the 2020 ACM SIGIR on International Conference on Theory of Information Retrieval*. ACM. DOI: 10.1145/3409256.3409838.

Gao, L., Z. Dai, and J. Callan. (2021a). "COIL: Revisit Exact Lexical Match in Information Retrieval with Contextualized Inverted List". In: *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. Association for Computational Linguistics. DOI: 10.18653/v1/2021.naacl-main.241.

Gao, L., Z. Dai, and J. Callan. (2021b). "Rethink Training of BERT Rerankers in Multi-stage Retrieval Pipeline". In: *Lecture Notes in Computer Science*. Springer International Publishing. 280–286. DOI: 10.1007/978-3-030-72240-1_26.

Gao, L., Z. Dai, Z. Fan, and J. Callan. (2020c). "Complementing Lexical Retrieval with Semantic Residual Embedding". *CoRR*. abs/2004.13969. arXiv: 2004.13969.

Gao, T., A. Fisch, and D. Chen. (2021c). "Making Pre-trained Language Models Better Few-shot Learners". In: *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*. Association for Computational Linguistics. DOI: 10.18653/v1/2021.acl-long.295.

Gillick, D., A. Presta, and G. S. Tomar. (2018). "End-to-End Retrieval in Continuous Space". *CoRR*. abs/1811.08008. arXiv: 1811.08008.

Grbovic, M., N. Djuric, V. Radosavljevic, F. Silvestri, R. Baeza-Yates, A. Feng, E. Ordentlich, L. Yang, and G. Owens. (2016). "Scalable Semantic Matching of Queries to Ads in Sponsored Search Advertising". In: *Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval*. ACM. DOI: 10.1145/2911451.2911538.

Grbovic, M., N. Djuric, V. Radosavljevic, F. Silvestri, and N. Bhamidipati. (2015). "Context- and Content-aware Embeddings for Query Rewriting in Sponsored Search". In: *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM. DOI: 10.1145/2766462.2767709.

Grover, A. and J. Leskovec. (2016). "node2vec: Scalable feature learning for networks". In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM. DOI: 10.1145/2939672.2939754.

Guo, J., Y. Cai, Y. Fan, F. Sun, R. Zhang, and X. Cheng. (2022). "Semantic Models for the First-Stage Retrieval: A Comprehensive Review". *ACM Transactions on Information Systems*. 40(4): 1–42. DOI: 10.1145/3486250.

Guo, J., Y. Fan, Q. Ai, and W. B. Croft. (2016). "A Deep Relevance Matching Model for Ad-hoc Retrieval". In: *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management*. ACM. DOI: 10.1145/2983323.2983769.

Guo, J., Y. Fan, L. Pang, L. Yang, Q. Ai, H. Zamani, C. Wu, W. B. Croft, and X. Cheng. (2020). "A Deep Look into neural ranking models for information retrieval". *Information Processing and Management*. 57(6): 102067. DOI: 10.1016/j.ipm.2019.102067.

Guu, K., K. Lee, Z. Tung, P. Pasupat, and M. Chang. (2020). "REALM: Retrieval-Augmented Language Model Pre-Training". *CoRR*. abs/2002.08909. arXiv: 2002.08909.

Gysel, C. V., M. de Rijke, and E. Kanoulas. (2018). "Neural Vector Spaces for Unsupervised Information Retrieval". *ACM Transactions on Information Systems*. 36(4): 1–25. DOI: 10.1145/3196826.

Habibi, M., P. Mahdabi, and A. Popescu-Belis. (2016). "Question answering in conversations: Query refinement using contextual and semantic information". *Data and Knowledge Engineering*. 106: 38–51. DOI: 10.1016/j.datak.2016.06.003.

Han, X., W. Zhao, N. Ding, Z. Liu, and M. Sun. (2021a). "PTR: Prompt Tuning with Rules for Text Classification". *CoRR*. abs/2105.11259. arXiv: 2105.11259.

Han, Y., G. Huang, S. Song, L. Yang, H. Wang, and Y. Wang. (2021b). "Dynamic Neural Networks: A Survey". *CoRR*. abs/2102.04906. arXiv: 2102.04906.

Hashemi, H., H. Zamani, and W. B. Croft. (2020). "Guided transformer: Leveraging multiple external sources for representation learning in conversational search". In: *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM. DOI: 10.1145/3397271.3401061.

Herbrich, R. (1999). "Support vector learning for ordinal regression". In: *9th International Conference on Artificial Neural Networks: ICANN '99*. IEE. DOI: 10.1049/cp:19991091.

Hinton, G. E., O. Vinyals, and J. Dean. (2015). "Distilling the Knowledge in a Neural Network". *CoRR*. abs/1503.02531. arXiv: 1503.02531.

Hofstätter, S., S. Althammer, M. Schröder, M. Sertkan, and A. Hanbury. (2020). "Improving Efficient Neural Ranking Models with Cross-Architecture Knowledge Distillation". *CoRR*. abs/2010.02666. arXiv: 2010.02666.

Hofstätter, S., S.-C. Lin, J.-H. Yang, J. Lin, and A. Hanbury. (2021a). "Efficiently Teaching an Effective Dense Retriever with Balanced Topic Aware Sampling". In: *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM. DOI: 10.1145/3404835.3462891.

Hofstätter, S., B. Mitra, H. Zamani, N. Craswell, and A. Hanbury. (2021b). "Intra-Document Cascading: Learning to Select Passages for Neural Document Ranking". In: *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM. DOI: 10.1145/3404835.3462889.

Houlsby, N., A. Giurgiu, S. Jastrzebski, B. Morrone, Q. de Laroussilhe, A. Gesmundo, M. Attariyan, and S. Gelly. (2019). "Parameter-Efficient Transfer Learning for NLP". In: *Proceedings of the 36th International Conference on Machine Learning, ICML 2019, 9-15 June 2019, Long Beach, California, USA*. Ed. by K. Chaudhuri and R. Salakhutdinov. Vol. 97. *Proceedings of Machine Learning Research*. PMLR. 2790–2799. URL: http://proceedings.mlr.press/v97/houlsby19a.html.

Howard, J. and S. Ruder. (2018). "Universal Language Model Fine-tuning for Text Classification". In: *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Association for Computational Linguistics. DOI: 10.18653/v1/p18-1031.

Hu, B., Z. Lu, H. Li, and Q. Chen. (2014). "Convolutional Neural Network Architectures for Matching Natural Language Sentences". In: *Advances in Neural Information Processing Systems 27: Annual Conference on Neural Information Processing Systems 2014, December 8-13 2014, Montreal, Quebec, Canada*. Ed. by Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence, and K. Q. Weinberger. 2042–2050. URL: https://proceedings.neurips.cc/paper/2014/hash/b9d487a30398d42ecff55c228ed5652b-Abstract.html.

Hu, E. J., Y. Shen, P. Wallis, Z. Allen-Zhu, Y. Li, S. Wang, and W. Chen. (2021). "LoRA: Low-Rank Adaptation of Large Language Models". *CoRR*. abs/2106.09685. arXiv: 2106.09685.

Huang, H., Y. Liang, N. Duan, M. Gong, L. Shou, D. Jiang, and M. Zhou. (2019). "Unicoder: A Universal Language Encoder by Pre-training with Multiple Cross-lingual Tasks". 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. DOI: 10.18653/v1/d19-1252.

Huang, J.-T., A. Sharma, S. Sun, L. Xia, D. Zhang, P. Pronin, J. Padmanabhan, G. Ottaviano, and L. Yang. (2020). "Embedding-based Retrieval in Facebook Search". In: *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM. DOI: 10.1145/3394486.3403305.

Huang, P.-S., X. He, J. Gao, L. Deng, A. Acero, and L. Heck. (2013). "Learning deep structured semantic models for web search using clickthrough data". In: *Proceedings of the 22nd ACM international conference on Conference on information and knowledge management - CIKM '13*. ACM Press. DOI: 10.1145/2505515.2505665.

Huang, Y., Z. Yu, J. Guo, Y. Xiang, and Y. Xian. (2021). "Element graph-augmented abstractive summarization for legal public opinion news with graph transformer". *Neurocomputing*. 460: 166–180. DOI: 10.1016/j.neucom.2021.07.013.

Huo, Y., M. Zhang, G. Liu, H. Lu, Y. Gao, G. Yang, J. Wen, H. Zhang, B. Xu, W. Zheng, *et al.* (2021). "WenLan: Bridging vision and language by large-scale multi-modal pre-training". *arXiv preprint arXiv:2103.06561*.

