

Die Transformation von zbMATH zu einer offenen Plattform für die Mathematik. II: Ein Zwischenbericht nach zwei Jahren

Klaus Hulek und Olaf Teschke

Leibniz Universität Hannover und FIZ Karlsruhe

Vor zwei Jahren haben wir in [3] die Vision von zbMATH Open als offenem Dienst für die mathematische Forschung skizziert. Wir berichten über die seither erreichten Ergebnisse. Eine englische Version dieses Artikels ist im EMS Magazine zur Veröffentlichung angenommen.

1 Einleitung

Vor zwei Jahren haben wir in [3] beschrieben, wie wir zbMATH zu einen offenen Dienst für die mathematische Forschung umgestalten wollen. Vieles davon ist inzwischen Wirklichkeit geworden. Ermöglicht wurde dies durch die Bewilligung eines sogenannten Sondertatbestands der Leibniz-Gemeinschaft, der zunächst die Transformation während eines Zeitraums von zwei Jahren förderte. In unserem ursprünglichen Antrag hatten wir mehrere Ziele skizziert, um wesentliche Bedürfnisse der mathematischen Community zu unterstützen. Diese haben wir auch in [3] beschrieben und diskutiert. Natürlich ist der Zeitraum von zwei Jahren nicht lang genug, um all diese langfristigen Ziele vollständig zu erreichen. Darüber hinaus brachte die Pandemie zusätzliche besondere Herausforderungen mit sich, die parallel gelöst werden mussten. Trotzdem konnten wir einige wichtige Meilensteine erreichen, und die Evaluierung am Ende des Transformationszeitraums be-

stätigte, dass der Zuschuss verstetigt werden konnte. Damit erhält zbMATH Open nun eine nachhaltige Finanzierung durch die Leibniz-Gemeinschaft. In diesem Artikel berichten wir über die Fortschritte, die seit Anfang 2020 gemacht wurden.

2 Vorbereitungen (I) - Rechtliche Aspekte

Ob als lange Reihe von gelben Zentralblatt-Bänden oder später als Subskriptionsdatenbank zbMATH – seit seiner Gründung 1931 durch den Springer-Verlag hatte unser Reviewingdienst auch eine kommerzielle Komponente. Entsprechend beruhten alle rechtlichen Grundlagen, vom Redaktionsvertrag bis zu den Indexierungsvereinbarungen mit den zahlreichen Verlagen, auf dem Modell eines Abonnementdienstes, der von einer definierten und kontrollierbaren Menge von Subskribenten (jeweils mit speziellen Lizenzverträgen und unterschiedlichen Laufzeiten) genutzt wurde. Der Übergang zu einer Open-Access- und darüber hinaus zu einer Open-Data-Plattform erforderte eine vollständige Erneuerung dieser Verträge und damit verbunden, komplizierte Verhandlungen. Insbesondere wurde ein neuer Herausgebervertrag zwischen der European Mathematical Society (EMS), FIZ Karlsruhe – Leibniz-Institut für Informationsinfrastruktur und der Heidelberger Akademie der Wissenschaften abgeschlossen. Der Übergang zu dem Open Access Modell hatte auch zur Folge, dass die Rolle eines kommerziellen Distributors obsolet wurde. Dies brachte die Notwendigkeit einer Trennung von einem Partner mit sich, der unser Referateorgan von Beginn an und besonders in schwierigen Zeiten sehr unterstützt hat. Bei dieser Gelegenheit möchten wir uns bei unseren Springer-Kollegen für ihr jahrzehntelanges Engagement bedanken, das mit der Ermöglichung eines koordinierten Übergangs von Abonnements zu Open Access bis Ende 2020 abgeschlossen wurde. Die Rolle der European Mathematical Society und der Heidelberger Akademie besteht darin, die wissenschaftliche Qualität des Dienstes zu gewährleisten und die Beteiligung der mathematischen Gemeinschaft zu fördern.

