

The integration of OEIS links in zbMATH Open

Dariusz Ehsani, Matteo Petrera and Olaf Teschke

The transition towards an Open Data Platform enabled zbMATH Open to build a network of open resources. Important components in the evolving information system are mathematical research data, which are of quite heterogeneous nature. For their interlinking, zbMATH Open provides Application Programming Interface (API) solutions to offer mathematical research data to the community. Among other APIs recently implemented at zbMATH Open, the so-called Links API is aimed to document interconnections between our database and external platforms which display mathematical literature indexed at zbMATH Open. The Digital Library of Mathematical Functions (DLMF) has been our first partner and their data have been integrated in our database in 2021. Recently we interlinked with the second platform, the Online Encyclopedia of Integer Sequences (OEIS), a renowned digital database of number sequences that cites a lot of mathematical literature, especially from number theory and graph theory. The purpose of this short contribution is to announce and discuss the links to OEIS data in zbMATH Open.

1 Introduction

As outlined in [3, 4], zbMATH Open¹ is currently transformed into an information system that connects a broad variety of resources relevant for mathematics research, facilitated by the new opportunities provided in the framework of Open Access and Open Data. Mathematical Research Data are an essential additional information layer in such a network. Our commitment, in addition to being diversified on various development fronts, always aims to be exhaustive and publicly accessible, thus offering a complete, easily usable and open service. Our ultimate goal is to consolidate zbMATH Open as a solid reference point and a modern research tool for the entire scientific community.

Recently we developed Application Programming Interface (API) solutions to facilitate and optimize the open access to our mathematical research data. We already documented some of our

APIs in previous publications, see, e.g., [5, 6]. At present we have three distinct APIs, an Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) API,² designed to harvest the entire zbMATH Open dataset or some specific subsets of it, a Representational State Transfer (REST) API,³ still in the staging phase, and the so-called Links API,⁴ designed to document the interconnections between zbMATH Open and external digital platforms with mathematical contents.

The motivation behind the implementation of APIs at zbMATH Open is twofold. On the one hand, we want to provide the community with machine-readable tools to benefit from the open access of our data. On the other hand, we wish to expose and document the dynamic interactions between our data and those coming from other digital resources used by the community.

Among the potential users of our API endpoints are

1. bibliographic consumers (e.g., MathOverflow, Wikimedia) displaying references to scientific publications;
2. aggregators (e.g., research data infrastructures, Semantic Scholar) extracting data;
3. archives (e.g., research/software digital archives) storing flows of data;
4. search engines (e.g., Google) implementing the OpenSearch standard;
5. researchers consulting literature for research purposes.

In this short note we announce the integration into the Links API of our second partner, the Online Encyclopedia of Integer Sequences (OEIS). In Section 2 we briefly discuss the main purpose and functionalities of this API. Section 3 will be focused on the role of the OEIS in our database. The concluding Section 4 is devoted to a discussion of future perspectives.

² <https://oai.zbmhath.org>

³ <https://donald.zentralblatt-math.org/zbmath-api-test/docs>

⁴ https://donald.zentralblatt-math.org/linksapi-test/links_api

¹ <https://zbmath.org>

2 The Links API

The Links API has been developed between 2021 and 2022 to document the interconnections (more specifically, the *links*) between zbMATH Open and external platforms (called *partners*) which display and use mathematical bibliographic data (see Figure 1). As a matter of fact, there exist several well-established digital services serving the mathematical community which cite documents indexed in our database. Among such platforms we mention MathOverflow,⁵ the Online Encyclopedia of Integer Sequences,⁶ the Digital Library of Mathematical Functions⁷ and arXiv.⁸ In addition to having retrieved the bibliographic data of the matched documents from the platforms themselves, we have considered making the links to these digital resources directly accessible from our web page. We take this as a valid support for anyone using zbMATH Open for research purposes. All of this has been made possible thanks to the implementation of the Links API.