Jaleel, N. A., J. Allan, W. B. Croft, F. Diaz, L. S. Larkey, X. Li, M. D. Smucker, and C. Wade. (2004). "UMass at TREC 2004: Novelty and HARD". In: *Proceedings of the Thirteenth Text REtrieval Conference, TREC 2004, Gaithersburg, Maryland, USA, November 16-19, 2004*. Ed. by E. M. Voorhees and L. P. Buckland. Vol. 500-261. *NIST Special Publication*. National Institute of Standards and Technology (NIST). URL: http://trec.nist.gov/pubs/trec13/papers/umass.novelty.hard.pdf.

Jang, K., J. Kang, G. Hong, S. Myaeng, J. Park, T. Yoon, and H. Seo. (2021). "UHD-BERT: Bucketed Ultra-High Dimensional Sparse Representations for Full Ranking". *CoRR*. abs/2104.07198. arXiv: 2104.07198.

Jégou, H., M. Douze, and C. Schmid. (2011). "Product Quantization for Nearest Neighbor Search". *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 33(1): 117–128. DOI: 10.1109/tpami.2010.57.

Jiang, J.-Y. and W. Wang. (2018). "RIN: Reformulation inference network for context-aware query suggestion". In: *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*. ACM. DOI: 10.1145/3269206.3271808.

Jiang, Z., F. F. Xu, J. Araki, and G. Neubig. (2020). "How Can We Know What Language Models Know". *Trans. Assoc. Comput. Linguistics*. 8: 423–438. URL: https://transacl.org/ojs/index.php/tacl/article/view/1983.

Jiao, X., Y. Yin, L. Shang, X. Jiang, X. Chen, L. Li, F. Wang, and Q. Liu. (2020). "TinyBERT: Distilling BERT for Natural Language Understanding". In: *Findings of the Association for Computational Linguistics: EMNLP 2020*. Association for Computational Linguistics. DOI: 10.18653/v1/2020.findings-emnlp.372.

Johnson, J., A. Karpathy, and L. Fei-Fei. (2016). "DenseCap: Fully Convolutional Localization Networks for Dense Captioning". In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE. DOI: 10.1109/cvpr.2016.494.

Jung, E., J. Choi, and W. Rhee. (2021). "Semi-Siamese Bi-encoder Neural Ranking Model Using Lightweight Fine-Tuning". *CoRR*. abs/2110.14943. arXiv: 2110.14943.

Kågebäck, M., O. Mogren, N. Tahmasebi, and D. Dubhashi. (2014). "Extractive Summarization using Continuous Vector Space Models". In: *Proceedings of the 2nd Workshop on Continuous Vector Space Models and their Compositionality (CVSC)*. Association for Computational Linguistics. DOI: 10.3115/v1/w14-1504.

Karpukhin, V., B. Oguz, S. Min, P. Lewis, L. Wu, S. Edunov, D. Chen, and W.-t. Yih. (2020). "Dense Passage Retrieval for Open-Domain Question Answering". In: *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Association for Computational Linguistics. DOI: 10.18653/v1/2020.emnlp-main.550.

Ke, P., H. Ji, S. Liu, X. Zhu, and M. Huang. (2019). "SentiLR: Linguistic Knowledge Enhanced Language Representation for Sentiment Analysis". *CoRR*. abs/1911.02493. arXiv: 1911.02493.

Kenter, T. and M. de Rijke. (2015). "Short Text Similarity with Word Embeddings". In: *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management*. *CIKM '15*. Melbourne, Australia: Association for Computing Machinery. 1411–1420. DOI: 10.1145/2806416.2806475.

Khattab, O. and M. Zaharia. (2020). "Colbert: Efficient and effective passage search via contextualized late interaction over bert". In: *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM. DOI: 10.1145/3397271.3401075.

Kobayashi, H., M. Noguchi, and T. Yatsuka. (2015). "Summarization Based on Embedding Distributions". In: *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics. DOI: 10.18653/v1/d15-1232.

Kulkarni, S., S. Chammas, W. Zhu, F. Sha, and E. Ie. (2020). "AQuaMuSe: Automatically Generating Datasets for Query-Based Multi-Document Summarization". *CoRR*. abs/2010.12694. arXiv: 2010.12694.

Kumar, V., V. Raunak, and J. Callan. (2020). "Ranking Clarification Questions via Natural Language Inference". In: *Proceedings of the 29th ACM International Conference on Information and Knowledge Management*. ACM. DOI: 10.1145/3340531.3412137.

Kurland, O. and L. Lee. (2004). "Corpus structure, language models, and ad hoc information retrieval". In: *Proceedings of the 27th annual international conference on Research and development in information retrieval - SIGIR '04*. ACM Press. DOI: 10.1145/1008992.1009027.

Kuzi, S., D. Carmel, A. Libov, and A. Raviv. (2017). "Query Expansion for Email Search". In: *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM. DOI: 10.1145/3077136.3080660.

Kuzi, S., A. Shtok, and O. Kurland. (2016). "Query Expansion Using Word Embeddings". In: *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management*. ACM. DOI: 10.1145/2983323.2983876.

Kuzi, S., M. Zhang, C. Li, M. Bendersky, and M. Najork. (2020). "Leveraging Semantic and Lexical Matching to Improve the Recall of Document Retrieval Systems: A Hybrid Approach". *CoRR*. abs/2010.01195. arXiv: 2010.01195.

Kwiatkowski, T., J. Palomaki, O. Redfield, M. Collins, A. P. Parikh, C. Alberti, D. Epstein, I. Polosukhin, J. Devlin, K. Lee, K. Toutanova, L. Jones, M. Kelcey, M. Chang, A. M. Dai, J. Uszkoreit, Q. Le, and S. Petrov. (2019). "Natural Questions: a Benchmark for Question Answering Research". *Trans. Assoc. Comput. Linguistics*. 7: 452–466. URL: https://transacl.org/ojs/index.php/tacl/article/view/1455.

Lafferty, J. and C. Zhai. (2003). "Probabilistic relevance models based on document and query generation". In: *Language modeling for information retrieval*. Springer. 1–10.

Lan, Z., M. Chen, S. Goodman, K. Gimpel, P. Sharma, and R. Soricut. (2020). "ALBERT: A Lite BERT for Self-supervised Learning of Language Representations". In: *8th International Conference on Learning Representations, ICLR 2020, Addis Ababa, Ethiopia, April 26-30, 2020*. OpenReview.net. URL: https://openreview.net/forum?id=H1eA7AEtvS.

Laskar, M. T. R., E. Hoque, and J. Huang. (2020a). "Query Focused Abstractive Summarization via Incorporating Query Relevance and Transfer Learning with Transformer Models". In: *Advances in Artificial Intelligence*. Springer International Publishing. 342–348. DOI: 10.1007/978-3-030-47358-7_35.

Laskar, M. T. R., E. Hoque, and J. X. Huang. (2020b). "WSL-DS: Weakly Supervised Learning with Distant Supervision for Query Focused Multi-Document Abstractive Summarization". In: *Proceedings of the 28th International Conference on Computational Linguistics*. International Committee on Computational Linguistics. DOI: 10.18653/v1/2020.coling-main.495.

Lavrenko, V. and W. B. Croft. (2017). "Relevance-Based Language Models". *ACM SIGIR Forum*. 51(2): 260–267. DOI: 10.1145/3130348.3130376.

Le, Q. V. and T. Mikolov. (2014). "Distributed Representations of Sentences and Documents". In: *Proceedings of the 31th International Conference on Machine Learning, ICML 2014, Beijing, China, 21-26 June 2014*. Vol. 32. *JMLR Workshop and Conference Proceedings*. JMLR.org. 1188–1196. URL: http://proceedings.mlr.press/v32/le14.html.

Lee, J., M. Seo, H. Hajishirzi, and J. Kang. (2020). "Contextualized Sparse Representations for Real-Time Open-Domain Question Answering". In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics. DOI: 10.18653/v1/2020.acl-main.85.

Lee, J., W. Yoon, S. Kim, D. Kim, S. Kim, C. H. So, and J. Kang. (2019a). "BioBERT: a pre-trained biomedical language representation model for biomedical text mining". *Bioinformatics*. Ed. by J. Wren. DOI: 10.1093/bioinformatics/btz682.

Lee, K., M.-W. Chang, and K. Toutanova. (2019b). "Latent Retrieval for Weakly Supervised Open Domain Question Answering". In: *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics. DOI: 10.18653/v1/p19-1612.

Lee, K.-H., X. Chen, G. Hua, H. Hu, and X. He. (2018). "Stacked Cross Attention for Image-Text Matching". In: *Computer Vision – ECCV 2018*. Springer International Publishing. 212–228. DOI: 10.1007/978-3-030-01225-0_13.