Im Rückblick ist es erstaunlich, wie viele Verhandlungen und Vereinbarungen schon im Jahr 2020 abgeschlossen werden konnten. Dazu gehört eine große Anzahl erneuerter Indexierungsverträge mit einem Großteil der Mathematikverlage, von denen ein beträchtlicher Teil auch Open-Data-Vereinbarungen innerhalb der neuen Plattform zbMATH Open zugestimmt hat. Während beschlossen wurde, dass alle Daten, die im Rahmen des zbMATH-Redaktionsprozesses erzeugt werden (z.B. Reviews, Daten zur Autorendisambiguierung oder semantische und Verlinkungsdaten), unter der CC-BY-SA-Lizenz zur Verfügung gestellt werden sollen, passen nicht alle Verlagsdaten in diesen Rahmen, wie etwa Abstracts, die normalerweise anderen Copyright-Regelungen unterliegen. Dennoch konnte die Freigabe für einen beachtlichen Teil der Informationen erreicht werden, und dieser Anteil kann hoffentlich in weiteren Verhandlungen ausgebaut werden.

Auch die Nutzungsbedingungen – sowohl für das Webangebot als auch für Reviewer – mussten entsprechend angepasst und vereinbart werden. Schließlich wurde auch die Oberfläche überarbeitet, wobei zunächst ein besonderer Schwerpunkt auf der Minimierung der Speicherung und Verarbeitung von personenbezogener Daten lag. Während in der Vergangenheit Abonnements für die Bibliotheken mit der Anforderung von detaillierten Nutzungsberichten und entsprechenden Daten einhergingen, ermöglicht uns Open Access eine Plattform, die auf dem Prinzip der Datenvermeidung und Datenminimierung aufbaut. Tatsächlich gehört zbMATH Open so seit 2021 zu den wenigen komplexen Portalen, die komplett ohne Cookies (und entsprechend ohne die lästige Notwendigkeit von Cookie-Genehmigungen) genutzt werden können (es kommen nur optionale Cookies zum Einsatz, wenn Nutzungspräferenzen gespeichert werden sollen).

Die genannten Vorarbeiten konnten im Laufe des Jahres 2020 abgeschlossen werden, so dass zbMATH Open Anfang 2021 an den Start gehen konnte.

3 Feedback und Nutzung nach der Öffnung von zbMATH

2021 war das erste volle Kalenderjahr in dem zbMATH Open als kostenfreier Dienst zur Verfügung stand. Die Auswertung sowohl der Nutzung als auch des Feedbacks liefert ausgesprochen positive Ergebnisse. Dabei sind die Nutzungszahlen zweifellos auch durch das gewandelte Arbeitsumfeld in der Pandemiesituation beeinflusst. Als Subskriptionsdienst hatte zbMATH bis 2020 ca. 1.200 institutionelle Abonnements mit etwa 22 Millionen Suchanfragen pro Jahr, wobei der Zugriff hier häufig über Proxies der Bibliothek erfolgte und dies wenig Rückschlüsse auf individuelle Nutzung erlaubte. Im durchschnittlichen Monat des Jahres 2021 nutzten mehr als 60.000 Besucher die Website, mit mehr als 32 Millionen Suchanfragen im Jahr 2021. Dies deutet darauf hin, dass Open Access eine viel breitere Nutzerbasis ermöglichte, die immer noch wächst (die Zahlen im ersten Halbjahr 2022 deuten auf knapp 40 Millionen Abfragen in diesem Jahr hin).

