



**FIZ Karlsruhe**

Leibniz Institute for Information Infrastructure

# Software information and the swMATH service

MMS Days  
2018-03-01  
Leipzig

Wolfram Sperber

# AGENDA

INTRODUCTION

STATE OF THE ART

THE SWMATH APPROACH

ENCODING OF SOFTWARE CITATIONS

FINAL REMARKS

# A few general remarks about scientific software (I)

- Scientific software is research data but software has unique features → (scientific) software is a tool implementing an algorithm in a special environment
- Reproducing of research results achieved by software requires not only the software but also the data
- Reproducing of research results achieved by software requires knowledge about the used version of the software
- Reproducing and reuse of software presupposes a special hard- and software environment
- Reuse of software can be done in different way by APIs, webservice (as cloud computing) or by development of code

# A few general remarks about scientific software (I)

- Scientific software is research data but software has unique features → (scientific) software is a tool implementing an algorithm in a special environment
- Reproducing of research results achieved by software requires not only the software but also the data
- Reproducing of research results achieved by software requires knowledge about the used version of the software
- Reproducing and reuse of software presupposes a special hard- and software environment
- Reuse of software can be done in different way by APIs, webservice (as cloud computing) or by development of code

# A few general remarks about scientific software (II)

- Software is more than code  
(the term software stands for an aggregated object consisting of software code, and accompanied documents for the human user as documentations, tutorials, implementation guides, license about licenses and usability, dependencies, developers, etc. )
  - Software information is distributed on software portals, software repositories, software websites, or in publications
  - Scientific software development is often distributed.  
For that developer platforms have been established, e.g., Git (Github), Maven, etc.
- requirements to software information and knowledge management must be redefined and specified

# Needs to Software Information and Knowledge Management

## General needs to information for research products:

- accessibility
- citeability
- persistency of research objects and persistent identifiers
- a few useful metadata
- quality

## Specification for software:

- structured information about software products and versions
- information about licenses and usability of software
- information about technical environment of software
- information how software can be reused: (Semantic) APIs, webservices, Open Source code
- scientific background: algorithms, models
- acceptance and use cases

# State of the art

- The current information services for scientific software are fragmented (see, e.g., software repositories, or Wikipedia lists for special fields of software) and use proprietary standards, e.g., for citation.

But

- there is a broad discussion about standards reflecting the increasing role of scientific software (FORCE11, RDA, Software Sustainability Institute)  
A concrete result: `Software Citation Principles`  
(Implementation of the Software Citation Principles is under work)
- Next question: How can we build up an useful and efficient information and knowledge management system for scientific software?

# A distributed system for scientific software

- The initial situation: Various resources (e.g., git sources, software websites, software repostories, publications, etc. ) provide different information about it.
- A cooperative distributed system would be the best solution.
- Therefore we need standards and standardization and new services, e.g., for long-term archiving of software.

# The swMATH approach

- swMATH is a node in such a distributed system and a (general) portal for (mathematical) software.
- A short overview about the swMATH concepts and features:

The first concept:

## The publication-based approach

This means:

- Identification of software citations in “publications” (by heuristic means), sources:
  - zbMATH,
  - arXiv,
  - software journals

## In which journals should I publish my software?

By Neil Chue Hong.

Until there is a radical change in the way that academic credit is given, the principal record of scientific research is still the peer-reviewed publication. Given that software is a fundamental part of doing science in the digital age, the question we are often asked is: *where can I publish papers which are primarily focused on my scientific software?*

The following is a list of journals which accept submissions that are primarily about the software, and not necessarily on new algorithms or new science. There is an expectation that the use of the software will enable new research to be carried out. This list of journals is not exhaustive - if you know of others, please [let us know](#).