Lester, B., R. Al-Rfou, and N. Constant. (2021). "The Power of Scale for Parameter-Efficient Prompt Tuning". In: *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics. DOI: 10.18653/v1/2021.emnlp-main.243.

Lewis, M., Y. Liu, N. Goyal, M. Ghazvininejad, A. Mohamed, O. Levy, V. Stoyanov, and L. Zettlemoyer. (2020a). "BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension". In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics. DOI: 10.18653/v1/2020.acl-main.703.

Lewis, P. S. H., E. Perez, A. Piktus, F. Petroni, V. Karpukhin, N. Goyal, H. Küttler, M. Lewis, W. Yih, T. Rocktäschel, S. Riedel, and D. Kiela. (2020b). "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks". In: *Advances in Neural Information Processing Systems 33: Annual Conference on Neural Information Processing Systems 2020, NeurIPS 2020, December 6-12, 2020, virtual.* Ed. by H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan, and H. Lin. URL: https://proceedings.neurips.cc/paper/2020/hash/6b493230205f780e1bc26945df7481e5-Abstract.html.

Li, C., A. Yates, S. MacAvaney, B. He, and Y. Sun. (2020a). "PARADE: Passage Representation Aggregation for Document Reranking". *CoRR*. abs/2008.09093. arXiv: 2008.09093.

Li, H. (2014). "Learning to Rank for Information Retrieval and Natural Language Processing, Second Edition". *Synthesis Lectures on Human Language Technologies*. 7(3): 1–121. DOI: 10.2200/s00607ed2v01y201410hlt026.

Li, H. and J. Xu. (2014). "Semantic Matching in Search". *Foundations and Trends® in Information Retrieval*. 7(5): 343–469. DOI: 10.1561/1500000035.

Li, L. H., M. Yatskar, D. Yin, C.-J. Hsieh, and K.-W. Chang. (2020b). "What Does BERT with Vision Look At?" In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics. DOI: 10.18653/v1/2020.acl-main.469.

Li, P., C. J. C. Burges, and Q. Wu. (2007). "McRank: Learning to Rank Using Multiple Classification and Gradient Boosting". In: *Advances in Neural Information Processing Systems 20, Proceedings of the Twenty-First Annual Conference on Neural Information Processing Systems, Vancouver, British Columbia, Canada, December 3-6, 2007*. Ed. by J. C. Platt, D. Koller, Y. Singer, and S. T. Roweis. Curran Associates, Inc. 897–904. URL: https://proceedings.neurips.cc/paper/2007/hash/b86e8d03fe992d1b0e19656875ee557c-Abstract.html.

Li, W., C. Gao, G. Niu, X. Xiao, H. Liu, J. Liu, H. Wu, and H. Wang. (2021). "UNIMO: Towards Unified-Modal Understanding and Generation via Cross-Modal Contrastive Learning". In: *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*. Association for Computational Linguistics. DOI: 10.18653/v1/2021.acl-long.202.

Li, W., Y. Zhang, Y. Sun, W. Wang, M. Li, W. Zhang, and X. Lin. (2020c). "Approximate Nearest Neighbor Search on High Dimensional Data — Experiments, Analyses, and Improvement". *IEEE Transactions on Knowledge and Data Engineering*. 32(8): 1475–1488. DOI: 10.1109/tkde.2019.2909204.

Li, X. L. and P. Liang. (2021). "Prefix-Tuning: Optimizing Continuous Prompts for Generation". In: *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*. Association for Computational Linguistics. DOI: 10.18653/v1/2021.acl-long.353.

Li, X., X. Yin, C. Li, P. Zhang, X. Hu, L. Zhang, L. Wang, H. Hu, L. Dong, F. Wei, Y. Choi, and J. Gao. (2020d). "Oscar: Object-Semantics Aligned Pre-training for Vision-Language Tasks". In: *Computer Vision – ECCV 2020*. Springer International Publishing. 121–137. DOI: 10.1007/978-3-030-58577-8_8.

Liang, D., P. Xu, S. Shakeri, C. N. dos Santos, R. Nallapati, Z. Huang, and B. Xiang. (2020). "Embedding-based Zero-shot Retrieval through Query Generation". *CoRR*. abs/2009.10270. arXiv: 2009.10270.

Lim, Y., D. Seo, and Y. Jung. (2020). "Fine-tuning BERT Models for Keyphrase Extraction in Scientific Articles". *Journal of Advanced Information Technology and Convergence*. 10(1): 45–56.

Lin, J. and X. Ma. (2021). "A Few Brief Notes on DeepImpact, COIL, and a Conceptual Framework for Information Retrieval Techniques". *CoRR*. abs/2106.14807. arXiv: 2106.14807.

Lin, J., R. Nogueira, and A. Yates. (2021a). "Pretrained Transformers for Text Ranking: BERT and Beyond". *Synthesis Lectures on Human Language Technologies*. 14(4): 1–325. DOI: 10.2200/s01123ed1v01y202108hlt053.

Lin, S.-C., J.-H. Yang, and J. Lin. (2021b). "In-Batch Negatives for Knowledge Distillation with Tightly-Coupled Teachers for Dense Retrieval". In: *Proceedings of the 6th Workshop on Representation Learning for NLP (RepL4NLP-2021)*. Association for Computational Linguistics. DOI: 10.18653/v1/2021.repl4nlp-1.17.

Lin, S., J. Yang, R. Nogueira, M. Tsai, C. Wang, and J. Lin. (2020). "Query Reformulation using Query History for Passage Retrieval in Conversational Search". *CoRR*. abs/2005.02230. arXiv: 2005.02230.

Liu, B., H. Zamani, X. Lu, and J. S. Culpepper. (2021a). "Generalizing Discriminative Retrieval Models using Generative Tasks". In: *Proceedings of the Web Conference 2021*. ACM. DOI: 10.1145/3442381.3449863.

Liu, H., M. Chen, Y. Wu, X. He, and B. Zhou. (2021b). "Conversational Query Rewriting with Self-Supervised Learning". In: *ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE. DOI: 10.1109/icassp39728.2021.9413557.

Liu, P., W. Yuan, J. Fu, Z. Jiang, H. Hayashi, and G. Neubig. (2021c). "Pre-train, Prompt, and Predict: A Systematic Survey of Prompting Methods in Natural Language Processing". *CoRR*. abs/2107.13586. arXiv: 2107.13586.

Liu, R., Z. Lin, and W. Wang. (2021d). "Addressing Extraction and Generation Separately: Keyphrase Prediction With Pre-Trained Language Models". *IEEE/ACM Transactions on Audio, Speech, and Language Processing*. 29: 3180–3191. DOI: 10.1109/taslp.2021.3120587.

Liu, T.-Y. (2007). "Learning to Rank for Information Retrieval". *Foundations and Trends® in Information Retrieval*. 3(3): 225–331. DOI: 10.1561/1500000016.

Liu, X., F. Zhang, Z. Hou, L. Mian, Z. Wang, J. Zhang, and J. Tang. (2021e). "Self-supervised Learning: Generative or Contrastive". *IEEE Transactions on Knowledge and Data Engineering*: 1–1. DOI: 10.1109/tkde.2021.3090866.

Liu, X., Y. Zheng, Z. Du, M. Ding, Y. Qian, Z. Yang, and J. Tang. (2021f). "GPT Understands, Too". *CoRR*. abs/2103.10385. arXiv: 2103.10385.

Liu, X. and W. B. Croft. (2004). "Cluster-based retrieval using language models". In: *Proceedings of the 27th annual international conference on Research and development in information retrieval - SIGIR '04*. ACM Press. DOI: 10.1145/1008992.1009026.

Liu, X. and W. B. Croft. (2006). "Statistical language modeling for information retrieval". *Annual Review of Information Science and Technology*. 39(1): 1–31. DOI: 10.1002/aris.1440390108.

Liu, Y. and M. Lapata. (2019). "Text Summarization with Pretrained Encoders". 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. DOI: 10.18653/v1/d19-1387.

Liu, Y., W. Lu, S. Cheng, D. Shi, S. Wang, Z. Cheng, and D. Yin. (2021g). "Pre-trained Language Model for Web-scale Retrieval in Baidu Search". In: *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. ACM. DOI: 10.1145/3447548.3467149.

Liu, Y., M. Ott, N. Goyal, J. Du, M. Joshi, D. Chen, O. Levy, M. Lewis, L. Zettlemoyer, and V. Stoyanov. (2019). "RoBERTa: A Robustly Optimized BERT Pretraining Approach". *CoRR*. abs/1907.11692. arXiv: 1907.11692.

Lotze, T., S. Klut, M. Aliannejadi, and E. Kanoulas. (2021). "Ranking Clarifying Questions Based on Predicted User Engagement". *CoRR*. abs/2103.06192. arXiv: 2103.06192.