Die Mitte 2021 durchgeführte Umfrage bekräftigte diesen Eindruck. Viele spezifische Fragen zu den Funktionen, dem Service und der Datenqualität von zbMATH Open bestätigten eine deutlich positive Entwicklung im Vergleich zu den bereits sehr guten Ergebnissen von 2016 (siehe den Bericht [1]). Der Schwerpunkt der Rückmeldungen lag aber deutlich auf der einhelligen Wertschätzung der Öffnung von zbMATH. Darüber hinaus gab es eine Reihe von Anregungen für weitere Verbesserungen, von denen einige bereits umgesetzt werden konnten, wie z. B. eine Verbesserung der Bibtex-Ausgabe in Richtung eines standardisierten Formats und die Umsetzung von Richtlinien zur Barrierefreiheit. Wir erlebten auch eine neue überwältigende Bereitschaft zum Beitrag von Reviews. Dies bedeutete, dass wir fast 1.200 neue Reviewer gewinnen konnten (etwa doppelt so viel wie in den Vor-

jahren), und insgesamt wurden 2021 13% mehr Rezensionen veröffentlicht als 2020.

Es ist vielleicht problematisch, aus der Menge ein einzelnes Beispiel herauszugreifen, aber die deutlich erhöhte Sichtbarkeit und Reichweite wird auch durch Peter Scholzes Review des Publ. RIMS Bandes mit Shinichi Mochizukis Inter-Universal Teichmüller Theory¹ illustriert. Durch den offenen Zugang wurde es innerhalb weniger Tage auf einer Vielzahl von Plattformen (Reddit, Twitter, MathOverflow. . .) verbreitet, verlinkt, diskutiert und mehr als 10.000 Mal gelesen, was es vermutlich zum meistgelesenen zbMATH-Review aller Zeiten macht.

4 Vorbereitungen (II) - neues technisches Framework

Der Übergang zu offenen Diensten erforderte eine Reihe von begleitenden technischen Entwicklungen. Dabei konzentrierte sich die Hauptanstrengung auf neue Funktionen zur Unterstützung der Verknüpfung der nunmehr offenen Daten mit anderen freien Quellen. Wichtig war an dieser Stelle der vollständige Austausch der Indexierungssoftware. Seit dem Release des Onlinedienstes Mitte der 90er Jahre basierte zbMATH auf einem hauseigenen Code, der für unsere spezifischen Daten optimiert war. Die zunehmende Vernetzung mit naturgemäß sehr heterogenen Daten aus freien Quellen führte ihn jedoch an seine Grenzen. Daher wurde der gesamte Indexierungscode auf Basis des wesentlich flexibleren Systems Elasticsearch aktualisiert. Die gewohnten Features von zbMATH wurden dabei erhalten, zusätzliche konnten eingerichtet werden.

Eine weitere Schlüsselkomponente war die Entwicklung eines neuen Reviewer-Backends, das 2020 auch im Hinblick auf die Erfahrungen der veränderten Workflows aus dem Homeoffice optimiert wurde. Ohne dieses System wäre es unmöglich gewesen, das Wachstum von Reviews und Reviewern im Jahr 2021 zu bewältigen. Gleichzeitig wurde der Code aller Backend- und Frontend-Komponenten von Python 2 auf Python 3 aktualisiert (ein für die Pflegbarkeit des Dienstes zentrales Projekt, auch wenn es nach außen hin keine Auswirkungen hat).

5 Neue Funktionen

Die erste neu verfügbare Open-Data-Funktion wurde sogar noch vor der Öffnung der Weboberfläche freigegeben: Im September 2020 wurde die erste Version der zbMATH Open OAI-PMH API veröffentlicht, die auf den Daten des Jahrbuchs über die Fortschritte der Mathematik (JFM) basiert (s. [10] für mehr Details zur

¹<https://zbmath.org/1465.14002>

zbMATH Open Documents Authors Serials Classification Software Formulae

Structured Search

an0920.33001 Q Fields Operators Help

But a book of this size covers thousands of formulas, and by Murphy's law, a few of them should be incorrect. I tried to find such misprints, in particular in the sections 3.11 and 3.12 about summation methods, since there I could use my Maple software for purposes of detection [W. Koepf, Hypergeometric summation. An algorithm approach to summation and special function identities (1998; Zbl 0909.33001)]. Not surprisingly, this search was successful: Formula (3.11.10) is incorrect by a factor $-n$, both identities in the middle of p. 175 are incorrect restatements of the corresponding contiguous relations (3.11.12) and (3.11.15) on p. 173; furthermore in formula (3.12.1) the upper parameter $z + n - 1$ should read $z + n + 1$ (I would like to thank George Andrews for sending me the corrected formula).