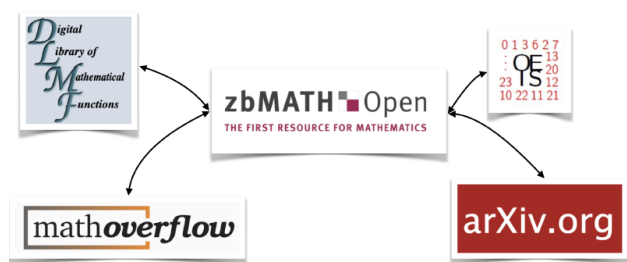


Figure 1. zbMATH Open interconnected with external platforms.

The current partners of the Links API are:

1. The Digital Library of Mathematical Functions (DLMF); the DLMF is a well-established web resource that enlarges and digitally translates the classical “Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables,”⁹ edited by M. Abramowitz and I. A. Stegun in 1964. The DLMF presents its contents in 36 chapters, and the bibliography currently consists of almost 2,800 references,¹⁰ of which about 75 % directly link to zbMATH Open.¹¹ For more details about the integration of the DLMF data into zbMATH Open we refer to [1, 5].

2. The Online Encyclopedia of Integer Sequences (OEIS); the OEIS is an online database of sequences of numbers. Please see Section 3.

The datasets coming from our partners are automatically updated in our database every six months. In the future we plan to integrate into the Links API our already existing datasets for MathOverflow and arXiv.

2.1 Response body and endpoints of the Links API

The JSON response body of the Links API is modeled on the Scholix metadata schema.¹² The models used to pack the data are explicitly reported in the API web interface. Scholix is a well-established framework to exchange information between data and literature links. The architecture of the schema is designed to allow for bulk exchange of link information, which contains all necessary data to keep track of bibliographic parameters identifying scholarly links.

The current version of the API offers twelve endpoints (seven GET routes, one DELETE route, two POST routes, one PATCH route and one PUT route):

- GET `/link`. It retrieves links for given zbMATH objects.
- DELETE `/link/item`. It deletes a link from the database.
- POST `/link/item`. It creates a new link related to a zbMATH object.
- GET `/link/item`. It checks relations between a given link and a given zbMATH object.
- PATCH `/link/item`. It edits an existing link.
- GET `/link/item/{doc_id}`. It retrieves links for a given zbMATH object.
- GET `/partner`. It retrieves data of a given zbMATH partner.
- PUT `/partner`. It edits data of a given zbMATH partner.
- POST `/partner`. It creates a new partner related to zbMATH.
- GET `/source`. It produces a list of all links of a given zbMATH partner.
- GET `/statistics/msc`. It shows the occurrence of primary MSC codes¹³ (2-digit level) of zbMATH objects in the set of links of a given partner.
- GET `/statistics/year`. It shows the occurrence of years of publication of zbMATH objects in the set of links of a given partner.

These allow a diversified and targeted search thanks to the use of appropriate filters.

In Figure 2 we see an example of an excerpt of the JSON response body. It corresponds to the output coming from the search with input field `author=Levine` in the endpoint GET `/link`. As a result, it turns out that in the OEIS sequence A104246,¹⁴ called

⁵ <https://mathoverflow.net>

⁶ <https://oeis.org>

⁷ <https://dlmf.nist.gov>

⁸ <https://arxiv.org>

⁹ <https://zbmath.org/0171.38503>

¹⁰ <https://dlmf.nist.gov/bib>

¹¹ The remaining 25 % of publications not linked to zbMATH Open refers to documents not indexed in the zbMATH Open database.

¹² <http://www.scholix.org/schema/3-0>

¹³ Mathematics Subject Classification scheme 2020, <https://msc2020.org>

¹⁴ <https://oeis.org/A104246>

“Minimal number of tetrahedral numbers (A000292(k) = $k(k+1)(k+2)/6$) needed to sum to n ,” has a document whose author is N. Levine among its references. This document is indexed as Zbl 0083.04002.

2.2 Usage

The Links API represents a tool that can be used in various ways, both from machines and humans. Here, we present some instances where a user of either a given partner or zbMATH Open can benefit from the service:

- A DLMF/OEIS user can easily access all bibliographic resources indexed at zbMATH Open relating to a specific topic of interest (e.g., a special function in the DLMF, a sequence in the OEIS).
- A researcher interested in a publication indexed at zbMATH Open can use our API to verify if and possibly where that publication is cited in DLMF/OEIS. A search of this type can also be diversified thanks to the filters that our routes offer. For example, one might be interested in identifying which DLMF/OEIS links are related to a particular Mathematical Subject Classification (MSC) code or a particular author.
- A researcher more interested in the history of mathematics can use the Links API to trace the bibliography related to a certain topic covered in DLMF/OEIS and observe the historical development of the topic itself in terms of the literature related to it.