Lu, J., D. Batra, D. Parikh, and S. Lee. (2019). "ViLBERT: Pretraining Task-Agnostic Visiolinguistic Representations for Vision-and-Language Tasks". In: *Advances in Neural Information Processing Systems 32: Annual Conference on Neural Information Processing Systems 2019, NeurIPS 2019, December 8-14, 2019, Vancouver, BC, Canada*. Ed. by H. M. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. B. Fox, and R. Garnett. 13–23. URL: https://proceedings.neurips.cc/paper/2019/hash/c74d97b01eae257e44aa9d5bade97baf-Abstract.html.

Lu, S., D. He, C. Xiong, G. Ke, W. Malik, Z. Dou, P. Bennett, T.-Y. Liu, and A. Overwijk. (2021). "Less is More: Pretrain a Strong Siamese Encoder for Dense Text Retrieval Using a Weak Decoder". In: *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics. DOI: 10.18653/v1/2021.emnlp-main.220.

Luan, Y., J. Eisenstein, K. Toutanova, and M. Collins. (2021). "Sparse, Dense, and Attentional Representations for Text Retrieval". *Transactions of the Association for Computational Linguistics*. 9: 329–345. DOI: 10.1162/tacl_a_00369.

Luo, C., Y. Zheng, Y. Liu, X. Wang, J. Xu, M. Zhang, and S. Ma. (2017a). "SogouT-16". In: *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM. DOI: 10.1145/3077136.3080694.

Luo, C., Y. Zheng, J. Mao, Y. Liu, M. Zhang, and S. Ma. (2017b). "Training Deep Ranking Model with Weak Relevance Labels". In: *Lecture Notes in Computer Science*. Springer International Publishing. 205–216. DOI: 10.1007/978-3-319-68155-9_16.

Ma, J., I. Korotkov, Y. Yang, K. Hall, and R. McDonald. (2021a). "Zero-shot Neural Passage Retrieval via Domain-targeted Synthetic Question Generation". In: *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume*. Association for Computational Linguistics. DOI: 10.18653/v1/2021.eacl-main.92.

Ma, X., J. Guo, R. Zhang, Y. Fan, X. Ji, and X. Cheng. (2021b). "PROP: Pre-training with Representative Words Prediction for Ad-hoc Retrieval". In: *Proceedings of the 14th ACM International Conference on Web Search and Data Mining*. ACM. DOI: 10.1145/3437963.3441777.

Ma, X., J. Guo, R. Zhang, Y. Fan, Y. Li, and X. Cheng. (2021c). "B-PROP". In: *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM. DOI: 10.1145/3404835.3462869.

Ma, Z., Z. Dou, W. Xu, X. Zhang, H. Jiang, Z. Cao, and J.-R. Wen. (2021d). "Pre-training for Ad-hoc Retrieval: Hyperlink is Also You Need". In: *Proceedings of the 30th ACM International Conference on Information and Knowledge Management*. ACM. DOI: 10.1145/3459637.3482286.

MacAvaney, S., F. M. Nardini, R. Perego, N. Tonellotto, N. Goharian, and O. Frieder. (2020). "Efficient Document Re-Ranking for Transformers by Precomputing Term Representations". In: *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM. DOI: 10.1145/3397271.3401093.

MacAvaney, S., A. Yates, A. Cohan, and N. Goharian. (2019). "CEDR: Contextualized Embeddings for Document Ranking". In: *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 2019, Paris, France, July 21-25, 2019*. ACM. 1101–1104. DOI: 10.1145/3331184.3331317.

Mahata, D., J. Kuriakose, R. R. Shah, and R. Zimmermann. (2018). "Key2Vec: Automatic Ranked Keyphrase Extraction from Scientific Articles using Phrase Embeddings". In: *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 2 (Short Papers)*. Association for Computational Linguistics. DOI: 10.18653/v1/n18-2100.

Mallia, A., O. Khattab, T. Suel, and N. Tonellotto. (2021). "Learning Passage Impacts for Inverted Indexes". In: *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM. DOI: 10.1145/3404835.3463030.

Manning, C. D., P. Raghavan, and H. Schütze. (2008). *Introduction to Information Retrieval*. Cambridge University Press. DOI: 10.1017/cbo9780511809071.

Mass, Y., B. Carmeli, H. Roitman, and D. Konopnicki. (2020). "Unsupervised FAQ Retrieval with Question Generation and BERT". In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics. DOI: 10.18653/v1/2020.acl-main.74.

McCann, B., J. Bradbury, C. Xiong, and R. Socher. (2017). "Learned in Translation: Contextualized Word Vectors". In: *Annual Conference on Neural Information Processing Systems 2017, December 4-9, 2017, Long Beach, CA, USA*. 6294–6305. URL: https://proceedings.neurips.cc/paper/2017/hash/20c86a628232a67e7bd46f76fba7ce12-Abstract.html.

McCreery, C. H., N. Katariya, A. Kannan, M. Chablani, and X. Amatriain. (2020). "Effective Transfer Learning for Identifying Similar Questions". In: *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM. DOI: 10.1145/3394486.3412861.

Mei, Q., J. Guo, and D. Radev. (2010). "DivRank". In: *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '10*. ACM Press. DOI: 10.1145/1835804.1835931.

Melamud, O., J. Goldberger, and I. Dagan. (2016). "context2vec: Learning Generic Context Embedding with Bidirectional LSTM". In: *CoNLL*.

Metzler, D., Y. Tay, D. Bahri, and M. Najork. (2021). "Rethinking search". *ACM SIGIR Forum*. 55(1): 1–27. DOI: 10.1145/3476415.3476428.

Mikolov, T., K. Chen, G. Corrado, and J. Dean. (2013a). "Efficient Estimation of Word Representations in Vector Space". In: *1st International Conference on Learning Representations, ICLR 2013, Scottsdale, Arizona, USA, May 2-4, 2013, Workshop Track Proceedings*. URL: http://arxiv.org/abs/1301.3781.

Mikolov, T., I. Sutskever, K. Chen, G. S. Corrado, and J. Dean. (2013b). "Distributed Representations of Words and Phrases and their Compositionality". In: *27th Annual Conference on Neural Information Processing Systems 2013. Proceedings of a meeting held December 5-8, 2013, Lake Tahoe, Nevada, United States*. 3111–3119. URL: https://proceedings.neurips.cc/paper/2013/hash/9aa42b31882ec039965f3c4923ce901b-Abstract.html.

Mikolov, T., W. Yih, and G. Zweig. (2013c). "Linguistic Regularities in Continuous Space Word Representations". In: *Human Language Technologies: Conference of the North American Chapter of the Association of Computational Linguistics, Proceedings, June 9-14, 2013, Westin Peachtree Plaza Hotel, Atlanta, Georgia, USA*. The Association for Computational Linguistics. 746–751. URL: https://aclanthology.org/N13-1090/.

Mitra, B. and N. Craswell. (2018). "An Introduction to Neural Information Retrieval t". *Foundations and Trends® in Information Retrieval*. 13(1): 1–126. DOI: 10.1561/1500000061.

Mitra, B., F. Diaz, and N. Craswell. (2017). "Learning to Match using Local and Distributed Representations of Text for Web Search". In: *Proceedings of the 26th International Conference on World Wide Web*. International World Wide Web Conferences Steering Committee. DOI: 10.1145/3038912.3052579.

Mitra, B., E. T. Nalisnick, N. Craswell, and R. Caruana. (2016). "A Dual Embedding Space Model for Document Ranking". *CoRR*. abs/1602.01137. arXiv: 1602.01137.

Mitra, R., M. Gupta, and S. Dandapat. (2020). "Transformer Models for Recommending Related Questions in Web Search". In: *Proceedings of the 29th ACM International Conference on Information and Knowledge Management*. ACM. DOI: 10.1145/3340531.3412067.

Mohd, M., R. Jan, and M. Shah. (2020). "Text document summarization using word embedding". *Expert Systems with Applications*. 143: 112958. DOI: 10.1016/j.eswa.2019.112958.

Mustar, A., S. Lamprier, and B. Piwowarski. (2020). "Using BERT and BART for Query Suggestion". In: *Proceedings of the First Joint Conference of the Information Retrieval Communities in Europe (CIRCLE 2020), Samatan, Gers, France, July 6-9, 2020*. Vol. 2621. *CEUR Workshop Proceedings*. CEUR-WS.org. URL: http://ceur-ws.org/Vol-2621/CIRCLE20%5C_06.pdf.

Naseri, S., J. Dalton, A. Yates, and J. Allan. (2021). "CEQE: Contextualized Embeddings for Query Expansion". In: *Lecture Notes in Computer Science*. Springer International Publishing. 467–482. DOI: 10.1007/978-3-030-72113-8_31.