### General Journals

- [Computing in Science & Engineering](#)
- [Concurrency and Computation: Practice and Experience](#)
- [Journal of Open Research Software](#)
- [The Journal of Open Source Software](#) *[example]*
- [Journal of Software: Practice and Experience](#)
- [Nature Toolbox](#)
- [Research Ideas and Outcomes \(RIO\)](#)
- [SIAM Journal on Scientific Computing \(SISC\) Software section](#)
- [SoftwareX](#)

### Engineering

- [Advances in Engineering Software](#)
- [Coastal Engineering](#)
- [Renewable Energy](#)

### Humanities and Social Sciences

- [Digital Humanities Quarterly](#)
- [Journal of Artificial Societies and Social Simulation](#)
- [Journal of Economic Dynamics and Control](#)

### Software and Research Blog

21-February-2018 - [CW17: My experience](#) - By Nikoleta Glynatsi, Software Sustainability Institute fellow The...

16-February-2018 - [Collaborations Workshop discussion session and collaborative ideas explained](#) - By Raniere Silva, Community Officer. Two activities that form part of all...

15-February-2018 - [Cloud-first: Simple automated testing using Drone](#) - By Mark Woodbridge, Research Software Engineering Team Lead This is the...

09-February-2018 - [Bridging the gap: Convincing researchers with different backgrounds to adopt good \(enough\) software development practices](#) - By Stuart Grieve, Research Software Developer, University College London,...

08-February-2018 - [How do we help adapt software best practices to make them more applicable to domain researchers?](#) - By Adam Tomkins (Chair), University of Sheffield, James Grant, University,...

### Latest News

19-February-2018 - [Call for posters & registration open for Workshop: Impact of international collaborations in research software](#) - The Software Sustainability Institute is organising the 'Workshop: Impact...

14-February-2018 - [UXLS: UX Toolkit for Life Sciences](#) - The culmination of two years' work by the Pistoia Alliance project team,...

## Informatics, Mathematics and Statistics

- [ACM Transactions on Mathematical Software](#)
- [The Archive of Numerical Software](#)
- [Future Generation Computer Systems](#)
- [Journal of Machine Learning Research](#) (Machine Learning Open Source Software track)
- [Journal of Multiscale Modelling and Simulation](#)
- [Journal of Parallel and Distributed Computing](#)
- [Journal of Software for Algebra and Geometry](#) [*example*]
- [Journal of Statistical Software](#)
- [Knowledge-Based Systems](#)
- [LMS Journal of Computation and Mathematics](#) [*example*]
- [The Mathematica Journal](#) [*example*]
- [Mathematical Programming Computation](#)
- [Numerical Algorithms](#)
- [PeerJ Computer Science](#) [*example*]
- [The R Journal](#)
- [Science of Computer Programming](#) [*example*]
- [The Stata Journal](#) [*example*]

## Life Sciences

- [American Journal of Human Genetics](#)
- [Artificial Intelligence in Medicine](#)
- [Artificial Life](#)
- [Behaviour Research Methods](#) [*example*]
- [Bioinformatics \(Application Notes\)](#)
- [Bioinformatics and Biology Insights](#)
- [Biophysical Journal](#)
- [BMC Bioinformatics](#)
- [BMC Neuroscience](#) [*example*]
- [BMC Systems Biology](#)
- [BMC Source Code for Biology and Medicine](#)
- [Bone](#)
- [Computer Methods and Programs in Biomedicine](#) [*example*]
- [Current Protocols in Bioinformatics](#)
- [Database: The Journal of Biological Databases and Curation](#)
- [Ecography \(Software Notes\)](#) [*example*]
- [eLife \(Tools and Resources\)](#) [*example*]
- [Epidemiology](#)
- [Evolutionary Bioinformatics](#)
- [F1000 Research](#)
- [Frontiers in Neuroinformatics](#)
- [Gigascience](#)
- [Methods in Ecology and Evolution](#)
- [Nature Methods](#) [*example*]
- [Neurocomputing](#)
- [Neuroinformatics](#)
- [Nucleic Acids Research \(special issues\)](#)
- [PeerJ](#) [*example*]
- [PLoS Computational Biology: Software collection](#)
- [PLoS ONE](#)
- [Trends in Parasitology](#)

