Teaching Information Retrieval with a Shared Task Across Universities: First Steps and Findings

Maik Fröbe[6], Christopher Akiki[1,2], Timo Breuer[3], Thomas Eckart[4], Annemarie Friedrich[5], Lukas Gienapp[1,2], Jan Heinrich Merker[6], Martin Potthast[2,7,8], Harrisen Scells[1], Philipp Schaer[3] and Benno Stein[9]

[1]*Leipzig University,* [2]*ScaDS.AI,* [3]*TH Köln,* [4]*Sächsische Akademie der Wissenschaften zu Leipzig,* [5]*University of Augsburg,* [6]*Friedrich-Schiller-Universität Jena,* [7]*University of Kassel,* [8]*hessian.AI,* [9]*Bauhaus-Universität Weimar*

Abstract

Many universities offer information retrieval (IR) courses with different specializations as part of their computer science or information science programs. Student involvement and collaboration in these courses can increase engagement in the course and improve learning outcomes. We report on our first steps towards creating synergies between information retrieval courses at four German universities by conducting a shared task assignment with a document collection combining the IR and ACL anthologies, which we enrich with relevance judgments. We prepared two versions of this collection. First, a minimal test collection with 100 documents for which students can manually obtain results from intermediate results. Second, a more extensive test collection of 126 958 documents is used in a shared task setup, where students create topics, relevance judgments, and retrieval systems. The shared task setup therefore covers a broad spectrum of applied IR research. Our collaborative teaching initiative can help students learn from their peers locally and across universities. The cross-university setup means that institutes and degree programs with different academic backgrounds are involved, which leads to a broad spectrum of perspectives in the construction of topics and also in system development.

Keywords

Open Educational Resources, Shared Tasks, IR Evaluation, Corpus Construction

1. Overview

The information retrieval (IR) landscape has changed dramatically with the introduction of the Transformer technology [1] and the subsequent evolution to the Retrieval Augmented Generation paradigm [2]. In 2023, at a one-day workshop organized by the German Special Interest Group on Information Retrieval[1] in Cologne, we brainstormed on how to reflect these drastic changes in our IR teaching material, also considering related efforts [3, 4, 5, 6, 7].

The basic objectives of IR teaching remain the same, i.e., they depend on the background of the students and range from the professional use of retrieval systems and the assessment of the suitability of systems for a specific task to the development and improvement of existing systems. However, our analysis has shown that there is a lack of shared open resources that can be used by all universities. In particular, we are not aware of any collections that are specifically tailored to IR teaching, where small collections are required that highlight important concepts while being available under a permissive license. To address this shortcoming, we conducted a cross-university shared task in the summer semester 2024, in which four IR courses participated (held at the University of Augsburg, the Friedrich-Schiller-Universität Jena, the TH Köln, and Leipzig University). We report on our experiences with this shared task and its current outcomes, which can form the basis for shared teaching resources, and we ask the community for feedback for future developments.

The primary goal of our shared task is to familiarize students with IR technology and research methods. The secondary goal is to establish and expand an IR collection specifically tailored to teaching. As document collection, we use the IR Anthology [8] and the ACL Anthology [9], 126 958 documents combined (40MB compressed). We provide two versions: (1) the collection ir_teach_100, which consists of 100 documents, and (2) the collection ir_teach_full, which contains all documents. We make a

LWDA'24: Lernen, Wissen, Daten, Analysen. September 23–25, 2024, Würzburg, Germany

[1] https://fg-retrieval.gi.de/

browser-based search engine for the smaller ir_teach_100 collection publicly available,[2] which enables students without programming skills to carry out experiments. The ir_teach_full collection is intended for programming-based processing. This collection is also integrated into ir_datasets [10] to provide students with simplified access via a standardized API. For the shared task, we follow a predefined teaching workflow in which students evaluate their own systems using their annotations [11]: First, students create topics describing information needs. For each topic, we obtain a pool of documents from a set of 10 different baseline retrieval systems, for which the students provide relevance judgments. We then extract a set of training topics an make them available to the students for the development of their own retrieval systems. In total, 154 topics have been created in this way for our document collection (68 topics last semester, and 86 topics this semester).