Nema, P., M. M. Khapra, A. Laha, and B. Ravindran. (2017). "Diversity driven attention model for query-based abstractive summarization". In: *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Association for Computational Linguistics. DOI: 10.18653/v1/p17-1098.

Ni, J., C. Qu, J. Lu, Z. Dai, G. H. Ábrego, J. Ma, V. Y. Zhao, Y. Luan, K. B. Hall, M. Chang, and Y. Yang. (2021a). "Large Dual Encoders Are Generalizable Retrievers". *CoRR*. abs/2112.07899. arXiv: 2112.07899.

Ni, M., H. Huang, L. Su, E. Cui, T. Bharti, L. Wang, D. Zhang, and N. Duan. (2021b). "M3P: Learning Universal Representations via Multitask Multilingual Multimodal Pre-Training". In: *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2021, virtual, June 19-25, 2021*. Computer Vision Foundation / IEEE. 3977–3986. URL: https://openaccess.thecvf.com/content/CVPR2021/html/Ni%5C_M3P%5C_Learning%5C_Universal%5C_Representations%5C_via%5C_Multitask%5C_Multilingual%5C_Multimodal%5C_Pre-Training%5C_CVPR%5C_2021%5C_paper.html.

Nogueira, R. and K. Cho. (2019). "Passage Re-ranking with BERT". *CoRR*. abs/1901.04085. arXiv: 1901.04085.

Nogueira, R., Z. Jiang, R. Pradeep, and J. Lin. (2020). "Document Ranking with a Pretrained Sequence-to-Sequence Model". In: *Findings of the Association for Computational Linguistics: EMNLP 2020*. Association for Computational Linguistics. DOI: 10.18653/v1/2020.findings-emnlp.63.

Nogueira, R., J. Lin, and A. Epistemic. (2019a). "From doc2query to docTTTTTquery". *Online preprint*. 6.

Nogueira, R., W. Yang, K. Cho, and J. Lin. (2019b). "Multi-Stage Document Ranking with BERT". *CoRR*. abs/1910.14424. arXiv: 1910.14424.

Onal, K. D., Y. Zhang, I. S. Altingovde, M. M. Rahman, P. Karagoz, A. Braylan, B. Dang, H.-L. Chang, H. Kim, Q. McNamara, A. Angert, E. Banner, V. Khetan, T. McDonnell, A. T. Nguyen, D. Xu, B. C. Wallace, M. de Rijke, and M. Lease. (2017). "Neural information retrieval: at the end of the early years". *Information Retrieval Journal*. 21(2-3): 111–182. DOI: 10.1007/s10791-017-9321-y.

Padaki, R., Z. Dai, and J. Callan. (2020). "Rethinking Query Expansion for BERT Reranking". In: *Advances in Information Retrieval - 42nd European Conference on IR Research, ECIR 2020, Lisbon, Portugal, April 14-17, 2020, Proceedings, Part II*. Vol. 12036. *Lecture Notes in Computer Science*. Springer. 297–304. DOI: 10.1007/978-3-030-45442-5_37.

Papagiannopoulou, E. and G. Tsoumakas. (2018). "Local word vectors guiding keyphrase extraction". *Information Processing and Management*. 54(6): 888–902. DOI: 10.1016/j.ipm.2018.06.004.

Park, S. and C. Caragea. (2020). "Scientific Keyphrase Identification and Classification by Pre-Trained Language Models Intermediate Task Transfer Learning". In: *Proceedings of the 28th International Conference on Computational Linguistics*. International Committee on Computational Linguistics. DOI: 10.18653/v1/2020.coling-main.472.

Pennington, J., R. Socher, and C. Manning. (2014). "Glove: Global Vectors for Word Representation". In: *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Association for Computational Linguistics. DOI: 10.3115/v1/d14-1162.

Peters, M., M. Neumann, M. Iyyer, M. Gardner, C. Clark, K. Lee, and L. Zettlemoyer. (2018). “Deep Contextualized Word Representations”. In: *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*. Association for Computational Linguistics. DOI: 10.18653/v1/n18-1202.

Petroni, F., T. Rocktäschel, S. Riedel, P. Lewis, A. Bakhtin, Y. Wu, and A. Miller. (2019). “Language Models as Knowledge Bases?” 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. DOI: 10.18653/v1/d19-1250.

Ponte, J. M. and W. B. Croft. (2017). “A Language Modeling Approach to Information Retrieval”. *ACM SIGIR Forum*. 51(2): 202–208. DOI: 10.1145/3130348.3130368.

Pradeep, R., R. Nogueira, and J. Lin. (2021). “The Expando-Mono-Duo Design Pattern for Text Ranking with Pretrained Sequence-to-Sequence Models”. *CoRR*. abs/2101.05667. arXiv: 2101.05667.

Puri, R. and B. Catanzaro. (2019). “Zero-shot Text Classification With Generative Language Models”. *CoRR*. abs/1912.10165. arXiv: 1912.10165.

Qiao, Y., C. Xiong, Z. Liu, and Z. Liu. (2019). “Understanding the Behaviors of BERT in Ranking”. *CoRR*. abs/1904.07531. arXiv: 1904.07531.

Qiu, Q., Z. Xie, L. Wu, and W. Li. (2019). “Geoscience keyphrase extraction algorithm using enhanced word embedding”. *Expert Systems with Applications*. 125: 157–169. DOI: 10.1016/j.eswa.2019.02.001.

Qiu, X., T. Sun, Y. Xu, Y. Shao, N. Dai, and X. Huang. (2020). “Pretrained models for natural language processing: A survey”. *Science China Technological Sciences*. 63(10): 1872–1897. DOI: 10.1007/s11431-020-1647-3.

Qu, Y., Y. Ding, J. Liu, K. Liu, R. Ren, W. X. Zhao, D. Dong, H. Wu, and H. Wang. (2021). "RocketQA: An Optimized Training Approach to Dense Passage Retrieval for Open-Domain Question Answering". In: *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. Association for Computational Linguistics. DOI: 10.18653/v1/2021.naacl-main.466.

Radford, A., K. Narasimhan, T. Salimans, and I. Sutskever. (2018). "Improving language understanding by generative pre-training".

Radford, A., J. Wu, R. Child, D. Luan, D. Amodei, I. Sutskever, *et al.* (2019). "Language models are unsupervised multitask learners". *OpenAI blog*. 1(8): 9.

Raffel, C., N. Shazeer, A. Roberts, K. Lee, S. Narang, M. Matena, Y. Zhou, W. Li, and P. J. Liu. (2020). "Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer". *J. Mach. Learn. Res.* 21: 140:1–140:67. URL: http://jmlr.org/papers/v21/20-074.html.

Reddy, R. G., V. Yadav, M. A. Sultan, M. Franz, V. Castelli, H. Ji, and A. Sil. (2021). "Towards Robust Neural Retrieval Models with Synthetic Pre-Training". *arXiv preprint arXiv:2104.07800*.

Reimers, N. and I. Gurevych. (2019). "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. DOI: 10.18653/v1/d19-1410.

Ren, R., Y. Qu, J. Liu, W. X. Zhao, Q. She, H. Wu, H. Wang, and J.-R. Wen. (2021). "RocketQAv2: A Joint Training Method for Dense Passage Retrieval and Passage Re-ranking". In: *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics. DOI: 10.18653/v1/2021.emnlp-main.224.

Robertson, S. E. and K. S. Jones. (1976). "Relevance weighting of search terms". *Journal of the American Society for Information Science*. 27(3): 129–146. DOI: 10.1002/asi.4630270302.

Robertson, S. E. and S. Walker. (1994). "Some Simple Effective Approximations to the 2-Poisson Model for Probabilistic Weighted Retrieval". In: *SIGIR '94*. Springer London. 232–241. DOI: 10.1007/978-1-4471-2099-5_24.

Robertson, S. and H. Zaragoza. (2009). "The Probabilistic Relevance Framework: BM25 and Beyond". *Foundations and Trends® in Information Retrieval*. 3(4): 333–389. DOI: 10.1561/1500000019.

Robertson, S. E., S. Walker, S. Jones, M. Hancock-Beaulieu, and M. Gatford. (1994). "Okapi at TREC-3". In: *Proceedings of The Third Text REtrieval Conference, TREC 1994, Gaithersburg, Maryland, USA, November 2-4, 1994*. Vol. 500-225. *NIST Special Publication*. National Institute of Standards and Technology (NIST). 109–126. URL: http://trec.nist.gov/pubs/trec3/papers/city.ps.gz.