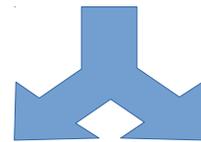
## Physical Sciences and Geosciences

- [AAS: The Astronomy Journal](#)
- [AAS: The Astrophysical Journal](#)
- [AAS: The Astrophysical Journal Supplement Series](#) [*example*]
- [Astronomy and Computing](#)
- [Communications in Computational Physics](#)
- [Computational Astrophysics and Cosmology](#)
- [Computer Physics Communications](#)
- [Computers and Geosciences](#)
- [Computing and Software for Big Science](#)
- [Geoscientific Model Development](#)
- [International Journal of Quantum Chemistry](#)
- [Journal of Chemical Theory and Computation](#)
- [Journal of Chemical Information and Modelling](#)
- [Journal of Cheminformatics](#)
- [Journal of Computational Chemistry \(special articles - software news and updates\)](#)
- [Molecular Simulation](#)
- [Wiley Interdisciplinary Reviews: Computational Molecular Science \(Software Focus\)](#) [*example*]

# The next step

- Extraction and analysis of software information in the publications

## Publications citing a software



Publications with software  
as subject  
*(describe the problems which  
can be solved with a software)*

Publications which cite  
software as tool  
*(provide information about  
acceptance and use cases of a  
software)*

# Problems of the publication-based approach: Identification

- Identification of software citations is done by heuristic means because software citations in publications are heterogeneous and often sparse:
  - Software citations refer to **different objects**, especially publications, documentation, websites, etc.
  - Often citations contain not more than the name of the software, no information about the version used.  
In other words, typically software citations refer to only to **software products or software container**.

# The software profiles in swMATH

By an analysis of software citation in publications we get:

- title
- description (abstract)
- (weighted list of) keywords  
*(weights result from the frequency of keywords)*
- classification of mathematical subjects and application areas  
*(corresponding to the Mathematical Subject Classification)*
- citation and usage profile  
*(frequency of citations of a software per year)*
- similar software  
*(by counting the common references of a software in citations, Google's page rank is under implementation)*
- list of publications citing a software

# Adding information to swMATH pages from other resources

Enrichment of information by non-publication resources, especially

- › *the website of a software,*
- › *repositories (CRAN),*
- › *and Internet Archive/Tempas*

in detail:

- › searching for a website of a software (by a Google search)
- › looking for versions of the software in the Internet Archive or repositories
- › analysis, structuring, and combining the additional information with the existing information

[Advanced search](#)
[Browse](#)



## SCIP

SCIP is currently one of the fastest non-commercial solvers for mixed integer programming (MIP) and mixed integer nonlinear programming (MINLP). It is also a framework for constraint integer programming and branch-cut-and-price. It allows for total control of the solution process and the access of detailed information down to the guts of the solver. SCIP is part of the SCIP Optimization Suite, which also contains the LP solver SoPlex, the modelling language ZIMPL, the parallelization framework UG and the generic column generation solver GCG.

This software is also **peer reviewed** by journal MPC.

### Keywords for this software



**URL:** [scip.zib.de/](http://scip.zib.de/)

**InternetArchive**

**Versions:** [Info](#)

**Authors:** Gerald Gamrath, Ambros Gleixner, Gregor Hendel, Stephen J. Maher, Matthias Miltenberger, Benjamin Müller, Marc Pfetsch, Felipe Serrano, Dieter Weninger, Jakob Witzig

**Platforms:** Linux, Windows, Mac OS

**Licence:** ZIB academic license

**Current version:** 3.2

**Dependencies:** LP-solver, e.g. SoPlex, CPLEX, XPress, ...

Add information on this software.

