



ELSEVIER

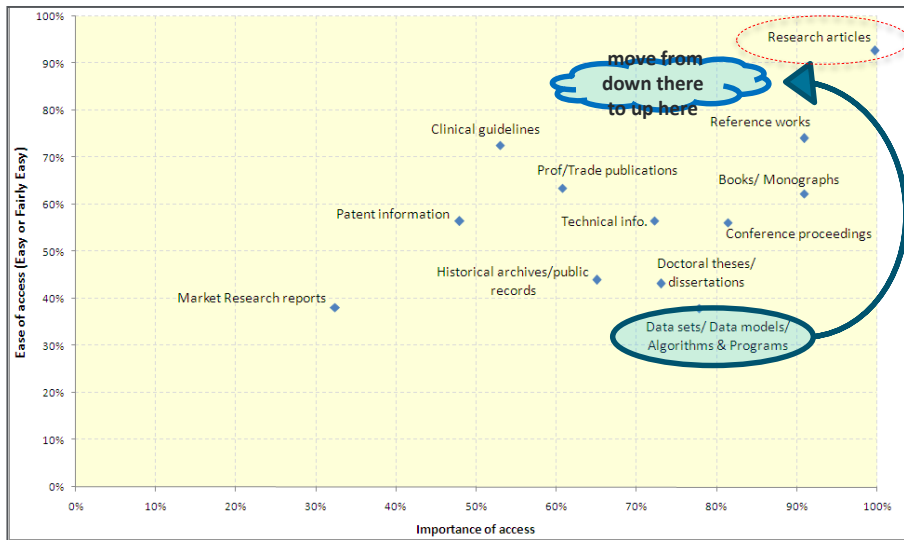
# Publishing software and data

Dr. Chiara Farinelli – Publisher

CERN Workshop, 31 October 2017

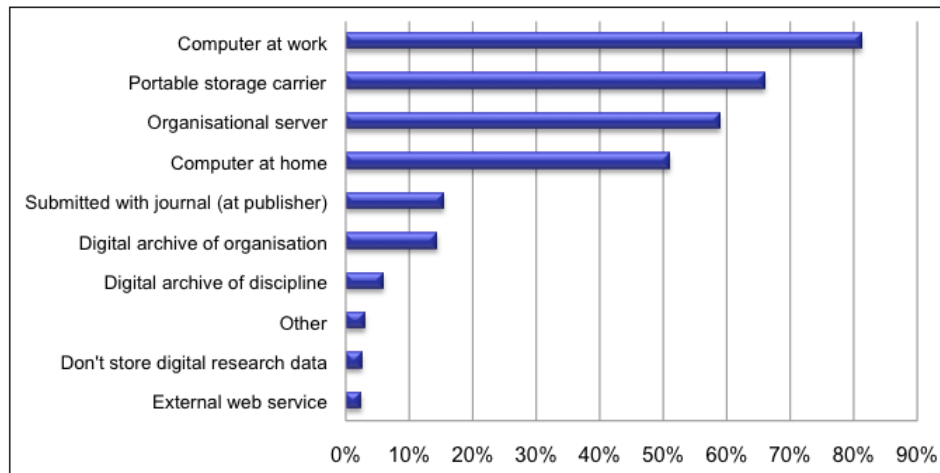


# Why publish data and software?



1. Data and software are important, but hard to access

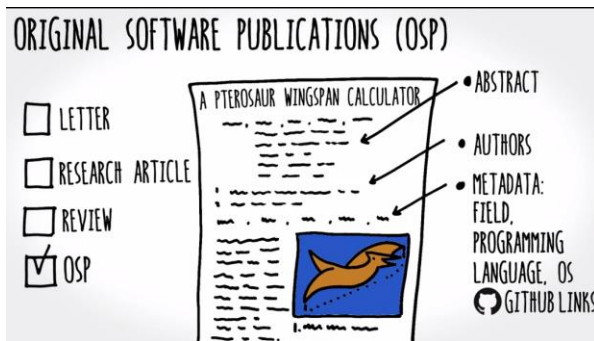
Where do you currently store your research data? (researchers/multiple answers, N=1202)