Last but not least, thanks to the Links API, we are offering a new service directly visible at our public web site. Precisely, we are displaying all URLs linking a given document to the external web page of the corresponding partner (see Figure 3).

At present, we are displaying 6,312 links from zbMATH Open to the DLMF and 67,436 links from zbMATH Open to the OEIS. From these numbers we can therefore conclude that, as a byproduct, we are also increasing the visibility of our partners.

3 OEIS references in zbMATH Open

3.1 Some OEIS figures

The OEIS is an online database of sequences of numbers and most people use it to get information about a particular number sequence. It was launched online in November 2010, based on the book “A Handbook of Integer Sequences” by N. J. A. Sloane published in 1973.¹⁵ This digital platform is very well maintained,¹⁶ daily updated and currently contains over 358,000 sequences. The entry for each sequence gives among other metadata, the beginning of the sequence, its name or description, formulas, references to books and articles where the sequence has appeared.

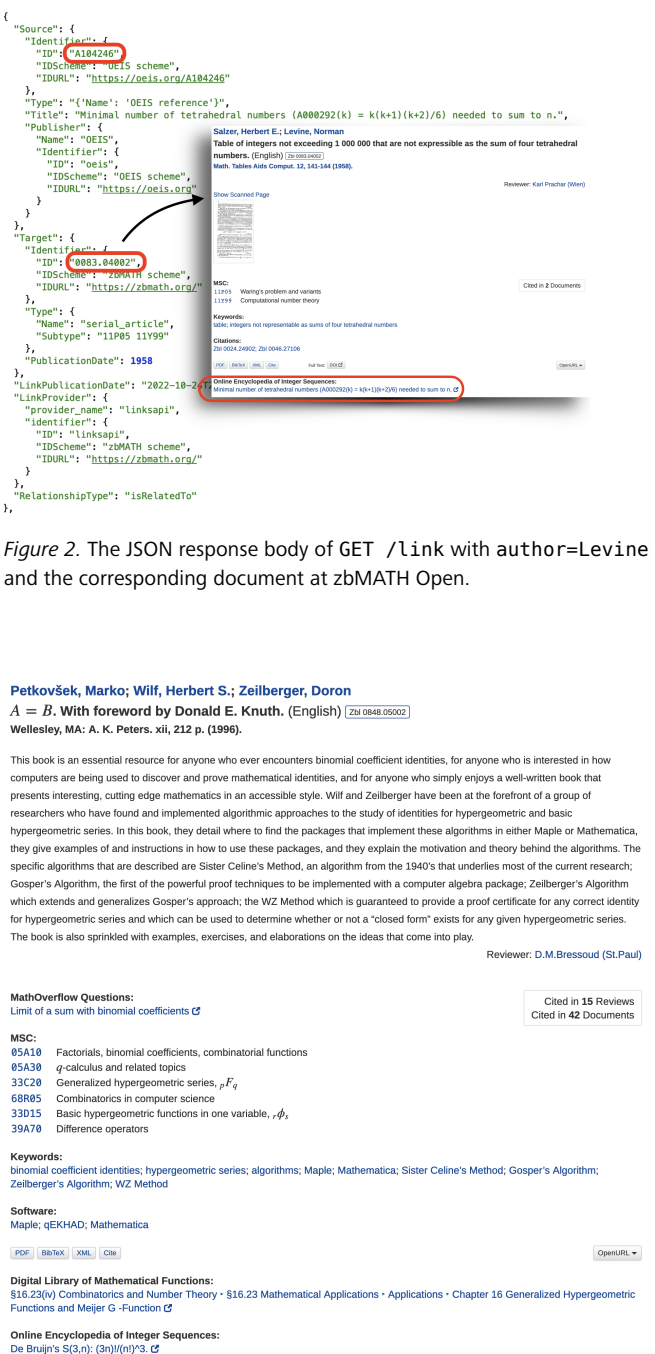


Figure 2. The JSON response body of GET /link with author=Levine and the corresponding document at zbMATH Open.