### Related software:

- CPLEX
- MIPLIB
- MIPLIB2003
- Gurobi
- SoPlex
- XPRESS
- LINDO
- Benchmarks for Optimization...
- Bonmin
- Ipopt

Show more...

Showing results 1 to 20 of 268.

Sorted by year (citations) 20

1 2 3 ... 12 13 14 next

1. Altherr, Lena C.; Dörig, Bastian; Ederer, Thorsten; Pelz, Peter F.; Pfetsch, Marc E.; Wolf, Jan: A mixed-integer nonlinear program for the design of gearboxes (2018) [archived SW](#)
2. Berthold, Timo: A computational study of primal heuristics inside an MI(NL)P solver (2018) [archived SW](#)
3. Chen, Wei-Kun; Chen, Liang; Yang, Mu-Ming; Dai, Yu-Hong: Generalized coefficient strengthening cuts for mixed integer programming (2018) [archived SW](#)
4. Andrea Callia D'Iddio, Michael Huth: Manyopt: An Extensible Tool for Mixed, Non-Linear Optimization Through SMT Solving (2017) [arXiv](#) [archived SW](#)
5. Assarf, Benjamin; Gawrilow, Evgenij; Herr, Katrin; Joswig, Michael; Lorenz, Benjamin; Paffenholz, Andreas; Rehn, Thomas: Computing convex hulls and counting integer points with polymake (2017) [archived SW](#)
6. Bartlett, Mark; Cussens, James: Integer linear programming for the Bayesian network structure learning problem (2017) [archived SW](#)
7. Beck, Amir; Pan, Dror: A branch and bound algorithm for nonconvex quadratic optimization with ball and linear constraints (2017) [archived SW](#)
8. Belotti, Pietro; Berthold, Timo: Three ideas for a feasibility pump for nonconvex MINLP (2017) [archived SW](#)
9. Berg, Jeremias; Järvisalo, Matti: Cost-optimal constrained correlation clustering via weighted partial maximum satisfiability (2017) [archived SW](#)
10. Berthold, Timo: Improving the performance of MIP and MINLP solvers by integrated heuristics (2017) [archived SW](#)
11. Brinkmann, Philip; Ziegler, Günter M.: A flag vector of a 3-sphere that is not the flag vector of a 4-polytope (2017) [archived SW](#)
12. Cheung, Kevin K.H.; Gleixner, Ambros; Steffy, Daniel E.: Verifying integer programming results (2017) [archived SW](#)
13. Cussens, James; Järvisalo, Matti; Korhonen, Janne H.; Bartlett, Mark: Bayesian network structure learning with integer programming: polytopes, facets and complexity (2017) [archived SW](#)
14. Dilkina, Bistra; Khalil, Elias B.; Nemhauser, George L.: Comments on: "On learning and branching: a survey" (2017) [archived SW](#)
15. Firsching, Moritz: Realizability and inscribability for simplicial polytopes via nonlinear optimization (2017) [archived SW](#)
16. Gerard, D.; Köppe, M.; Louveaux, Q.: Guided dive for the spatial branch-and-bound (2017) [archived SW](#)
17. Ghasemi, Mohammad S.; Afzaljan, Ali A.: Robust tube-based MPC of constrained piecewise affine systems with bounded additive disturbances (2017) [archived SW](#)
18. Gleixner, Ambros M.; Berthold, Timo; Müller, Benjamin; Weltge, Stefan: Three enhancements for optimization-based bound tightening (2017) [archived SW](#)
19. Göttlich, Simone; Potschka, Andreas; Ziegler, Ute: Partial outer convexification for traffic light optimization in road networks (2017) [archived SW](#)
20. Guns, Tias; Dries, Anton; Nijssen, Siegfried; Tack, Guido; De Raedt, Luc: MiningZinc: A declarative framework for constraint-based mining (2017) [archived SW](#)

