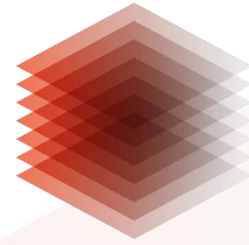

LEIBNIZ INFORMATION CENTRE
FOR SCIENCE AND TECHNOLOGY
UNIVERSITY LIBRARY



TIB

Preserving information on mathematical software via web archives

Helge Holzmann, Mila Runnwerth

San Diego, January 11th 2018

AMS Special Session on Mathematical Information in the Digital Age of
Science, Joint Mathematical Meeting 2018



Why consider software? (1/2)

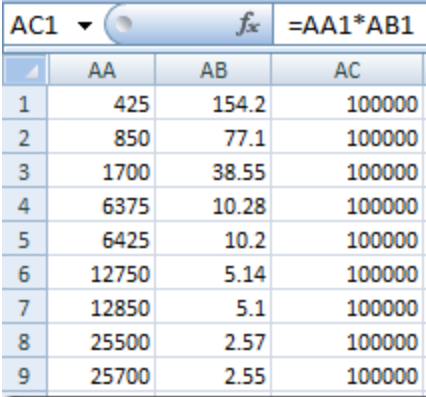
- Improving the status of software and its developers in the scientific publication process.
 - Especially in mathematics scientific software comprises sophisticated knowledge.
- The digital revolution provides the logistics to augment all kinds of data to traditional publications and make all data thoroughly accessible.
- Best practices for software in science:
 - Findable, Accessible, Interoperable, Reusable
 - Replicability, Reproducibility, Reusability
- Checks & balances for the tools we use in scientific work.



Why consider software? (2/2)

What can possibly go wrong?

- Excel bug(s) – the most commonly used mathematical software in non-mathematical science.
- [Knight Capital Group](#).
- Science alert: [“A Bug in FMRI Software Could Invalidate 15 Years of Brain Research”](#)
- Google: disasters caused by software / mathematical errors



	AA	AB	AC
1	425	154.2	100000
2	850	77.1	100000
3	1700	38.55	100000
4	6375	10.28	100000
5	6425	10.2	100000
6	12750	5.14	100000
7	12850	5.1	100000
8	25500	2.57	100000
9	25700	2.55	100000

Goal: Make software visible in science

- If the software code is available on a version controlled repository, always cite the SHA value.
- If the software code is freely available in a less standardised environment
 - Download the software for yourself;
 - Follow the citing instructions on the webpage;
 - Ask the authors to move the software to a version controlled repository;
 - **Create a web archive.**
- If the software code is not available
 - Follow the citing instructions on the webpage;
 - **Create a web archive.**

Evaluation of the situation in maths

Q1: How well is software represented by its surrogate on the web?

Q2: Which information about software is available on the web?

Q3: How many websites of software are archived?

Q4: For how many of these can we recover referenced versions?

Data basis:  swMATH
an information service for mathematical software

A1: High correlation between references in literature and in-links on the web. Good representation.

A2: Software pages are well structured with i. e. a documentation, download, update section.

Result 1: It is worthwhile considering software surrogates on the web.

A3: About half of the webpages have been archived.

A4: Only about 20% of these can be linked to a given referenced version in literature.

Result 2: There is work to be done.

For details see [„Archiving Software Surrogates on the Web for Future Reference”](#) by Holzmann, Sperber, Runnwerth (TPDL 2016).