Figure 3. A book reviewed in zbMATH Open and linked to both the DLMF and the OEIS (as well as to a MathOverflow thread and mathematical software packages).

¹⁵ <https://zbmath.org/0286.10001>
¹⁶ <https://oeis.org/wiki/Welcome>

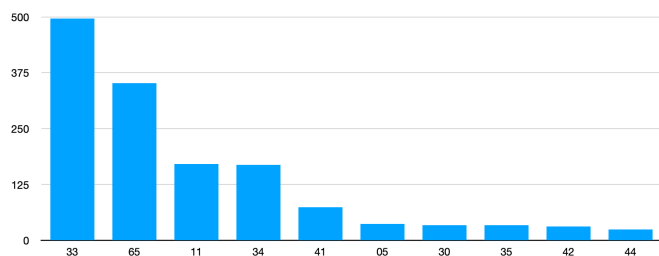


Figure 4. Distribution of primary 2-digit MSC codes (top ten) in the DLMF bibliography.

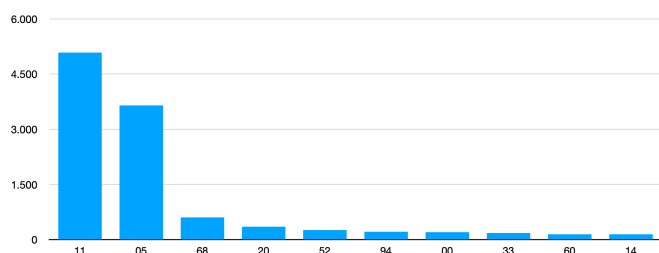


Figure 5. Distribution of primary 2-digit MSC codes (top ten) in the OEIS bibliography.

This last piece of data is what allowed us to match all bibliographic references in the OEIS with those in our database, thus constructing a dataset of all documents indexed at zbMATH Open cited by the OEIS. It turned out that there exist almost 60,000 references in OEIS matched to about 14,000 documents indexed in zbMATH Open (excluding the original printed version of Sloane's handbook which is, naturally, cited additionally for the majority of sequences). As a comparison, only about 2,000 documents (out of about 2,800) cited by the DLMF are indexed at zbMATH Open. As the figures indicate, the distribution is quite skewed: Several (38, as of January 1, 2023) publications are referenced in OEIS more than 100 times, while almost half (6567) documents have been cited just once in OEIS. A first analysis of their chronological and thematic distribution is given in the next section.

It is worth mentioning that the OEIS is very popular, and not only within the mathematical community, thanks to their commitment to disseminate the fascination and the ubiquity of number sequences. They are visible also on Facebook¹⁷ and YouTube.¹⁸

3.2 Analysis of available data

Based on our current DLMF/OEIS dataset, it is possible to draw a simple statistical analysis about the documents referenced by our partners. In fact, the two statistics routes of the Links API show the

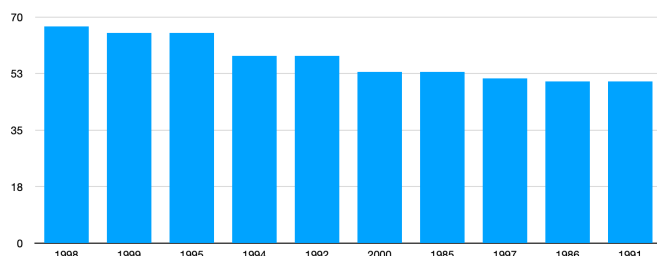


Figure 6. Distribution of years of publication (top ten) in the DLMF bibliography.

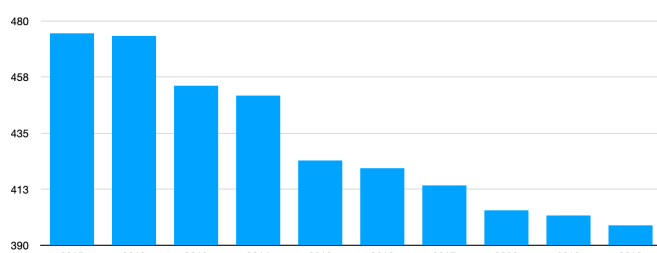


Figure 7. Distribution of years of publication (top ten) in the OEIS bibliography.

distribution of both primary MSC codes (2-digit level) and years of publication of the references for a given partner.