1 2 3 ... 12 13 14 next

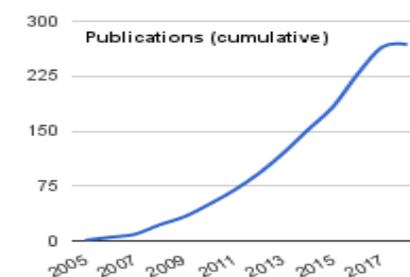
Search for articles

**MSC classification / top**

- Top MSC classes
  - 05 Combinatorics
  - 52 Convex and discrete...
  - 65 Numerical analysis
  - 68 Computer science
  - 90 Optimization
- Other MSC classes

**Publication year**

- 2010 - today
- 2005 - 2009
- 2000 - 2004
- before 2000

**Chart: cumulative / absolute**

# Links to versions

Software SCIP in  
*Altherr, Lena C.: D"orig, Bastian; Ederer, Thorsten; Pelz, Peter F.; Pfetsch, Marc E.; Wolf, Jan: A mixed-integer nonlinear program for the design o... (2018)*

swMATH

home source code publications updates documentation artifacts

SCIP Optimization Suite

SCIP SoPlex ZIMPL UG GCG Documentation

Temporal TimePortal

Q Search

live current web

2018/01/04 publication year

2017/01/06 year before publ.

Internet Archive powered by Wayback Machine

L3S ALEXANDRIA TIB

Get in touch!

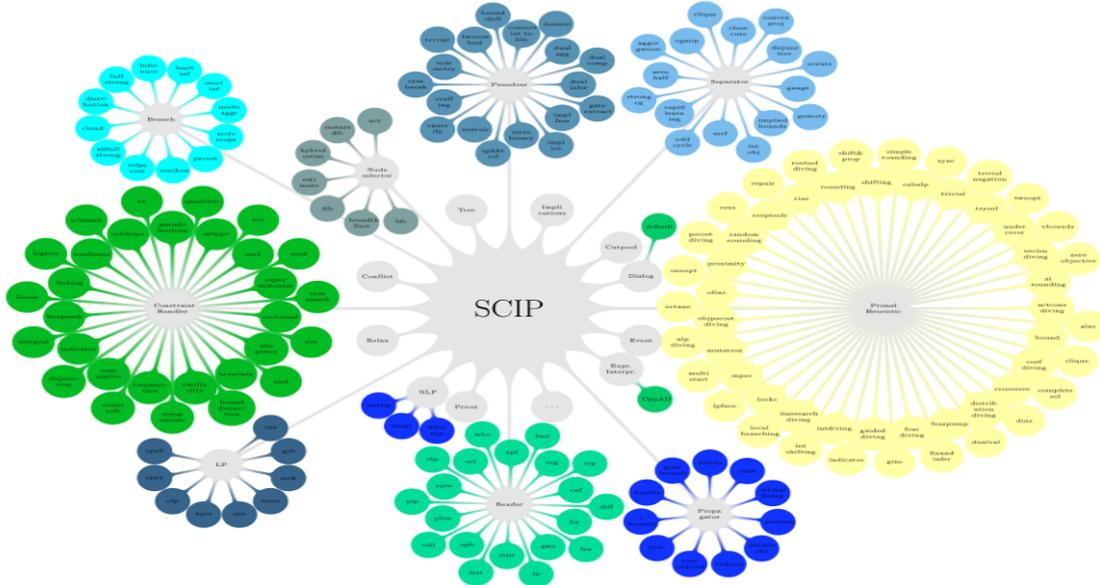
home source code publications updates documentation artifacts

SCIP Optimization Suite

SCIP SoPlex ZIMPL UG GCG Documentation

## About

SCIP is currently one of the fastest non-commercial solvers for mixed integer programming (MIP) and mixed integer nonlinear programming (MINLP). It is also a framework for constraint integer programming and branch-cut-and-price. It allows for total control of the solution process and the access of detailed information down to the guts of the solver.