In spite of these minor shortcomings, I recommend this book warmly as a rich source of information to everybody who is interested in "Special Functions".

Reviewer: Wolfram Koepf (Leipzig)

MSC:
33-02 Research exposition (monographs, survey articles) pertaining to special functions
33B15 Gamma, beta and polygamma functions
33Cxx Hypergeometric functions
33Dxx Basic hypergeometric functions

Keywords:
hypergeometric functions; basic hypergeometric functions; orthogonal polynomials; Bessel functions

Citations:
Zbl 0573.30014; Zbl 0743.30021; Zbl 0909.33001

Software:
Maple

PDF BIBTeX HTML CDB OpenURL

Digital Library of Mathematical Functions:
§10.22(vi) Compendia • §10.22 Integrals • Bessel and Hankel Functions • Chapter 10 Bessel Functions
§1.10(vii) Inverse Functions • §1.10 Functions of a Complex Variable • Areas • Chapter 1 Algebraic and Analytic Methods
§1.15(vii) Summability of Fourier Series • §1.15 Summability Methods • Areas • Chapter 1 Algebraic and Analytic Methods
§1.15(viii) Fractional Derivatives • §1.15 Summability Methods • Areas • Chapter 1 Algebraic and Analytic Methods
§1.15(vi) Fractional Integrals • §1.15 Summability Methods • Areas • Chapter 1 Algebraic and Analytic Methods
§13.9(v) Orthogonal Polynomials • §13.6 Relations to Other Functions • Kummer Functions • Chapter 13 Confluent Hypergeometric Functions
§13.9(i) Zeros of $M(a, b, z)$ • §13.9 Zeros • Kummer Functions • Chapter 13 Confluent Hypergeometric Functions
Chapter 13 Confluent Hypergeometric Functions

Abbildung 1: DLMF-Referenzen aus dem zbMATH Open-Review von Andrews/Askey/Roy's „Special functions“

Geschichte des JFM und den verfügbaren Daten in zbMATH Open). Möglich war dies dank der European Mathematical Society und der SUB Göttingen, die diese Daten bereits 2020 unter einer CC-BY-SA 4.0-Lizenz zur Verfügung stellten. Die Erfahrungen aus dieser ersten Version ermöglichten 2021 die Erweiterung auf die zbMATH Open data in einer stabilen Version [9, 7], die nun eine Kernkomponente der angestrebten offenen Plattform bildet. Basierend auf dieser Schnittstelle werden weitere APIs mit dem Ziel entwickelt, erweiterte Standards zu unterstützen und spezifische Dienste miteinander zu verknüpfen.

Eine unmittelbare Anwendung war die Entwicklung der Links API, die in ihrer ersten Anwendung zbMATH Open mit der NIST Digital Library of Mathematical Function (DLMF) verknüpft. Das System, das in Zukunft für die Verlinkung mit weiteren Forschungsdatenressourcen angepasst werden kann, wurde in [2] detailliert beschrieben, und die daraus resultierenden Verknüpfungsdaten sind nun auch in zbMATH Open sichtbar.

Aber auch die klassischen Aufgaben der Bereitstellung von Literatur- und Autoreninformationen profitieren stark vom offenen Ansatz. Die verbesserte und erweiterte zbMATcH-API ermöglichte die Verknüpfung mit deutlich mehr offenen Volltextquellen. Nunmehr sind nicht nur doppelt so viele arXiv-Links wie zuvor integriert, sondern es gibt auch fast 100.000 neue Links zu freien Unpaywall- und CiteSeer-Ressourcen. Autorenprofile enthalten Verknüpfungen mit bis zu fünfzehn externen Diensten. Umgekehrt können zusätzliche Informationen aus diesen Quellen aufgenommen werden, die z.B. die Anzeige verschiedener Nicht-ASCII-