Coupling swMATH and web archives

References in zbMATH (referenced in 1311 articles , 4 standard articles)

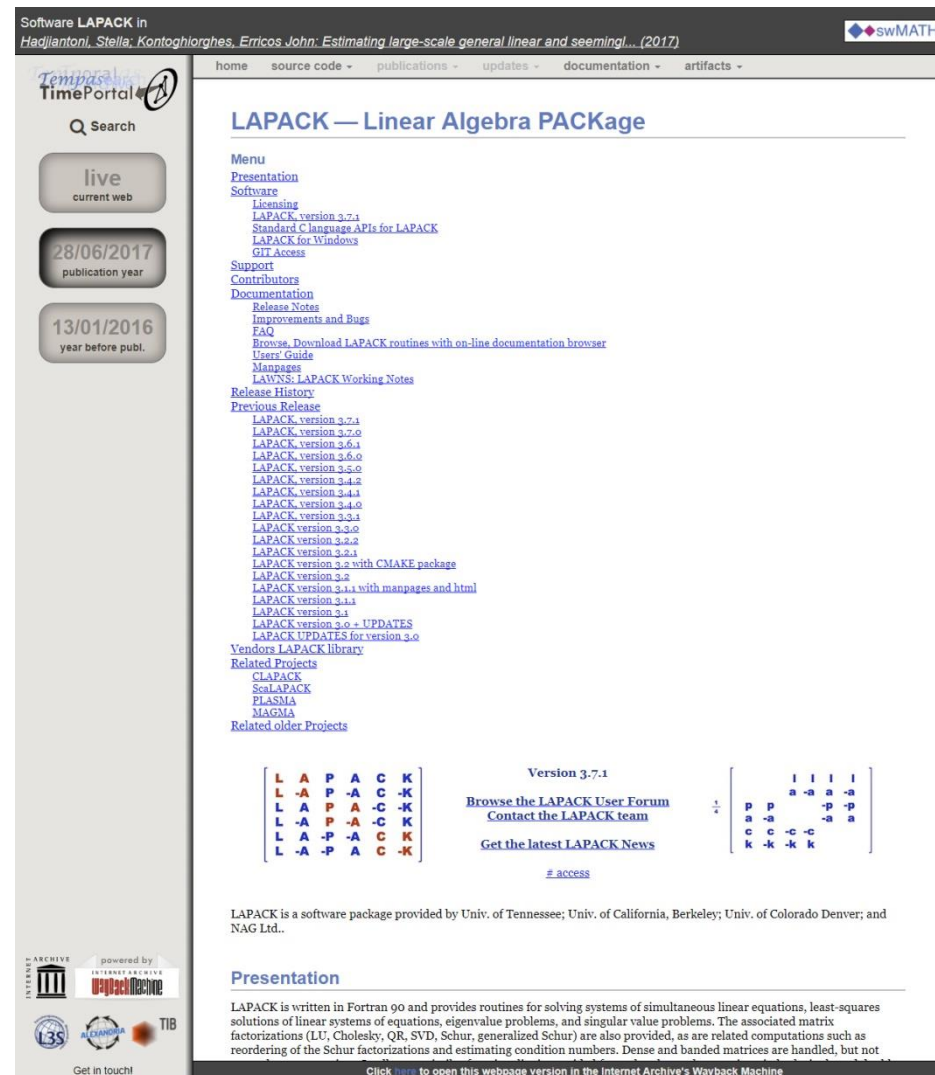
Showing results 1 to 20 of 1311.

Sorted by year (citations)