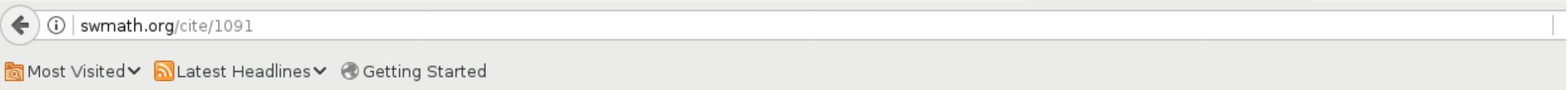


By default, SCIP comes with a bouquet of different plugins for solving MIPs and MINLPs.  
Click [here](#) to open this webpage version in the Internet Archive's Wayback Machine

# Software Citations Implementation: Standard TeX encoding

BibTeX standard  
BibLaTeX & Biber doesn't support the type software

You have to use the `misc' type



software citation proposal:

```
@MISC{sw01091,  
author = {Gerald Gamrath, Ambros Gleixner, Gregor Hendel, Stephen J. Maher, Matthias Miltenberger, Benjamin Müller, Marc Pfetsch, Felipe Serrano, Dieter Weninger, Jakob Witzig},  
title = {SCIP},  
note = {[SW] SOI:swmath.01091 (\url{http://www.swmath.org/software/01091})}  
}
```

# Software Citations Implementation: Customized TeX encoding

BibTeX extension:

Definition of new types

- software
- software versions

and the corresponding fields

The new types and the output must be defined in the files

- software.bbx
- software.cbx
- software.dbx

(local extensions)

# Software Citations Implementation: Customized TeX encoding

BibTeX extension:

Definition of new types

- software
- software versions

and the corresponding fields

The new types and the output must be defined in the files

- software.bbx
- software.cbx
- software.dbx

(local extensions)

```
software{MIPLIB,  
  id = {\url{http://swmath.org/software/4067}},  
  ti = {MIPLIB -- the Mixed Integer Programming LIBrary},  
  creator = {Robert E. Bixby and E.A. Boyd and Indovina, R.R and D. Dee},  
  maintainer = {--},  
  = {1992},  
}  
  
softwareversion{MIPLIB2010,  
  id = {\url{http://swmath.org/software/4067}},  
  ti = {MIPLIB -- the Mixed Integer Programming LIBrary},  
  creator = {--},  
  maintainer = {Thorsten Koch and Daniel Rehfeldt},  
  version = {5.0.0},  
  versiondate = {2010},  
}
```



- [1] Creator(s): **Robert E. Bixby, E.A. Boyd, R.R. Indovina, and D. Dee**  
*MIPLIB — the Mixed Integer Programming LIBrary* (1992)  
<http://swmath.org/software/4067>
- [2] Maintainer(s): **Thorsten Koch and Daniel Rehfeldt**  
*MIPLIB — the Mixed Integer Programming LIBrary*  
Version: 5.0.0 (2010)  
<http://swmath.org/software/4067>