Roitman, H., G. Feigenblat, D. Cohen, O. Boni, and D. Konopnicki. (2020). "Unsupervised Dual-Cascade Learning with Pseudo-Feedback Distillation for Query-Focused Extractive Summarization". In: *Proceedings of The Web Conference 2020*. ACM. DOI: 10.1145/3366423.3380009.

Rosset, C., C. Xiong, X. Song, D. Campos, N. Craswell, S. Tiwary, and P. Bennett. (2020). "Leading Conversational Search by Suggesting Useful Questions". In: *Proceedings of The Web Conference 2020*. ACM. DOI: 10.1145/3366423.3380193.

Roy, D., D. Ganguly, M. Mitra, and G. J. Jones. (2016a). "Representing documents and queries as sets of word embedded vectors for information retrieval". In: *Proceedings of Neu-IR: The SIGIR 2016 Workshop on Neural Information Retrieval*.

Roy, D., D. Paul, M. Mitra, and U. Garain. (2016b). "Using Word Embeddings for Automatic Query Expansion". *CoRR*. abs/1606.07608. arXiv: 1606.07608.

Sachan, D. S., S. Reddy, W. L. Hamilton, C. Dyer, and D. Yogatama. (2021). "End-to-End Training of Multi-Document Reader and Retriever for Open-Domain Question Answering". *CoRR*. abs/2106.05346. arXiv: 2106.05346.

Sahrawat, D., D. Mahata, H. Zhang, M. Kulkarni, A. Sharma, R. Gosangi, A. Stent, Y. Kumar, R. R. Shah, and R. Zimmermann. (2020). "Keyphrase Extraction as Sequence Labeling Using Contextualized Embeddings". In: *Lecture Notes in Computer Science*. Springer International Publishing. 328–335. DOI: 10.1007/978-3-030-45442-5_41.

Saito, I., K. Nishida, K. Nishida, and J. Tomita. (2020). "Abstractive Summarization with Combination of Pre-trained Sequence-to-Sequence and Saliency Models". *CoRR*. abs/2003.13028. arXiv: 2003.13028.

Sakata, W., T. Shibata, R. Tanaka, and S. Kurohashi. (2019). "FAQ Retrieval using Query-Question Similarity and BERT-Based Query-Answer Relevance". In: *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM. DOI: 10.1145/3331184.3331326.

Salton, G., A. Wong, and C. S. Yang. (1975). "A vector space model for automatic indexing". *Communications of the ACM*. 18(11): 613–620. DOI: 10.1145/361219.361220.

Sanh, V., L. Debut, J. Chaumond, and T. Wolf. (2019). "DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter". *CoRR*. abs/1910.01108. arXiv: 1910.01108.

Santos, C. N. dos, X. Ma, R. Nallapati, Z. Huang, and B. Xiang. (2020). "Beyond [CLS] through Ranking by Generation". In: *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Association for Computational Linguistics. DOI: 10.18653/v1/2020.emnlp-main.134.

Saracevic, T. (2016). "The Notion of Relevance in Information Science: Everybody knows what relevance is. But, what is it really?" *Synthesis Lectures on Information Concepts, Retrieval, and Services*. 8(3): i–109. DOI: 10.2200/s00723ed1v01y201607icr050.

Savery, M. E., A. B. Abacha, S. Gayen, and D. Demner-Fushman. (2020). "Question-Driven Summarization of Answers to Consumer Health Questions". *CoRR*. abs/2005.09067. arXiv: 2005.09067.

Schick, T. and H. Schütze. (2021a). "Exploiting Cloze-Questions for Few-Shot Text Classification and Natural Language Inference". In: *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume*. Association for Computational Linguistics. DOI: 10.18653/v1/2021.eacl-main.20.

Schick, T. and H. Schütze. (2021b). "It's Not Just Size That Matters: Small Language Models Are Also Few-Shot Learners". *ArXiv*. abs/2009.07118.

See, A., P. J. Liu, and C. D. Manning. (2017). "Get To The Point: Summarization with Pointer-Generator Networks". In: *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Association for Computational Linguistics. DOI: 10.18653/v1/p17-1099.

Sekulic, I., A. Soleimani, M. Aliannejadi, and F. A. Crestani. (2020). "Longformer for MS MARCO Document Re-ranking Task". *ArXiv*. abs/2009.09392.

Seo, M., J. Lee, T. Kwiatkowski, A. Parikh, A. Farhadi, and H. Hajishirzi. (2019). "Real-Time Open-Domain Question Answering with Dense-Sparse Phrase Index". In: *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics. DOI: 10.18653/v1/p19-1436.

Shen, Y., X. He, J. Gao, L. Deng, and G. Mesnil. (2014). "A Latent Semantic Model with Convolutional-Pooling Structure for Information Retrieval". In: *Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management*. ACM. DOI: 10.1145/2661829.2661935.

Sherman, G. and M. Efron. (2017). "Document Expansion Using External Collections". In: *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM. DOI: 10.1145/3077136.3080716.

Shi, P., H. Bai, and J. Lin. (2020). "Cross-Lingual Training of Neural Models for Document Ranking". In: *Findings of the Association for Computational Linguistics: EMNLP 2020*. Association for Computational Linguistics. DOI: 10.18653/v1/2020.findings-emnlp.249.

Shleifer, S. and A. M. Rush. (2020). "Pre-trained Summarization Distillation". *CoRR*. abs/2010.13002. arXiv: 2010.13002.

Singhal, A., C. Buckley, and M. Mitra. (2017). "Pivoted Document Length Normalization". *ACM SIGIR Forum*. 51(2): 176–184. DOI: 10.1145/3130348.3130365.

Soldaini, L. and A. Moschitti. (2020). "The Cascade Transformer: an Application for Efficient Answer Sentence Selection". In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics. DOI: 10.18653/v1/2020.acl-main.504.

Song, K., X. Tan, T. Qin, J. Lu, and T. Liu. (2019). "MASS: Masked Sequence to Sequence Pre-training for Language Generation". In: *Proceedings of the 36th International Conference on Machine Learning, ICML 2019, 9-15 June 2019, Long Beach, California, USA*. Ed. by K. Chaudhuri and R. Salakhutdinov. Vol. 97. *Proceedings of Machine Learning Research*. PMLR. 5926–5936. URL: http://proceedings.mlr.press/v97/song19d.html.

Sordoni, A., Y. Bengio, H. Vahabi, C. Lioma, J. G. Simonsen, and J.-Y. Nie. (2015). "A Hierarchical Recurrent Encoder-Decoder for Generative Context-Aware Query Suggestion". In: *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management*. ACM. DOI: 10.1145/2806416.2806493.

Su, D., Y. Xu, T. Yu, F. B. Siddique, E. Barezi, and P. Fung. (2020a). "CAiRE-COVID: A Question Answering and Query-focused Multi-Document Summarization System for COVID-19 Scholarly Information Management". In: *Proceedings of the 1st Workshop on NLP for COVID-19 (Part 2) at EMNLP 2020*. Association for Computational Linguistics. DOI: 10.18653/v1/2020.nlpcovid19-2.14.

Su, W., X. Zhu, Y. Cao, B. Li, L. Lu, F. Wei, and J. Dai. (2020b). "VL-BERT: Pre-training of Generic Visual-Linguistic Representations". In: *8th International Conference on Learning Representations, ICLR 2020, Addis Ababa, Ethiopia, April 26-30, 2020*. OpenReview.net. URL: https://openreview.net/forum?id=SygXPaEYvH.

Sun, C., A. Myers, C. Vondrick, K. Murphy, and C. Schmid. (2019a). "VideoBERT: A Joint Model for Video and Language Representation Learning". In: *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*. IEEE. DOI: 10.1109/iccv.2019.00756.

Sun, S., C. Xiong, Z. Liu, Z. Liu, and J. Bao. (2020). "Joint Keyphrase Chunking and Salience Ranking with BERT". *CoRR*. abs/2004.13639. arXiv: 2004.13639.

Sun, Y., S. Wang, Y. Li, S. Feng, X. Chen, H. Zhang, X. Tian, D. Zhu, H. Tian, and H. Wu. (2019b). "ERNIE: Enhanced Representation through Knowledge Integration". *CoRR*. abs/1904.09223. arXiv: 1904.09223.

Tahami, A. V., K. Ghajar, and A. Shakery. (2020). "Distilling Knowledge for Fast Retrieval-based Chat-bots". In: *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM. DOI: 10.1145/3397271.3401296.

Tang, H., X. Sun, B. Jin, J. Wang, F. Zhang, and W. Wu. (2021). "Improving Document Representations by Generating Pseudo Query Embeddings for Dense Retrieval". In: *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*. Association for Computational Linguistics. DOI: 10.18653/v1/2021.acl-long.392.