1 2 3 ... 64 65 66 next

1. Bylina, Beata: The block WZ factorization (2018) [archived SW](#)
2. González-Calderón, Alfredo; Vivas-Cruz, Luis X.; Herrera-Hernández, Erik César: Application of the Θ -method to a telegraphic model of fluid flow in a dual-porosity medium (2018) [archived SW](#)
3. Wang, Xuezhong; Che, Maolin; Wei, Yimin: Partial orthogonal rank-one decomposition of complex symmetric tensors based on the Takagi factorization (2018) [archived SW](#)
4. Amestoy, Patrick; Buttari, Alfredo; L'Excellent, Jean-Yves; Mary, Theo: On the complexity of the block low-rank multifrontal factorization (2017) [archived SW](#)
5. Bezanson, Jeff; Edelman, Alan; Karpinski, Stefan; Shah, Viral B.: Julia: a fresh approach to numerical computing (2017) [archived SW](#)
6. Birgin, E. G.; Martínez, J.M.: The use of quadratic regularization with a cubic descent condition for unconstrained optimization (2017) [archived SW](#)
7. Boiko, Andrey V.; Demyanko, Kirill V.; Nechepurenko, Yuri M.: On computing the location of laminar-turbulent transition in compressible boundary layers (2017) [archived SW](#)
8. Bosner, Nela; Karlsson, Lars: Parallel and heterogeneous m -Hessenberg-triangular-triangular reduction (2017) [archived SW](#)
9. Crivellaro, Alberto; Perotto, Simona; Zonca, Stefano: Reconstruction of 3D scattered data via radial basis functions by efficient and robust techniques (2017) [archived SW](#)
10. Diao, Huai-An: On condition numbers for least squares with quadric inequality constraint (2017) [archived SW](#)
11. Gianluca Frison, Dimitris Kouzoupis, Andrea Zanelli, Moritz Diehl: BLASFEO: Basic linear algebra subroutines for embedded optimization (2017) arXiv [archived SW](#)
12. Glinskiy, Boris; Kuchin, Nikolay; Kostin, Victor; Solovyev, Sergey: Parallel computations for solving 3D Helmholtz problem by using direct solver with low-rank approximation and HSS technique (2017) [archived SW](#)
13. Hadjiantoni, Stella; Kontogiorgos, Erricos John: Estimating large-scale general linear and seemingly unrelated regressions models after deleting observations (2017) [archived SW](#)
14. Huang, Wen; Absil, P.-A.; Gallivan, K.A.: Intrinsic representation of tangent vectors and vector transports on matrix manifolds (2017) [archived SW](#)
15. Jandron, Michael A.; Ruffa, Anthony A.; Baglama, James: An asynchronous direct solver for banded linear systems (2017) [archived SW](#)
16. Maréchal, Alexandre; Périn, Michaël: Efficient elimination of redundancies in polyhedra by raytracing (2017) [archived SW](#)
17. Martinsson, Per-Gunnar; Quintana Orti, Gregorio; Heavner, Nathan; van de Geijn, Robert: Householder QR factorization with randomization for column pivoting (HQRRP) (2017) [archived SW](#)
18. M. N. Gevorkyan, A. V. Demidova, A. V. Korolkova, D. S. Kulyabov, L. A. Sevastianov: The Stochastic Processes Generation in OpenModelica (2017) arXiv [archived SW](#)
19. Ostanin, Igor A.; Zorin, Denis N.; Oseledets, Ivan V.: Fast topological-shape optimization with boundary elements in two dimensions (2017) [archived SW](#)
20. Pouransari, Hadi; Coulier, Pieter; Darve, Eric: Fast hierarchical solvers for sparse matrices using extended sparsification and low-rank approximation (2017) [archived SW](#)

1 2 3 ... 64 65 66 next



Software LAPACK in
Hadjiantoni, Stella; Kontogiorgos, Erricos John: Estimating large-scale general linear and seemingly... (2017)

home source code - publications - updates - documentation - artifacts -

LAPACK — Linear Algebra PACKage

Menu
Presentation
Software
 [Licensing](#)
 [LAPACK version 3.7.1](#)
 [Standard C language APIs for LAPACK](#)
 [LAPACK for Windows](#)
 [GIT access](#)
 [Support](#)
 [Contributors](#)
 [Documentation](#)
 [Release Notes](#)
 [Improvements and Bugs](#)
 [FAQ](#)
 [Browse, Download LAPACK routines with on-line documentation browser](#)
 [Users' Guide](#)
 [Manpages](#)
 [LAWNS: LAPACK Working Notes](#)
 [Release History](#)
 [Previous Release](#)
 [LAPACK version 3.7.1](#)
 [LAPACK version 3.7.0](#)
 [LAPACK version 3.6.1](#)
 [LAPACK version 3.6.0](#)
 [LAPACK version 3.5.0](#)
 [LAPACK version 3.4.2](#)
 [LAPACK version 3.4.1](#)
 [LAPACK version 3.4.0](#)
 [LAPACK version 3.3.1](#)
 [LAPACK version 3.3.0](#)
 [LAPACK version 3.2.2](#)
 [LAPACK version 3.2.1](#)
 [LAPACK version 3.0 with CMAKE package](#)
 [LAPACK version 3.0](#)
 [LAPACK version 3.1.1 with manpages and html](#)
 [LAPACK version 3.1.1](#)
 [LAPACK version 3.1](#)
 [LAPACK version 3.0 - UPDATES](#)
 [LAPACK UPDATES for version 3.0](#)
 [Vendors LAPACK library](#)
 [Related Projects](#)
 [CLAPACK](#)
 [ScalAPACK](#)
 [PLASMA](#)
 [MAGMA](#)
 [Related older Projects](#)