# Final remarks

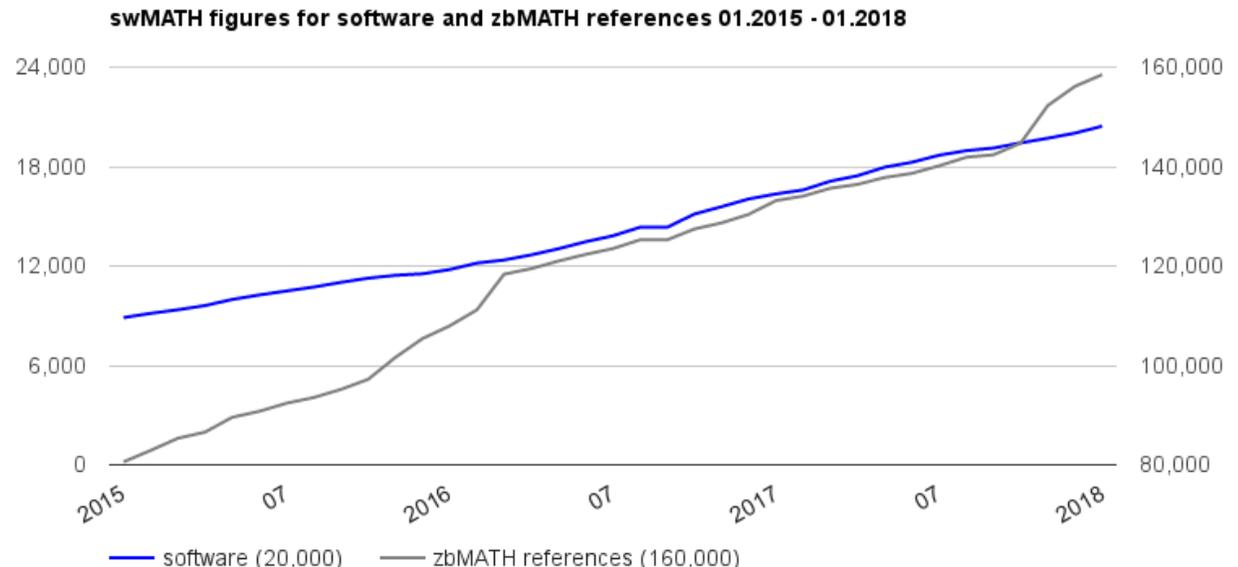
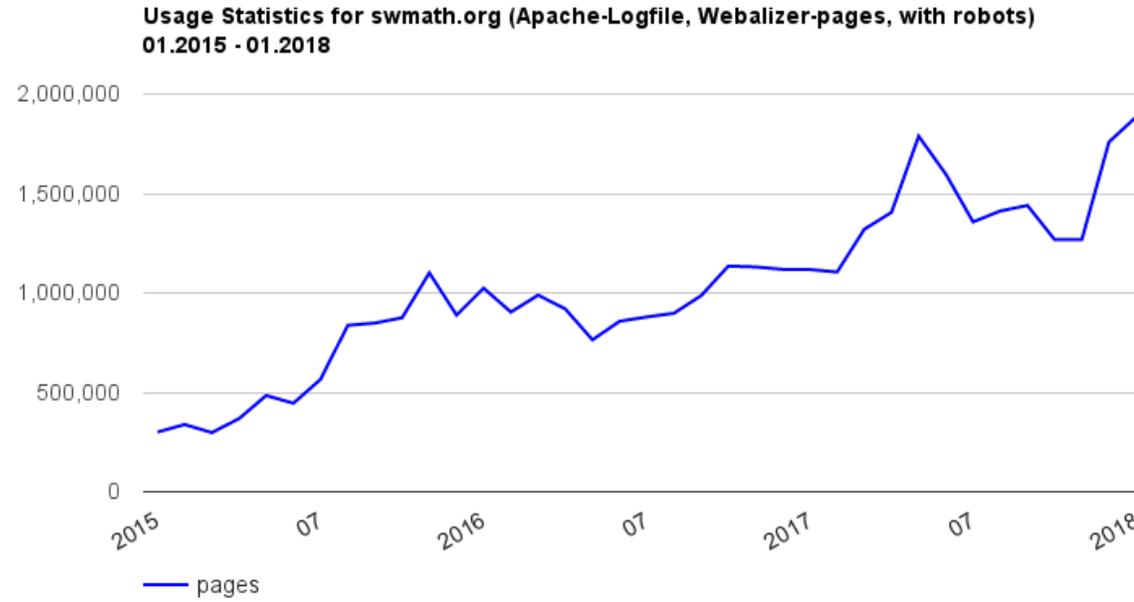
Software information comes from various sources, especially from the developers and users of the software.