As one may expect, the most frequent cited primary MSC codes for DLMF and OEIS are 33 (Special functions) and 11 (Number theory), respectively. Even though this result is not at all surprising, this confirms the consistency over time of the MSC assignment to documents. In Figures 4 and 5 we can see the distribution of the 10 most frequently cited primary MSC codes from DLMF and OEIS, respectively.

A similar analysis provides information about the most frequent years of publication of cited references. This analysis is summarized in Figures 6 and 7. From the comparison of data between DLMF and OEIS it may be argued that OEIS seems to pay more attention than DLMF to more recent publications (2010–2020).

4 Concluding remarks

The interlinking of research data resources like DLMF and OEIS provides an additional information layer which enables navigation and analysis beyond the classical publications. Both platforms were a somewhat natural starting point for interlinking, due to their relevance, widespread use, and ample literature references. But the ecosystem of mathematical research data is much more diverse and granular, as discussed earlier in this column [2]. The interlinking of further resources will continue, but may not always as straightforward. For example, the L -functions and modular forms database

¹⁷ <https://www.facebook.com/OEISF>

¹⁸ <https://www.youtube.com/watch?v=LCWglXljevY>

(LMFDB),¹⁹ is a highly relevant and widely used database in number theory and algebraic geometry, but does not come along naturally with references to the literature. Vice versa, identifying its objects (e.g., modular forms, varieties, or representations) in the literature is something that is currently beyond the reach of automatization, so the only hope is that LMFDB entries are properly referenced when used (which is increasingly, but not systematically, the case). Further mathematical research data resources are currently standardized and integrated within the MaRDI initiative,²⁰ supported within the German National Research Data Infrastructure (NFDI) framework.

On the other hand, interlinking with zbMATH Open is clearly only a first step: In an ideal world one would, e.g., like to have an integrated formula search that would include full texts as well as functions or sequences. This is currently far from becoming reality due to a variety of issues, such as copyright restrictions for full texts, heterogeneous formats, or the lack of semantic information. Technical progress will surely help to overcome some obstacles, but the mathematical community can also support this by a variety of activities, e.g., by further voluntary engagement in the platforms, or appropriately referencing research data in publications. By interlinking the platforms, we hope to increase the awareness of the mathematical community of these services, and hopefully also motivate a further engagement.

References

- [1] H. S. Cohl, O. Tschke 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)
- [2] K. Hulek, F. Müller, M. Schubotz and O. Tschke, Mathematical research data – an analysis through zbMATH references. *Eur. Math. Soc. Newsl.* **113**, 54–57 (2019)
- [3] K. Hulek and O. Tschke, The transition of zbMATH towards an open information platform for mathematics. *Eur. Math. Soc. Newsl.* **116**, 44–47 (2020)
- [4] K. Hulek and O. Tschke, The transition of zbMATH towards an open information platform for mathematics. II: A two-year progress report. *Eur. Math. Soc. Mag.* **125**, 44–47 (2022)
- [5] M. Petrera, D. Trautwein, I. Beckenbach, D. Ehsani, F. Müller, O. Tschke, B. Gipp and M. Schubotz, zbMATH Open: API solutions and research challenges. *CEUR Workshop Proceedings* **2976**, 4–13 (2021)
- [6] M. Schubotz and O. Tschke, zbMATH Open: Towards standardized machine interfaces to expose bibliographic metadata. *Eur. Math. Soc. Mag.* **119**, 50–53 (2021)

Dariusz Ehsani is a researcher at zbMATH Open.

dariusz.ehsani@fiz-karlsruhe.de

Matteo Petrera is editor and researcher at zbMATH Open.

matteo.petrera@fiz-karlsruhe.de

Olaf Tschke is managing editor of zbMATH Open and vice-chair of the EMS Committee on publications and electronic dissemination.

olaf.tschke@fiz-karlsruhe.de

¹⁹ <https://www.lmfdb.org>

²⁰ <https://www.mardi4nfdi.de>