Students submit their retrieval systems to TIRA/TIREx [12, 13], where the test topics are included as a private test set and the results are published at the end of the course. Overall, the students integrated various techniques into their retrieval systems, including traditional approaches such as query expansion, rank fusion, feature-based learning-to-rank, but also more recent approaches such as DocT5Query or BERT-based re-rankers. Some of the student contributions even outperformed state-of-the-art rankers such as RankZephyr. For presentation and review purposes, we developed a web-based application where students and instructors can browse their information needs, the systems developed, and the resulting leaderboards.[3]

2. Conclusions and Future Work

We present the results of a student-oriented shared task for IR courses at four universities in Germany. After an overview of results obtained in the past two terms, we see many opportunities for future extension. An important open question is how we can consolidate our resources and tutorials (e.g., on transformer model training) so that they can be reused in different courses. We have also received feedback that the IR/ACL anthology might not be an ideal teaching-oriented document collection, since it requires domain-specific knowledge, and we are deliberating other options. We invite IR teachers to join our initiative and to discuss the future of teaching IR with shared tasks across universities.

Acknowledgments

We would like to thank all students for their active participation in our courses. This includes students from the University of Augsburg (M.Sc. in Computer Science), the Friedrich-Schiller-Universität Jena (M.Sc. in Computer Science), TH Köln (B.Sc. in Information Science), and Leipzig University (B.Sc. in Computer Science and Digital Humanities).

References

[1] J. Lin, R. F. Nogueira, A. Yates, Pretrained Transformers for Text Ranking: BERT and Beyond, Synthesis Lectures on Human Language Technologies, Morgan & ClaypoolPublishers, 2021. doi:10.2200/S01123ED1V01Y202108HLT053.

[2] Y. Gao, Y. Xiong, X. Gao, K. Jia, J. Pan, Y. Bi, Y. Dai, J. Sun, Q. Guo, M. Wang, H. Wang, Retrieval-augmented generation for large language models: A survey, CoRR abs/2312.10997 (2023). doi:10.48550/ARXIV.2312.10997. arXiv:2312.10997.

[3] D. Blank, N. Fuhr, A. Henrich, T. Mandl, T. Rölleke, H. Schütze, B. Stein, Information retrieval: Concepts and practical considerations for teaching a rising topic, Datenbank-Spektrum 9 (2009) 30–41.

[2] https://irgroup.github.io/ir-teach/

[3] https://tira-io.github.io/ir-lab-sose-24/topics

[4] D. Blank, N. Fuhr, A. Henrich, T. Mandl, T. Rölleke, H. Schütze, B. Stein, Teaching IR: curricular considerations, in: E. N. Efthimiadis, J. M. Fernández-Luna, J. F. Huete, A. MacFarlane (Eds.), Teaching and Learning in Information Retrieval, volume 31 of *The Information Retrieval Series*, Springer, 2011, pp. 31–46. URL: https://doi.org/10.1007/978-3-642-22511-6_3. doi:10.1007/978-3-642-22511-6_3.

[5] I. Markov, M. de Rijke, What should we teach in information retrieval?, SIGIR Forum 52 (2018) 19–39. URL: https://doi.org/10.1145/3308774.3308780. doi:10.1145/3308774.3308780.

[6] C. Bauer, B. Carterette, N. Ferro, N. Fuhr, J. Beel, T. Breuer, C. L. A. Clarke, A. Crescenzi, G. Demartini, G. M. D. Nunzio, L. Dietz, G. Faggioli, B. Ferwerda, M. Fröbe, M. Hagen, A. Hanbury, C. Hauff, D. Jannach, N. Kando, E. Kanoulas, B. P. Knijnenburg, U. Kruschwitz, M. Li, M. Maistro, L. Michiels, A. Papenmeier, M. Potthast, P. Rosso, A. Said, P. Schaer, C. Seifert, D. Spina, B. Stein, N. Tintarev, J. Urbano, H. Wachsmuth, M. C. Willemsen, J. Zobel, Report on the Dagstuhl seminar on frontiers of information access experimentation for research and education, SIGIR Forum 57 (2023) 7:1–7:28. doi:10.1145/3636341.3636351.

