

Natural language search queries in zbMATH Open

Olaf Teschke

In June 2026, zbMATH Open has released an experimental search option facilitating the translation of natural language queries into faceted search commands, taking full advantage of the traditional precise query language without requiring knowledge of its details.

1 Introduction

Traditionally, zbMATH Open has offered two main access modes to its content: a one-line search function which allows both for simple searches and for a sophisticated syntax combining full logical expressions of now more than twenty fields, and a multi-line search providing a more traditional approach to combine these fields. The latter approach serves mainly as a reduced entrance barrier to the multitude of search dimensions. The various filter functions which can be applied to narrow results serve the same purpose. Overall, the guiding design philosophy is to facilitate both exhaustive and precise exploration instead of just convenient heuristics.

While expert users are usually aware of the core syntax and often employ complex pre-defined queries to obtain updates on their areas of interest, one cannot expect that all users have the full range of the ever-growing available functions in mind. E.g., information on arXiv versions [1], Open Access and licenses [4], or connections to Wikipedia [6] or research data like DLMF [2], OEIS [3], and the ATLAS of group representations have been added recently. At the same time, search habits change due to trends on other platforms, and may differ significantly among generations. Today, one may not always expect that the exact spelling of an author's name is needed, or that a precise Mathematics Subject Classification code is known for area search.

Hence, it is natural to employ recent advances in Artificial Intelligence (AI) – namely, the rapid progress of large language models (LLMs) – to overcome these barriers.

2 Large language models as an interpreting tool for zbMATH Open – benefits and limitations

LLMs now appear to be ubiquitous, and may be employed for very different purposes. For example, as described in an earlier article in

this column [5], two LLMs have been very useful to overcome the long-lasting difficulties in translating scans of historical mathematical reviews into $\text{\LaTeX}$ and properly assigning them to metadata.

Still, the prevailing LLM application today appears to be a chatbot-like usage, which has also turned out to be helpful for math exploration – e.g., as a dialogue tool in pursuing proofs. However, such an installation at zbMATH Open appears not quite feasible; the main purpose of the service is not to create proofs, but to help identify what has been achieved, and what might be worth further investigation. Therefore, precision and comprehensiveness have been traditionally considered to be more important than convenient heuristics.

Instead, the new application of LLMs to zbMATH Open has a more modest goal: to allow for natural-language search queries without requiring full expert knowledge of the query syntax. The tool has been made available as a switchable experimental option at the One-Line Search for zbMATH Open documents in June 2026. When enabled, queries are translated into a Boolean combination of facet searches of the zbMATH query language that appear to fit the original question best, as in Figure 1, where the LLM correctly identifies the arXiv facet, the publication year range, and the content fields for a phrase from a natural language query.

There are some obvious advantages of this approach. Firstly, it reflects modern search customs, and overcomes the barrier of learning the query language. Even for expert users, it may no longer be necessary to keep track of all new developments and functions, as long as they are covered by appropriate prompt engineering.

Moreover, this usage allows users to avoid the mentioned problematic issues related to LLMs. No data from zbMATH Open needs to be ingested into the models, therefore avoiding license issues. Instead, the translation uses only appropriate prompt engineering based on general help and description texts for the zbMATH Open usage.

In addition, the technical implementation (see Section 3) also allows for a solution where no user input is employed, e.g., for training issues, and privacy is guaranteed.

On the other hand, users should be aware of the limitations of this feature. Naturally, only math-related queries which can be reasonably mapped to zbMATH Open fields will result in meaningful answers. Moreover, one cannot expect full reproducibility: Even