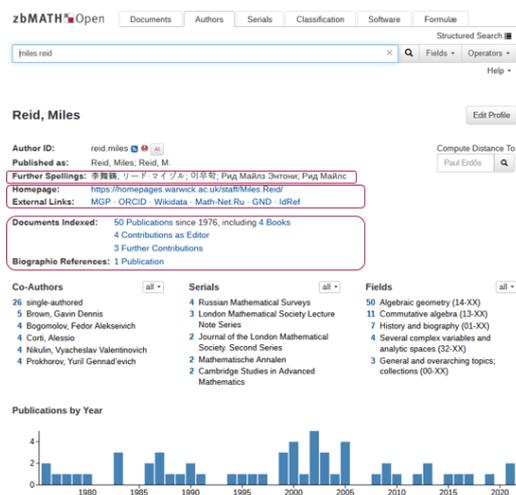


Abbildung 2: 2022 zbMATH Autorenprofil

Schreibweisen (wie Arabisch, Chinesisch, Farsi, Koreanisch, Kyrillisch...) ermöglichen. Darüber hinaus unterscheiden die Profile nun zwischen verschiedenen Rollen in der Publikation (Autor/-Herausgeber/weitere Beiträge, z.B. Appendix). Ebenso werden zusätzliche Informationen zum Koautornetzwerk angezeigt (die beiden letzteren Funktionen wurden explizit in der Umfrage gewünscht).

Interessant ist auch, dass die Schnittstellen der mathematischen Community auf verschiedene Weise ermöglichen, die für die Profile essentiellen Disambiguierungsdaten weiter zu verbessern. Neben dem schon länger existierenden direkten Interface in den Profilen [4] spielen zunehmend externe Community-Plattformen eine wichtige Rolle. Dies reicht von der Korrektur mehrerer Fehler bei der Zuschreibung von Werken von René Peiffer (bekannt durch die Peiffer-Identität) auf Twitter² bis hin zur Diskussion auf MathOverflow, dass der Trick von Rabinowitsch höchstwahrscheinlich nicht, wie zuvor angenommen, von G. Y. Rainich entdeckt wurde³.

Ein weiteres häufig nachgefragtes Feature wird aktuell entwickelt - eine Datenbank über die Zugehörigkeit zu Institutionen. Derzeit gibt es bereits eine interne Datenbank mit etwa 15.000 disambiguierten Einträgen mathematischer Institutionen, die jetzt mit Publikationen und Personen verknüpft werden. Danach werden ebenfalls offene Affiliation-Daten in zbMATH Open zur Verfügung stehen.

²<https://twitter.com/zbMATH/status/1368319234636447747>

³<https://mathoverflow.net/questions/416577/identity-of-j-1-rabinowitsch-of-rabinowitsch-trick>

6 Ausblick

Die laufende interne Entwicklung zur Integration und Verknüpfung weiterer Publikationen, Personen- und Affiliationdaten, Forschungsdaten und Community-Plattformen ist nur ein Aspekt des wachsenden Netzwerks. Genauso wichtig ist die Nutzung von zbMATH Open data in weltweit durchgeführten externen Projekten. Aktuell nutzen besonders Projekte Bereich der semantischen Analyse und der Geschichte der Mathematik intensiv Daten aus zbMATH Open. So enthält z.B. der Band „Framing global mathematics“ von Norbert Schappacher zum hundertjährigen Bestehen der IMU [8] eine detaillierte Analyse der ICM-Vortragenden, ihrer Netzwerke und Arbeitsgebiete auf der Basis von zbMATH Open-Daten. Bereits seit mehreren Jahren wird das vom IMU geleitete ISC-Projekt „Gender Gap In Science“⁴ durch zbMATH Open-Daten unterstützt, die ein integraler Bestandteil der Visualisierungsplattform „Gender Publication Gap“⁵ sind (s. a. [5]).