A distributed information system which provides links and access to all relevant information about a software would be useful. The idea of a such a link page called `Micro Web' was created by Helge Holzmann.

The swMATH pages of a software try to comprehensively collect the information about a software.

The swMATH pages are rudimentary Micro Webs for software.

# Usage statistics (swMATH)



# A special feature in swMATH: The MMS network collection



Search [Advanced search](#)

Browse

- [browse software by name](#)
- [browse software by keywords](#)
- [browse software by MSC](#)
- [browse software by types](#)

### Browse software by types

- 1 [Benchmarks](#) (53)
- 2 [Book Companion Software](#) (51)
- 3 [Data Collections](#) (31)
- 4 [Languages](#) (204)
- 5 [Educational](#) (48)
- 6 [Portals](#) (16)
- 7 [Services, Webservices](#) (18)

**special collections:**

- 1 [Math.Modeling and Simulation - MMS](#) (24)
- 2 [Theorem Prover Museum](#) (28)



Results 1 to 20 of 24

**Jstacs** Referenced in 1 article [[sw14436](#)]

Jstacs: a Java framework for statistical analysis and...

**CRP Toolbox** Referenced in 1 article [[sw17976](#)]

CRP toolbox for Matlab. CROSS RECURRENCE PLOT TOOLBOX...

**pyunicorn** Referenced in 2 articles [[sw19314](#)]

Unified functional network and nonlinear time series analysis...

**SimEnv** Referenced in 0 articles [[sw19587](#)]

The complexity of the Earth system, the intrinsic...

**TOCSY** Referenced in 1 article [[sw19588](#)]

TOCSY - Toolboxes for Complex Systems. ...

**ISALE** Referenced in 0 articles [[sw19589](#)]

ISALE (Impact-SALE) is a multi-material, multi...

**CoinCalc** Referenced in 1 article [[sw19590](#)]

TOCSY - Toolboxes for Complex Systems: CoinCalc -- A new...

**COPRA** Referenced in 0 articles [[sw19591](#)]

TOCSY - Toolboxes for Complex Systems: COConstructing Proxy Records...

**NESToolbox** Referenced in 1 article [[sw19592](#)]

TOCSY - Toolboxes for Complex Systems: The NESToolbox is...

**IOTA** Referenced in 4 articles [[sw19593](#)]

TOCSY - Toolboxes for Complex Systems: Inner composition alignment...

**Tigramite** Referenced in 0 articles [[sw19594](#)]

TOCSY - Toolboxes for Complex Systems: Tigramite is a...

**SOWAS** Referenced in 3 articles [[sw19595](#)]

TOCSY - Toolboxes for Complex Systems: SOWAS - Wavelet Spectral...

**PETROPY** Referenced in 1 article [[sw19596](#)]

TOCSY - Toolboxes for Complex Systems: PETROPY. Permutation entropy...

**K2** Referenced in 45 articles [[sw19597](#)]

TOCSY - Toolboxes for Complex Systems: K2 - Dynamical Invariants...

# MMS Network and swMATH

What can you do for better software information?

- Please, improve your software citations
- Please, give us information about your software (send an e-mail to: [contact@swmath.org](mailto:contact@swmath.org))

# THANKS!

## Contact

[contact@swmath.org](mailto:contact@swmath.org)

[swmath.org](http://swmath.org)

© FIZ Karlsruhe 2018

Leibniz-Institut für Informationsinfrastruktur GmbH

[www.fiz-karlsruhe.de](http://www.fiz-karlsruhe.de)

These documents are intended for presentation purposes only.  
Copyright lies with FIZ Karlsruhe.  
Any distribution or use of these documents or part thereof is  
subject to FIZ Karlsruhe's express approval.

© FIZ Karlsruhe – Leibniz-Institut für Informationsinfrastruktur GmbH

 **FIZ Karlsruhe**  
Leibniz Institute for Information Infrastructure

  
Leibniz  
Association