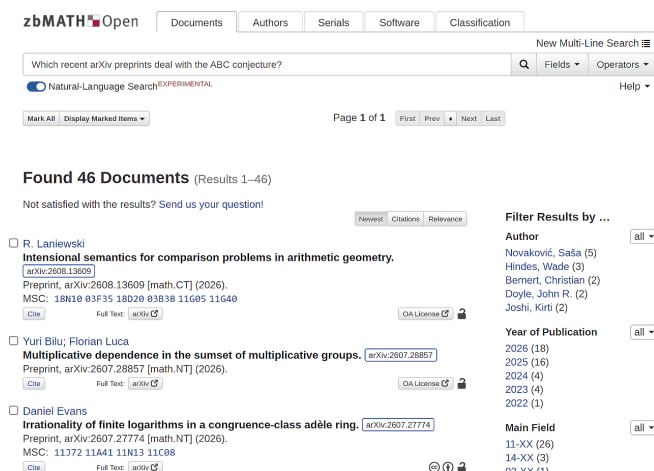


Figure 1. The query “Which recent arXiv preprints deal with the ABC conjecture?” is translated to “db:arxiv & py:2021-2026 & (ti:‘ABC conjecture’ | ab:‘ABC conjecture’ | ut:‘ABC conjecture’).”

the same language model will frequently produce different output for the same question, and the underlying LLM will be updated continuously to improve performance parameters (see Section 3).

Additionally, lacking billion-dollar resources of commercial AI companies, the service can only provide a limited amount of queries per specified units of time. When rate limits are reached, users may be forced to revert to traditional query types.

3 Technical implementation, models, and benchmarks

At a technical level, the natural language search is supported by the KISSKI LLM Service “Chat AI” provided by GWDG.¹ It offers a broad spectrum of recent open-weight models, both for web and API access. As mentioned, and perhaps most importantly, it guarantees no server-side storage of conversations. User input is not used for AI training or similar purposes, thereby avoiding any privacy conflicts.²

The list of currently available models can be found online³; versions and models are updated continuously. Experience from the test phase indicates that it makes sense to update the model underlying the zbMATH natural language search every few months, depending on the releases and the performance relative to the benchmarks.

For the initial benchmark suite, forty natural-language queries from typical questions occurring in the work of zbMATH section editors have been tested three times using several adapted prompts



Figure 2. LLMs allow also for typo correction in queries like “Which bok of Daubchies contains her lektures on wavlets?”

for the models, and evaluated with respect to robustness, output quality, performance, and token resources (with particular emphasis on the first two aspects).

At the time of writing, three models emerged at the top – Mistral Large Instruct 3 675B Instruct 2512, Qwen 3 32B, and OpenAI GPT OSS 120B. For all prompt versions, Mistral was by far the top model with respect to robustness and response time. The quality of the results varied depending on the prompts, but could be significantly improved for Mistral by prompt engineering to also beat the other models in this respect in the current setup. Hence, for the initial release, the Mistral Large Instruct 3 675B Instruct 2512 has been chosen (though at the time of publication, another model may be already in use).

4 Conclusions and future developments

The new query option intrinsically comes with some advantages which can be further exploited in the future. E.g., the LLM features allow for a reasonably good typo correction as in Figure 2.

Likewise, the models usually support (to a varying degree) multi-language queries, so English is no longer strictly necessary as input.

Obviously, it is not easy to foresee how effectively and successfully the natural language search will be adopted by zbMATH users. The further development depends very much upon the feedback, which can be provided through a specific contact form. In case of interest, we could also extend the current set of internal benchmark queries to a community-maintained open benchmark suite which helps keep pace with new LLM generations.

Currently, the natural language search has only been implemented for the document search, for the simple reason that it offers by far the most extensive search options, and the output is most helpful. Experiments with other search facets (like authors,

¹ <https://kisski.gwdg.de/en/leistungen/2-02-llm-service/>

² <https://docs.hpc.gwdg.de/services/ai-services/chat-ai/faq/>

³ <https://docs.hpc.gwdg.de/services/ai-services/chat-ai/models/>

journals, or software) give no clear indication whether it might be useful to implement a similar feature through adapted prompts. Instead, it might be more helpful to create a large unified index of these facets combined with document search. Obviously, this would also require much more fine-tuning of the LLM layer, and it makes sense to gain initial experience with the release of the search for documents only. Such an LLM-supported unified search would obviously also be convenient for mobile versions of the database.