Version 3.7.1

[Browse the LAPACK User Forum](#)
[Contact the LAPACK team](#)

[Get the latest LAPACK News](#)

[# access](#)

LAPACK is a software package provided by Univ. of Tennessee; Univ. of California, Berkeley; Univ. of Colorado Denver; and NAG Ltd.

Presentation

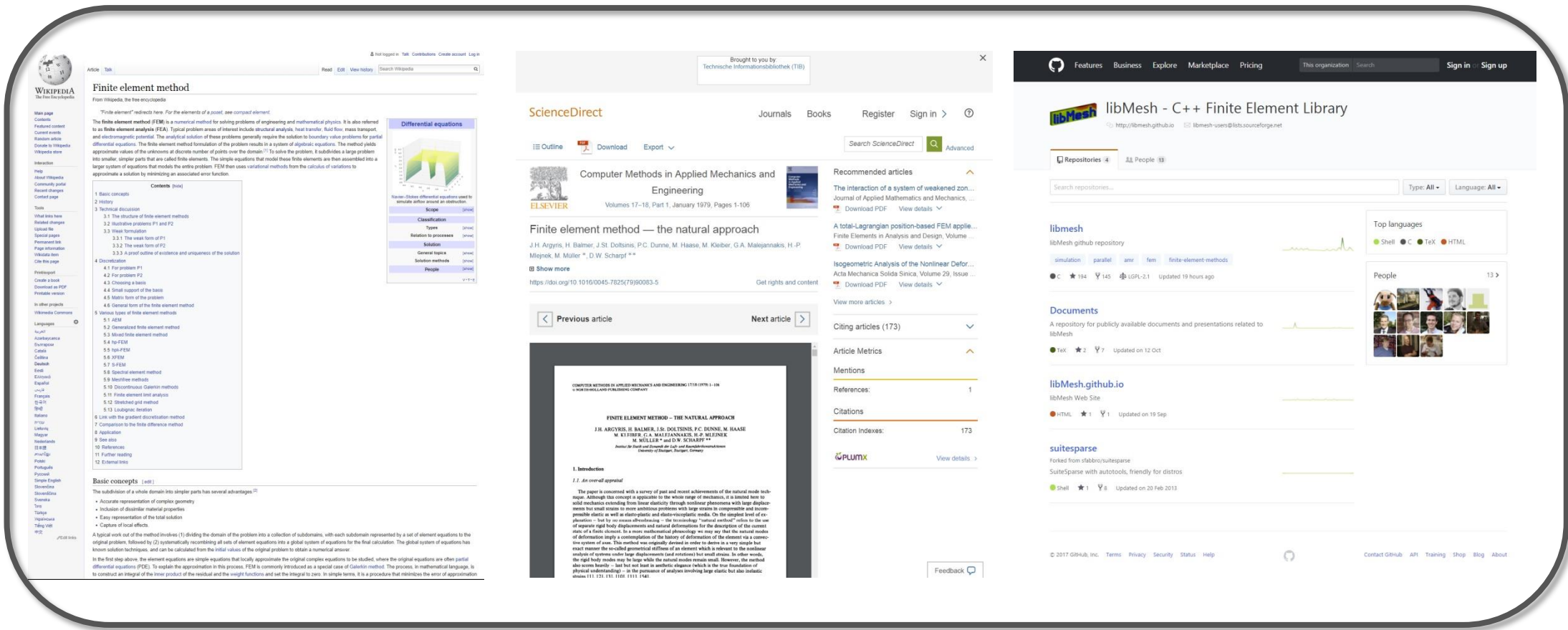
LAPACK is written in Fortran 90 and provides routines for solving systems of simultaneous linear equations, least-squares solutions of linear systems of equations, eigenvalue problems, and singular value problems. The associated matrix factorizations (LU, Cholesky, QR, SVD, Schur, generalized Schur) are also provided, as are related computations such as reordering of the Schur factorizations and estimating condition numbers. Dense and banded matrices are handled, but not