Und nicht zuletzt konnte sich FIZ Karlsruhe – die Institution, die zbMATH Open produziert, entwickelt und pflegt – auf Basis dieser Daten an mehreren Open Science-Projekten beteiligen. Das vielleicht wichtigste ist die Mathematical Research Data Initiative (MarDI)⁶, das als Mathematik-Konsortium im Rahmen der sich entwickelnden deutschen Nationalen Forschungsdateninfrastruktur (NF-DI) bewilligt wurde und Ende 2021 die Arbeit aufgenommen hat. Zwei weitere Projekte - die von der DFG unterstützte Entwicklung eines mathematikspezifischen Plagiatserkennungssystems (auf Grundlage der Untersuchungen in [6]) und die Aufnahme von Softwareinformationen in the EOSC-Cloud im Rahmen des FAIRCORE4EOSC-Projekts - haben am 01.04.2022 bzw. am 01.06.2022 angefangen.

Vor allem aber ist es die Mathematik-Community, die mit ihren Ideen und Beiträgen die Weiterentwicklung von zbMATH Open vorantreibt. Wir freuen uns sehr über die ständig wachsende Zahl wertvoller Rezensionen sowie über alle Vorschläge und Rückmeldungen an editor@zbmath.org!

Literatur

- [1] I. Brüggemann, K. Hulek, and O. Teschke, “Results of the 2016 EMS user survey for zbMATH”, *Eur. Math. Soc. Newsl.* 104, 67–68, 2017
- [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. Newsl.* 120, 66–67, 2021

⁴<https://gender-gap-in-science.org/>

⁵<https://gender-publication-gap.f4.htw-berlin.de/>

⁶<https://mardi4nfdi.de>

- [3] K. Hulek and O. Teschke, „Die Transformation von zbMATH zu einer offenen Plattform für die Mathematik.“ Mitt. Dtsch. Math.-Ver. 28, No. 2, 108–111, 2020
- [4] H. Mihaljević-Brandt and N. Roy, “zbMATH author profiles: open up for user participation”, Eur. Math. Soc. Newsl. 93, 53–5, 2014
- [5] H. Mihaljević and L. Santamaría, “Mathematics publications and authors’ gender: Learning from the Gender Gap in Science project”, Eur. Math. Soc. Mag. 123, 34–38, 2022
- [6] M. Schubotz, O. Teschke, V. Stange, N. Meuschke, and B. Gipp, “Forms of plagiarism in digital mathematical libraries”, Lect. Notes Comput. Sci. 11617, 258–274, 2019
- [7] M. Petrera, D. Trautwein, I. Beckenbach, D. Ehsani, F. Müller, O. Teschke, B. Gipp, and M. Schubotz, “zbMATH Open: API Solutions and Research Challenges”, Proc. DISCO 2021, JCDL 2021, 4–13, 2021
- [8] N. Schappacher, “Framing Global Mathematics. The International Mathematical Union between Theorems and Politics” Springer, 2022
- [9] M. Schubotz and O. Teschke, “zbMATH Open: Towards standardized machine interfaces to expose bibliographic metadata”, Eur. Math. Soc. Newsl. 119, 50–53, 2021
- [10] O. Teschke, “The ‘Jahrbuch über die Fortschritte der Mathematik’ as a part of zbMATH Open”, Eur. Math. Soc. Mag. 122, 62–64, 2021

Adresse der Autoren:

*Leibniz Universität Hannover, Fakultät für Mathematik und Physik, Institut für Algebraische Geometrie, Welfengarten 1, D-30167 Hannover, DEUTSCHLAND
email hulek@math.uni-hannover.de*

*FIZ Karlsruhe, Leibniz-Institut für Informationsinfrastruktur Abteilung Mathematik und Informatik, Franklinstr. 11, D-10587 Berlin, DEUTSCHLAND
email olaf.teschke@fiz-karlsruhe.de*