Acknowledgements. The author would like to thank Christian Bär, Andreas Kühnemund, and Fabian Müller for many ideas, discussions, implementations, and tests which made the new feature possible and formed the basis of this note.

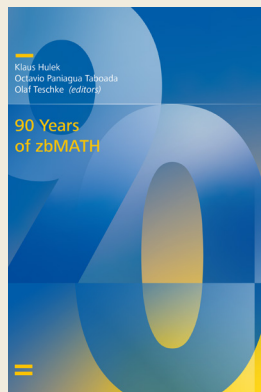
References

- [1] I. Beckenbach, K. Hulek and O. Teschke, [The extension of zbMATH open by arXiv preprints](#). *Eur. Math. Soc. Mag.* **131**, 54–58 (2024)
- [2] H. S. Cohl, O. Teschke and M. Schubotz, [Connecting islands: bridging zbMATH and DLMF with Scholix, a blueprint for connecting expert knowledge systems](#). *Eur. Math. Soc. Mag.* **120**, 66–67 (2021)
- [3] D. Ehsani, M. Petrera and O. Teschke, [The integration of OEIS links in zbMATH Open](#). *Eur. Math. Soc. Mag.* **128**, 60–64 (2023)
- [4] D. Ehsani and O. Teschke, [Open access status and licences in zbMATH Open](#). *Eur. Math. Soc. Mag.* **135**, 61–64 (2025)
- [5] B. Gipp, I. Pluzhnikov, M. Schubotz and O. Teschke, [Digitisation of scanned reviews from zbMATH Open](#). *Eur. Math. Soc. Mag.* **140**, 40–43 (2026)
- [6] H. Rahkooy, M. Schubotz, O. Teschke, M. Fuhrmann, N. Roy and D. Mietchen, [Wikipedia and zbMATH Open: connecting several layers of mathematical information](#). *Eur. Math. Soc. Mag.* **136**, 59–63 (2025)

Olaf Teschke studied mathematics at Humboldt University Berlin, and completed his PhD in algebraic geometry there. He moved to FIZ Karlsruhe in 2008, and has been working there since 2009 as the head of the Mathematics Department and managing editor of zbMATH (including a short intermediate term as editor-in-chief). He has also served in several EMS committee positions. His main occupation is information infrastructure for mathematics.

olaf.teschke@fiz-karlsruhe.de

EMS Press book



90 Years of zbMATH

Edited by
Klaus Hulek
(Leibniz University Hannover)
Octavio Paniagua Taboada
(FIZ Karlsruhe)
Olaf Teschke
(FIZ Karlsruhe)

ISBN 978-3-98547-073-0
eISBN 978-3-98547-573-5

2024. Softcover. 110 pages
€ 29.00

zbMATH Open, the world's most comprehensive and longest-running abstracting and reviewing service in pure and applied mathematics was founded by Otto Neugebauer in 1931. It celebrated its 90th anniversary by becoming an open access database. In December 2019, the Joint Science Conference (Gemeinsame Wissenschaftskonferenz) agreed that the Federal and State Governments of Germany would support FIZ Karlsruhe in transforming zbMATH into an open platform. In future, zbMATH Open will link mathematical services and platforms so as to provide considerably more content for further research and collaborative work in mathematics and related fields.

This book presents how zbMATH Open has reacted to a rapidly changing digital era. Topics covered include: the linkage of zbMATH Open with different community platforms and digital maths libraries, the use of zbMATH Open as a bibliographical tool, API solutions, current advancements in author profiles, the indexing of mathematical software packages (swMATH), and issues concerning mathematical formula search in zbMATH Open. We also reflect on the gender publication gap in mathematics, and focus on one of the central pillars of zbMATH Open: the community of reviewers.

**20% discount on any book purchases for individual members of the EMS, member societies or societies with a reciprocity agreement when ordering directly from EMS Press.*

EMS Press is an imprint of the European Mathematical Society – EMS – Publishing House GmbH
Straße des 17. Juni 136 | 10623 Berlin | Germany
<https://ems.press> | orders@ems.press



ADVERTISEMENT