2. Storage of data and software is very fragmented

## Original Software Publications

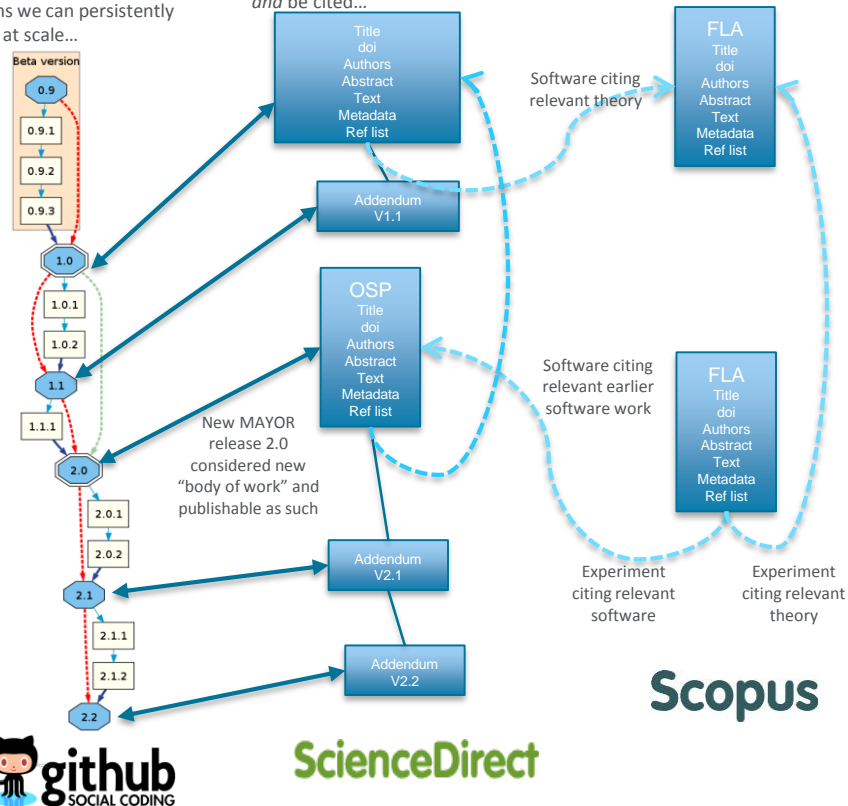
- Every researcher uses software, a big fraction writes software
- Distribution is not organized and authors don't get credit for their work
- It is often considered only a "support" and not part of publishable research.
- **software** makes software indexed, peer reviewed and discoverable. Developers have the same credits of any other author and their work can be found and re-used by others



1) Actual Software & code on developers platform use automated versioning systems we can persistently link to at scale...

2) Though OSP the Software & code can now be searched, indexed, discovered, archived, cite other work *and be cited...*

3) Creating a richer unified ecosystem for open science and discovery ...



- What is a **software** article?
- Submissions to SoftwareX are composed of:
  - A **short article** describing the software, with particular focus on the impact of the software in the research community and its re-usability across disciplines
  - A “**metadata table**” containing information about the software and software metrics



- A **permanent link** to a software repository (GitHub) where the software and code is stored and maintained by Elsevier and made freely available

## Code metadata

Current code version	2016.04.01.0
Permanent link to code/repository used of this code version	<a href="https://github.com/ElsevierSoftwareX/SOFTX-D-16-00022">https://github.com/ElsevierSoftwareX/SOFTX-D-16-00022</a>
Legal Code License	LGPLv3.0
Code versioning system used	git
Software code languages, tools, and services used	VB.NET, C#
Compilation requirements, operating environments & dependencies	.NET Framework 4.5
If available Link to developer documentation/manual	<a href="http://wiki.spectrafox.com">http://wiki.spectrafox.com</a>
Support email for questions	<a href="mailto:contact@spectrafox.com">contact@spectrafox.com</a>

Table options ▼

## Software metadata

Current software version	2016.04.01.0
Permanent link to executables of this version	<a href="https://github.com/ElsevierSoftwareX/SOFTX-D-16-00022">https://github.com/ElsevierSoftwareX/SOFTX-D-16-00022</a>
Legal Software License	LGPLv3.0
Computing platform/Operating System	Microsoft Windows Vista or higher (32- or 64-bit)
Installation requirements & dependencies	.NET Framework 4.5, data acquisition tools: SPECS/Nanonis, Omicron/Matrix, Createc, Nanotec/WSxM
If available, link to user manual—if formally published include a reference to the publication in the reference list	<a href="http://wiki.spectrafox.com">http://wiki.spectrafox.com</a>
Support email for questions	<a href="mailto:contact@spectrafox.com">contact@spectrafox.com</a>


Table options ▼

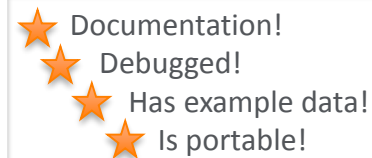
## How to review a Software Article?