Click [here](#) to open this webpage version in the Internet Archive's Wayback Machine



Make your own (micro) web archive for referencing software: in theory

MyFiniteElementMethod

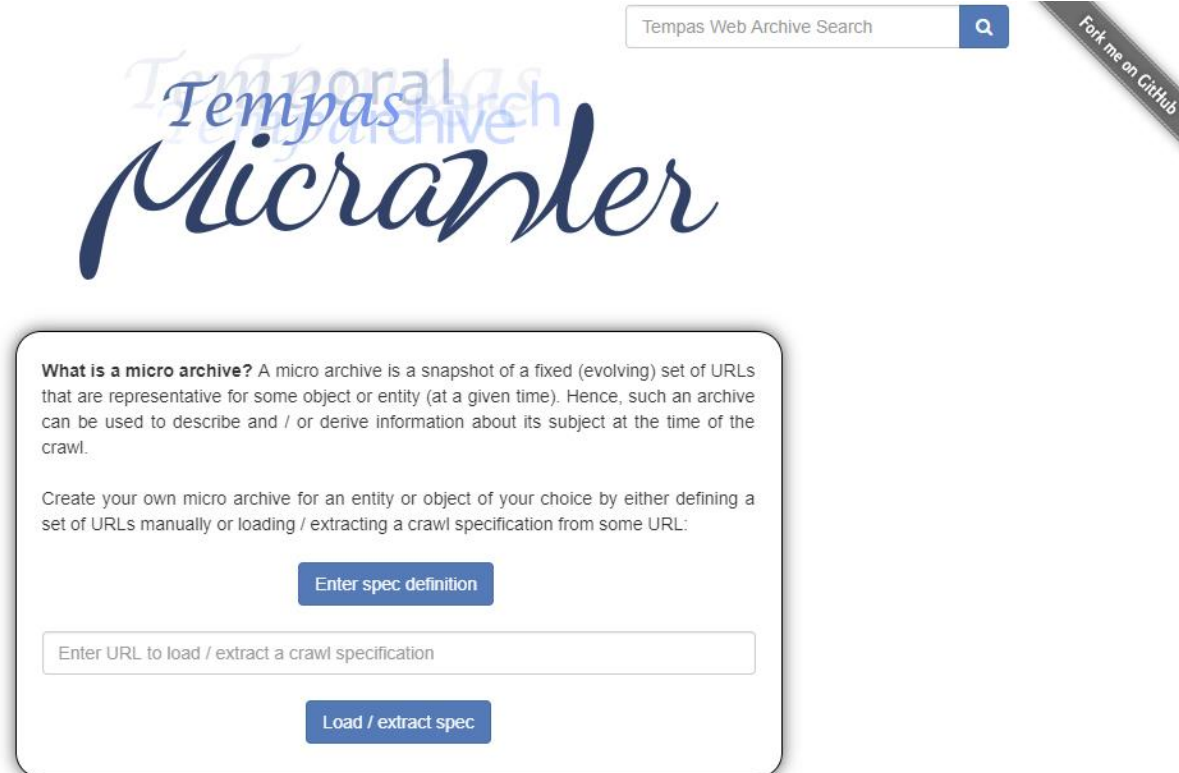


The screenshot displays three web pages related to finite element methods:

- Wikipedia:** The article "Finite element method" provides a general overview, defining it as a numerical method for solving problems of engineering and mathematical physics. It includes a table of contents with sections like "Basic concepts", "Technical discussion", and "Various types of finite element methods".
- ScienceDirect:** The article "Finite element method — the natural approach" by J.H. Argyris, H. Balmer, J.-St. Doltsinis, P.C. Dunne, M. Haase, M. Kleiber, G.A. Malejannakis, H.-P. Mlejnek, M. Müller, and D.W. Scharpf. It is part of the "Computer Methods in Applied Mechanics and Engineering" journal, Volumes 17-18, Part 1, January 1979, Pages 1-106.
- libMesh GitHub:** A repository for the "libMesh - C++ Finite Element Library". It shows repository statistics, top languages (Shell, C, C++, HTML), and a list of documents related to the library.

Make your own (micro) web archive for referencing software: in practice

<http://tempas.l3s.de/micrawler/>



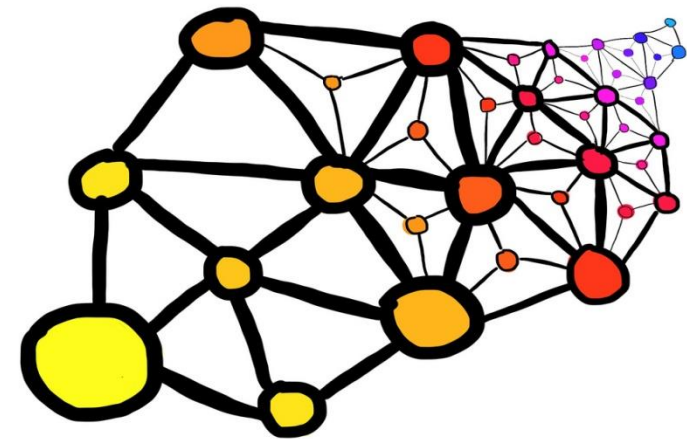
The screenshot shows the 'Tempas Web Archive Search' interface. At the top, there is a search bar with the text 'Tempas Web Archive Search' and a magnifying glass icon. Below the search bar is the title 'Temporal Arch' in a light blue font and 'Micrawler' in a large, dark blue, cursive font. To the right of the search bar is a diagonal button that says 'Fork me on GitHub'. Below the title is a text box with the following text: 'What is a micro archive? A micro archive is a snapshot of a fixed (evolving) set of URLs that are representative for some object or entity (at a given time). Hence, such an archive can be used to describe and / or derive information about its subject at the time of the crawl.' Below this text is another text box: 'Create your own micro archive for an entity or object of your choice by either defining a set of URLs manually or loading / extracting a crawl specification from some URL:'. Below this text box are two buttons: 'Enter spec definition' and 'Load / extract spec'. Below the 'Enter spec definition' button is a text input field with the placeholder text 'Enter URL to load / extract a crawl specification'.

© 2017 L3S Research Center (Helge Holzmann). All rights reserved.

Live Demo

What you gain with Microwler

- The information from the web is archived precisely as you processed it.
- Archive the website when you use its content, not when you reference / cite it.
- All your information from the web is verifiable.
- You create an archive according to your semantic specification.
- Your data is retraceable through time.

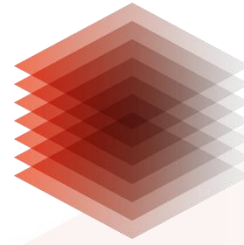


Outlook

- Enable referencing micro archives by handle (i. e. DOI)
 - Pointing to the micro archive from a publication
- Mining meta data from those archives automatically
 - Detect versions, features, etc.
- Consider scientifically relevant web content for web archives



LEIBNIZ INFORMATION CENTRE
FOR SCIENCE AND TECHNOLOGY
UNIVERSITY LIBRARY



TIB

To be continued...

Contact:

[Helge Holzmann](#)

[Mila Runnwerth](#)



Member of

Leibniz
Leibniz Association