Tang, M., P. Gandhi, M. A. Kabir, C. Zou, J. Blakey, and X. Luo. (2019). "Progress Notes Classification and Keyword Extraction using Attention-based Deep Learning Models with BERT". *CoRR*. abs/1910.05786. arXiv: 1910.05786.

Tay, Y., V. Q. Tran, M. Dehghani, J. Ni, D. Bahri, H. Mehta, Z. Qin, K. Hui, Z. Zhao, J. P. Gupta, T. Schuster, W. W. Cohen, and D. Metzler. (2022). "Transformer Memory as a Differentiable Search Index". *CoRR*. abs/2202.06991. arXiv: 2202.06991.

Thakur, N., N. Reimers, A. Rücklé, A. Srivastava, and I. Gurevych. (2021). "BEIR: A Heterogeneous Benchmark for Zero-shot Evaluation of Information Retrieval Models". In: *Thirty-fifth Conference on Neural Information Processing Systems Datasets and Benchmarks Track (Round 2)*. URL: https://openreview.net/forum?id=wCu6T5xFjeJ.

Tishby, N. and N. Zaslavsky. (2015). "Deep learning and the information bottleneck principle". In: *2015 IEEE Information Theory Workshop (ITW)*. IEEE. DOI: 10.1109/itw.2015.7133169.

Vaswani, A., N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin. (2017). "Attention is All you Need". In: *Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, December 4-9, 2017, Long Beach, CA, USA*. 5998–6008. URL: https://proceedings.neurips.cc/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html.

Vinyals, O., A. Toshev, S. Bengio, and D. Erhan. (2015). "Show and tell: A neural image caption generator". In: *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE. DOI: 10.1109/cvpr.2015.7298935.

Voorhees, E. (2004). "Overview of the TREC 2004 Robust Retrieval Track". In: *TREC*.

Vulić, I. and M.-F. Moens. (2015). "Monolingual and Cross-Lingual Information Retrieval Models Based on (Bilingual) Word Embeddings". In: *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval. SIGIR '15*. Santiago, Chile: Association for Computing Machinery. 363–372. DOI: 10.1145/2766462.2767752.

Wang, H., X. Wang, W. Xiong, M. Yu, X. Guo, S. Chang, and W. Y. Wang. (2019). "Self-Supervised Learning for Contextualized Extractive Summarization". In: *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics. DOI: 10.18653/v1/p19-1214.

Wang, R., W. Liu, and C. McDonald. (2014). "Corpus-independent generic keyphrase extraction using word embedding vectors". In: *Software engineering research conference*. Vol. 39. 1–8.

Wang, W., B. Bi, M. Yan, C. Wu, J. Xia, Z. Bao, L. Peng, and L. Si. (2020). "StructBERT: Incorporating Language Structures into Pre-training for Deep Language Understanding". In: *8th International Conference on Learning Representations, ICLR 2020, Addis Ababa, Ethiopia, April 26-30, 2020*. OpenReview.net. URL: https://openreview.net/forum?id=BJgQ4lSFPH.

Wang, X., T. Gao, Z. Zhu, Z. Zhang, Z. Liu, J. Li, and J. Tang. (2021). "KEPLER: A Unified Model for Knowledge Embedding and Pre-trained Language Representation". *Transactions of the Association for Computational Linguistics*. 9: 176–194. DOI: 10.1162/tacl_a_00360.

Wu, H., W. Liu, L. Li, D. Nie, T. Chen, F. Zhang, and D. Wang. (2021). "UniKeyphrase: A Unified Extraction and Generation Framework for Keyphrase Prediction". In: *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*. Association for Computational Linguistics. DOI: 10.18653/v1/2021.findings-acl.73.

Wu, S. and M. Dredze. (2019). "Beto, Bentz, Becas: The Surprising Cross-Lingual Effectiveness of BERT". 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. DOI: 10.18653/v1/d19-1077.

Wu, Z., J. Mao, Y. Liu, J. Zhan, Y. Zheng, M. Zhang, and S. Ma. (2020). "Leveraging Passage-level Cumulative Gain for Document Ranking". *Proceedings of The Web Conference 2020*.

Xin, J., R. Nogueira, Y. Yu, and J. Lin. (2020a). "Early Exiting BERT for Efficient Document Ranking". In: *Proceedings of SustaiNLP: Workshop on Simple and Efficient Natural Language Processing*. Association for Computational Linguistics. DOI: 10.18653/v1/2020.sustainlp-1.11.

Xin, J., R. Tang, J. Lee, Y. Yu, and J. Lin. (2020b). "DeeBERT: Dynamic Early Exiting for Accelerating BERT Inference". In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics. DOI: 10.18653/v1/2020.acl-main.204.

Xin, J., C. Xiong, A. Srinivasan, A. Sharma, D. Jose, and P. N. Bennett. (2021). "Zero-Shot Dense Retrieval with Momentum Adversarial Domain Invariant Representations". *CoRR*. abs/2110.07581. arXiv: 2110.07581.

Xiong, C., Z. Dai, J. Callan, Z. Liu, and R. Power. (2017a). "End-to-End Neural Ad-hoc Ranking with Kernel Pooling". *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*.

Xiong, C., Z. Dai, J. Callan, Z. Liu, and R. Power. (2017b). "End-to-End Neural Ad-hoc Ranking with Kernel Pooling". In: *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM. DOI: 10.1145/3077136.3080809.

Xiong, L., C. Xiong, Y. Li, K. Tang, J. Liu, P. N. Bennett, J. Ahmed, and A. Overwijk. (2021). "Approximate Nearest Neighbor Negative Contrastive Learning for Dense Text Retrieval". In: *9th International Conference on Learning Representations, ICLR 2021, Virtual Event, Austria, May 3-7, 2021*. OpenReview.net. URL: https://openreview.net/forum?id=zeFrfgyZln.

Xu, S., X. Zhang, Y. Wu, F. Wei, and M. Zhou. (2020). "Unsupervised Extractive Summarization by Pre-training Hierarchical Transformers". In: *Findings of the Association for Computational Linguistics: EMNLP 2020*. Association for Computational Linguistics. DOI: 10.18653/v1/2020.findings-emnlp.161.

Yamada, I., A. Asai, and H. Hajishirzi. (2021). "Efficient Passage Retrieval with Hashing for Open-domain Question Answering". In: *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*. Association for Computational Linguistics. DOI: 10.18653/v1/2021.acl-short.123.

Yan, M., C. Li, B. Bi, W. Wang, and S. Huang. (2021). "A Unified Pretraining Framework for Passage Ranking and Expansion". In: *Thirty-Fifth AAAI Conference on Artificial Intelligence, AAAI 2021, Thirty-Third Conference on Innovative Applications of Artificial Intelligence, IAAI 2021, The Eleventh Symposium on Educational Advances in Artificial Intelligence, EAAI 2021, Virtual Event, February 2-9, 2021*. AAAI Press. 4555–4563. URL: https://ojs.aaai.org/index.php/AAAI/article/view/16584.

Yang, L., M. Zhang, C. Li, M. Bendersky, and M. Najork. (2020). "Beyond 512 Tokens: Siamese Multi-depth Transformer-based Hierarchical Encoder for Long-Form Document Matching". In: *Proceedings of the 29th ACM International Conference on Information and Knowledge Management*. ACM. DOI: 10.1145/3340531.3411908.

Yang, Z., Z. Dai, Y. Yang, J. G. Carbonell, R. Salakhutdinov, and Q. V. Le. (2019). "XLNet: Generalized Autoregressive Pretraining for Language Understanding". In: *Advances in Neural Information Processing Systems 32: Annual Conference on Neural Information Processing Systems 2019, NeurIPS 2019, December 8-14, 2019, Vancouver, BC, Canada*. Ed. by H. M. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. B. Fox, and R. Garnett. 5754–5764. URL: https://proceedings.neurips.cc/paper/2019/hash/dc6a7e655d7e5840e66733e9ee67cc69-Abstract.html.

Yin, D., Y. Hu, J. Tang, T. Daly, M. Zhou, H. Ouyang, J. Chen, C. Kang, H. Deng, C. Nobata, J.-M. Langlois, and Y. Chang. (2016). "Ranking Relevance in Yahoo Search". In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. ACM. DOI: 10.1145/2939672.2939677.

Yin, W. and Y. Pei. (2015). "Optimizing Sentence Modeling and Selection for Document Summarization". In: *Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence, IJCAI 2015, Buenos Aires, Argentina, July 25-31, 2015*. AAAI Press. 1383–1389. URL: http://ijcai.org/Abstract/15/199.

Yu, S., Z. Liu, C. Xiong, T. Feng, and Z. Liu. (2021). "Few-Shot Conversational Dense Retrieval". In: *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM. DOI: 10.1145/3404835.3462856.