- Software articles are judged based mainly on their IMPACT on science and research
- Reviewers are asked to answer a set of questions to simplify their job and assess the value of the submitted software
- Reviewing criteria are public and authors are encouraged to keep them in mind when submitting their work. They can be found at: [http://www.elsevier.com/data/assets/pdf\\_file/0010/97066/ReviewerForm.pdf](http://www.elsevier.com/data/assets/pdf_file/0010/97066/ReviewerForm.pdf) (or via the Guide for Authors)
- Reviewers are not expected to compile, run or debug the code, but they need to be convinced by the author that the software runs correctly. This can be done via videos, screencasts, examples or anything else the author can provide

WORK IN  
PROGRESS



- Test the submitted software (current pilot with  CODE OCEAN BETA)
- Accompany software with example data (on Mendeley Data) – because software and data must go together!
- Introduce a star system to acknowledge the extra step from authors besides peer review, eg.

- 
- ★ Documentation!
  - ★ Debugged!
  - ★ Has example data!
  - ★ Is portable!

# Elsevier & research data

MISSION & POLICY

To help researchers store, share, discover and use data

*“Raw research data should be made freely available to all researchers wherever possible” – STM Brussels Declaration 2007*

Elsevier Research Data Policy

PROGRAMS & PILOTS

Data-linking program



Data Repository



Open Data



Data articles (Data in Brief)

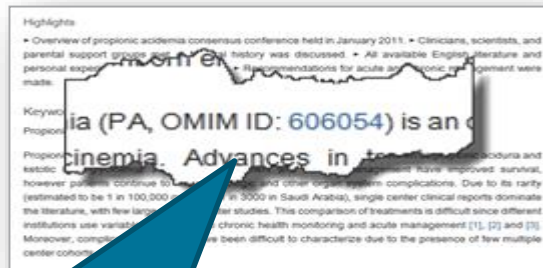


Standards bodies & working

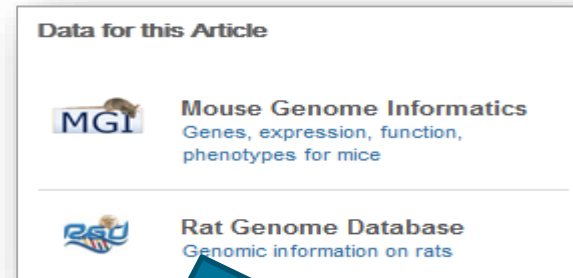


# Data-linking

- Elsevier has an extensive program with 90+ leading domain-specific data repositories to interlink articles and data
- Makes it easier to find relevant data and place data into the right context
- Linking through in-article accession numbers, data DOI's, or data banners



Linking through in-article data accession numbers



Database banners shown next to the article on ScienceDirect



- Place to describe datasets in any scientific discipline.
- Data articles are purely descriptive and do not provide functional data nor interpretation.
- Data articles are intended to facilitate data reuse and reproducibility
- All data described must be made publicly available, either with the article or in a public repository
- All authors fill in a standard data article template available at: <https://www.elsevier.com/dib-template>
- Specifications table with direct link to data
  - Value of the Data
  - Materials and Methods
- All *Data in Brief* articles have CC-BY license

Data in Brief



Data in Brief
  
Volume 11, April 2017, Pages 15–18
Open Access

Data Article  
**Data supporting the effects of lysozyme on mRNA and protein expression in a colonic epithelial scratch wound model**

Sarah K. Abey<sup>a</sup>, Yuana Yuana<sup>a</sup>, Paule V. Joseph<sup>a</sup>, Natnael D. Kenea<sup>a</sup>, Nicolaas H. Fourie<sup>a</sup>, LeeAnne B. Sherwin<sup>a</sup>, Gregory E. Gonye<sup>a</sup>, Paul A. Smyser<sup>a</sup>, Erin S. Stempinski<sup>a</sup>, Christina M. Boulineaux<sup>a</sup>, Kristen R. Weaver<sup>a</sup>, Christopher K.E. Bleck<sup>a</sup>, Wendy A. Henderson<sup>a</sup>  

[Show more](#)

<http://dx.doi.org/10.1016/j.dib.2016.12.043> [Get rights and content](#)  
 Under a Creative Commons license

**Refers To**  
 Sarah K. Abey, Yuana Yuana, Paule V. Joseph, Natnael D. Kenea, Nicolaas H. Fourie, LeeAnne B. Sherwin, Gregory E. Gonye, Paul A. Smyser, Erin S. Stempinski, Christina M. Boulineaux, Kristen R. Weaver, Christopher K.E. Bleck, Wendy A. Henderson  
 Lysozyme association with circulating RNA, extracellular vesicles, and chronic stress  
*BBA Clinical*, Volume 7, June 2017, Pages 23-35  
 PDF (3168 K) | [Supplementary content](#)

**Abstract**  
 Colonic epithelial health is implicated in a host of gastrointestinal (GI) diseases and disorders. Lysozyme is suspected to play a role in the ability of the epithelium to recover from injury (Abey et al., in press; Gallo, 2012; Rubio, 2014) [1], [2] and [3]. Disrupted

links back to related research article

Both articles cite each other in the reference list.