[7] C. Bauer, M. Fröbe, D. Jannach, U. Kruschwitz, P. Rosso, D. Spina, N. Tintarev, Overcoming methodological challenges in information retrieval and recommender systems through awareness and education, arXiv 2305.01509, 2023. doi:10.48550/arXiv.2305.01509.

[8] M. Potthast, S. Günther, J. Bevendorff, J. P. Bittner, A. Bondarenko, M. Fröbe, C. Kahmann, A. Niekler, M. Völske, B. Stein, M. Hagen, The information retrieval anthology, in: F. Diaz, C. Shah, T. Suel, P. Castells, R. Jones, T. Sakai (Eds.), SIGIR '21: The 44th International ACM SIGIR Conference on Research and Development in Information Retrieval, Virtual Event, Canada, July 11-15, 2021, ACM, 2021, pp. 2550–2555. URL: https://doi.org/10.1145/3404835.3462798. doi:10.1145/3404835.3462798.

[9] S. Bird, R. Dale, B. J. Dorr, B. R. Gibson, M. T. Joseph, M. Kan, D. Lee, B. Powley, D. R. Radev, Y. F. Tan, The ACL anthology reference corpus: A reference dataset for bibliographic research in computational linguistics, in: Proceedings of the International Conference on Language Resources and Evaluation, LREC 2008, 26 May - 1 June 2008, Marrakech, Morocco, European Language Resources Association, 2008, pp. 1755–1759.

[10] S. MacAvaney, A. Yates, S. Feldman, D. Downey, A. Cohan, N. Goharian, Simplified data wrangling with ir_datasets, in: F. Diaz, C. Shah, T. Suel, P. Castells, R. Jones, T. Sakai (Eds.), SIGIR '21: The 44th International ACM SIGIR Conference on Research and Development in Information Retrieval, Virtual Event, Canada, July 11-15, 2021, ACM, 2021, pp. 2429–2436. URL: https://doi.org/10.1145/3404835.3463254. doi:10.1145/3404835.3463254.

[11] M. Fröbe, H. Scells, T. Elstner, C. Akiki, L. Gienapp, J. H. Reimer, S. MacAvaney, B. Stein, M. Hagen, M. Potthast, Resources for Combining Teaching and Research in Information Retrieval Courses, in: G. H. Yang, H. Wang, S. Han, C. Hauff, G. Zuccon, Y. Zhang (Eds.), 47th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2024), ACM, 2024, pp. 1115–1125. doi:10.1145/3626772.3657886.

[12] M. Fröbe, M. Wiegmann, N. Kolyada, B. Grahm, T. Elstner, F. Loebe, M. Hagen, B. Stein, M. Potthast, Continuous Integration for Reproducible Shared Tasks with TIRA.io, in: J. Kamps, L. Goeuriot, F. Crestani, M. Maistro, H. Joho, B. Davis, C. Gurrin, U. Kruschwitz, A. Caputo (Eds.), Advances in Information Retrieval. 45th European Conference on IR Research (ECIR 2023), Lecture Notes in Computer Science, Springer, Berlin Heidelberg New York, 2023, pp. 236–241. doi:10.1007/978-3-031-28241-6_20.

[13] M. Fröbe, J. H. Reimer, S. MacAvaney, N. Deckers, S. Reich, J. Bevendorff, B. Stein, M. Hagen, M. Potthast, The Information Retrieval Experiment Platform, in: H. Chen, W. E. Duh, H. Huang, M. P. Kato, J. Mothe, B. Poblete (Eds.), 46th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2023), ACM, 2023, pp. 2826–2836. doi:10.1145/3539618.3591888.