Zamani, H. and W. B. Croft. (2016). "Embedding-based Query Language Models". In: *Proceedings of the 2016 ACM International Conference on the Theory of Information Retrieval*. ACM. DOI: 10.1145/2970398.2970405.

Zamani, H. and W. B. Croft. (2017). "Relevance-based Word Embedding". In: *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM. DOI: 10.1145/3077136.3080831.

Zamani, H. and W. B. Croft. (2018). "On the Theory of Weak Supervision for Information Retrieval". In: *Proceedings of the 2018 ACM SIGIR International Conference on Theory of Information Retrieval*. ACM. DOI: 10.1145/3234944.3234968.

Zamani, H., W. B. Croft, and J. S. Culpepper. (2018a). "Neural Query Performance Prediction using Weak Supervision from Multiple Signals". In: *The 41st International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM. DOI: 10.1145/3209978.3210041.

Zamani, H., M. Dehghani, W. B. Croft, E. Learned-Miller, and J. Kamps. (2018b). "From Neural Re-Ranking to Neural Ranking: Learning a Sparse Representation for Inverted Indexing". In: *Proceedings of the 27th ACM International Conference on Information and Knowledge Management. CIKM '18*. Torino, Italy: Association for Computing Machinery. 497–506. DOI: 10.1145/3269206.3271800.

Zhai, C. (2007). "Statistical Language Models for Information Retrieval A Critical Review". *Foundations and Trends® in Information Retrieval*. 2(3): 137–213. DOI: 10.1561/1500000008.

Zhai, C. and J. Lafferty. (2004). "A study of smoothing methods for language models applied to information retrieval". *ACM Transactions on Information Systems*. 22(2): 179–214. DOI: 10.1145/984321.984322.

Zhan, J., J. Mao, Y. Liu, J. Guo, M. Zhang, and S. Ma. (2021a). "Jointly Optimizing Query Encoder and Product Quantization to Improve Retrieval Performance". In: *Proceedings of the 30th ACM International Conference on Information and Knowledge Management*. ACM. DOI: 10.1145/3459637.3482358.

Zhan, J., J. Mao, Y. Liu, J. Guo, M. Zhang, and S. Ma. (2021b). "Optimizing Dense Retrieval Model Training with Hard Negatives". In: *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM. DOI: 10.1145/3404835.3462880.

Zhan, J., J. Mao, Y. Liu, J. Guo, M. Zhang, and S. Ma. (2022). "Learning Discrete Representations via Constrained Clustering for Effective and Efficient Dense Retrieval". In: *Proceedings of the Fifteenth ACM International Conference on Web Search and Data Mining*. ACM. DOI: 10.1145/3488560.3498443.

Zhan, J., J. Mao, Y. Liu, M. Zhang, and S. Ma. (2020a). "Learning To Retrieve: How to Train a Dense Retrieval Model Effectively and Efficiently". *CoRR*. abs/2010.10469. arXiv: 2010.10469.

Zhan, J., J. Mao, Y. Liu, M. Zhang, and S. Ma. (2020b). "RepBERT: Contextualized Text Embeddings for First-Stage Retrieval". *CoRR*. abs/2006.15498. arXiv: 2006.15498.

Zhang, H., H. Shen, Y. Qiu, Y. Jiang, S. Wang, S. Xu, Y. Xiao, B. Long, and W.-Y. Yang. (2021a). "Joint Learning of Deep Retrieval Model and Product Quantization based Embedding Index". In: *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM. DOI: 10.1145/3404835.3462988.

Zhang, H., H. Shen, Y. Qiu, Y. Jiang, S. Wang, S. Xu, Y. Xiao, B. Long, and W.-Y. Yang. (2021b). "Joint Learning of Deep Retrieval Model and Product Quantization based Embedding Index". In: *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM. DOI: 10.1145/3404835.3462988.

Zhang, H., J. Cai, J. Xu, and J. Wang. (2019a). "Pretraining-Based Natural Language Generation for Text Summarization". In: *Proceedings of the 23rd Conference on Computational Natural Language Learning (CoNLL)*. Association for Computational Linguistics. DOI: 10.18653/v1/k19-1074.

Zhang, J., Y. Zhao, M. Saleh, and P. Liu. (2020a). "Pegasus: Pre-training with extracted gap-sentences for abstractive summarization". In: *International Conference on Machine Learning*. PMLR. 11328–11339.

Zhang, K., C. Xiong, Z. Liu, and Z. Liu. (2020b). "Selective Weak Supervision for Neural Information Retrieval". In: *Proceedings of The Web Conference 2020*. ACM. DOI: 10.1145/3366423.3380131.

Zhang, X., F. Wei, and M. Zhou. (2019b). "HIBERT: Document Level Pre-training of Hierarchical Bidirectional Transformers for Document Summarization". In: *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics. DOI: 10.18653/v1/p19-1499.

Zhang, Z., X. Han, Z. Liu, X. Jiang, M. Sun, and Q. Liu. (2019c). "ERNIE: Enhanced Language Representation with Informative Entities". In: *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics. DOI: 10.18653/v1/p19-1139.

Zheng, G. and J. Callan. (2015). "Learning to Reweight Terms with Distributed Representations". In: *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval. SIGIR '15*. Santiago, Chile: Association for Computing Machinery. 575–584. DOI: 10.1145/2766462.2767700.

Zheng, Z., K. Hui, B. He, X. Han, L. Sun, and A. Yates. (2020). "BERT-QE: Contextualized Query Expansion for Document Re-ranking". In: *Findings of the Association for Computational Linguistics: EMNLP 2020*. Association for Computational Linguistics. DOI: 10.18653/v1/2020.findings-emnlp.424.

Zheng, Z., K. Hui, B. He, X. Han, L. Sun, and A. Yates. (2021). "Contextualized query expansion via unsupervised chunk selection for text retrieval". *Information Processing and Management*. 58(5): 102672. DOI: 10.1016/j.ipm.2021.102672.

Zhong, M., P. Liu, Y. Chen, D. Wang, X. Qiu, and X. Huang. (2020). "Extractive Summarization as Text Matching". In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics. DOI: 10.18653/v1/2020.acl-main.552.

Zhong, M., P. Liu, D. Wang, X. Qiu, and X. Huang. (2019). "Searching for Effective Neural Extractive Summarization: What Works and What's Next". In: *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics. DOI: 10.18653/v1/p19-1100.

Zhou, Y., Z. Dou, B. Wei, R. Xie, and J.-R. Wen. (2021a). "Group based Personalized Search by Integrating Search Behaviour and Friend Network". In: *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval.* ACM. DOI: 10.1145/3404835.3462918.

Zhou, Y., Z. Dou, and J.-R. Wen. (2020). "Encoding History with Context-aware Representation Learning for Personalized Search". In: *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval.* ACM. DOI: 10.1145/3397271.3401175.

Zhou, Y., Z. Dou, Y. Zhu, and J.-R. Wen. (2021b). "PSSL: Self-supervised Learning for Personalized Search with Contrastive Sampling". In: *Proceedings of the 30th ACM International Conference on Information and Knowledge Management.* ACM. DOI: 10.1145/3459637.3482379.

Zhou, Y., J. Yao, Z. Dou, L. Wu, and J. Wen. (2022). "DynamicRetriever: A Pre-training Model-based IR System with Neither Sparse nor Dense Index". *CoRR*. abs/2203.00537. DOI: 10.48550/arXiv.2203.00537. arXiv: 2203.00537.

Zhu, H., L. Dong, F. Wei, B. Qin, and T. Liu. (2019). "Transforming Wikipedia into Augmented Data for Query-Focused Summarization". *CoRR*. abs/1911.03324. arXiv: 1911.03324.

Zhu, Y., J.-Y. Nie, Z. Dou, Z. Ma, X. Zhang, P. Du, X. Zuo, and H. Jiang. (2021). "Contrastive Learning of User Behavior Sequence for Context-Aware Document Ranking". In: *Proceedings of the 30th ACM International Conference on Information and Knowledge Management.* ACM. DOI: 10.1145/3459637.3482243.

Zou, Y., X. Zhang, W. Lu, F. Wei, and M. Zhou. (2020). "Pre-training for Abstractive Document Summarization by Reinstating Source Text". In: *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP).* Association for Computational Linguistics. DOI: 10.18653/v1/2020.emnlp-main.297.

Zuccon, G., B. Koopman, P. Bruza, and L. Azzopardi. (2015). "Integrating and Evaluating Neural Word Embeddings in Information Retrieval". In: *Proceedings of the 20th Australasian Document Computing Symposium.* ACM. DOI: 10.1145/2838931.2838936.