**Abstract**  
*[Briefly describe the contents of this Data article (this should not be the same as a research article abstract). Tell the reader the repository name and reference number for the data here. If data is supplied in the article instead of a repository, please explicitly state this. If the data is related to a published research article, please also directly mention the article in the abstract]*

**Specifications Table** *[please fill in right-hand column of the table below]*

Subject area	Physics, Chemistry, Biology, Economics, Psychology
More specific subject area	Describe narrower subject area
Type of data	Table, image (x-ray, microscopy, etc), text file, graph, figure
How data was acquired	Microscope, survey, SEM, NMR, mass spectroscopy, etc. If an instrument was used, please provide the model and make of the instrument
Data format	Raw, filtered, analyzed, etc
Experimental factors	Brief description of any pretreatment of samples
Experimental features	Very brief experimental description
Data source location	City, Country and/or Latitude & Longitude (& GPS coordinates) for collected samples/data if applicable
Data accessibility	State if data is with this article or in public repository. If public repository, please explicitly name repository and data identification number and provide a direct URL to data

**Value of the data** *[describe in 3-5 bulleted points why this data is of value to the scientific community]*

- 

**Data, Experimental Design, Materials and Methods**  
*[Thoroughly describe your data here. Provide a description of the Experimental design and methods used to acquire or perform base level analysis of the data presented with this article and where applicable. Include any relevant figures/tables needed to fully understand the data. Please also attach, where applicable, any code files used to provide base-level analysis or filtering of the data.]*

**References** *[please include all references relevant to the data described here; references are not limited!]*





## Mendeley Data: Elsevier's data repository

- Active 'data share' section at submission to directly upload datasets to Mendeley
- Research data can be deposited, shared and cited

## Link between dataset/program on Mendeley and journal article:

- CPC programs on Mendeley are linked to the original research article on the CPC journal

Mendeley

Reference Management Research Network **Datasets** Careers Funding

Browse Datasets My Datasets New Dataset FAQ

**Computer Physics Communications**  
Published on behalf of Elsevier  
ISSN: 0010-4655  
Visit journal's website

Datasets associated with articles published in Computer Physics Communications

**Multiple elastic scattering of electrons in condensed matter**  
A. Jablonski  
22 Sep 2016 In: Natural Sciences  
Since the 1940s, much attention has been devoted to the problem of accurate theoretical description of electron transport in condensed matter. The needed information for describing different aspects of the electron transport is the angular distribution of this...

Linked article in: [Computer Physics Communications](#)

**OpenMP, OpenMP/MPI, and CUDA/MPI C programs for solving the time-dependent dipolar Gross-Pitaevskii equation**  
Vladimir Lincov, Luis E. Young-S., Srdjan Stelbic, Patsyony Mangonondam, Sudhan K. Adhikari, Antun Balaz  
14 Oct 2016 In: Natural Sciences  
We present new versions of the previously published C and CUDA programs for solving the dipolar Gross-Pitaevskii equation in one, two, and three spatial dimensions, which calculate stationary and non-stationary solutions by propagation in imaginary or real...

Linked article in: [Computer Physics Communications](#)

ScienceDirect

Journals Books Register Sign in >

Download PDF Export Search ScienceDirect Advanced search

**Computer Physics Communications**  
Volume 210, January 2017, Pages 92-102

**Multiple elastic scattering of electrons in condensed matter**  
A. Jablonski  
Show more  
<https://doi.org/10.1016/j.cpc.2016.08.018> Get rights and content

**Abstract**  
Since the 1940s, much attention has been devoted to the problem of accurate theoretical description of electron transport in condensed matter. The needed information for describing different aspects of the electron transport is the angular distribution of electron directions after multiple elastic collisions. This distribution can be expanded into a series of Legendre polynomials with coefficients,  $A_l$ . In the present work, a database of these coefficients for all elements up to uranium ( $Z=92$ ) and a dense grid of electron energies varying from 50 to 5000 eV has been created. The database makes possible the following applications: (i) accurate interpolation of coefficients  $A_l$  for any element and any energy from the above range, (ii) fast calculations of the differential and total elastic scattering cross sections, (iii) determination of the angular distribution of electrons after multiple collisions, (iv) calculations of the probability of elastic scattering from solids

**Open Data with this article**  
Research data on Mendeley Data  
Multiple elastic scattering of electrons in condensed matter  
Since the 1940s, much attention has been devoted to the problem of accurate theoretical description of electron transport in condensed matter. The needed information for describing different aspects of the electron transport is the angular distribution of electron directions after multiple elastic collisions. This distribution can be expanded into a series of Legendre polynomials with coefficients,  $A_l$ . In the present work, a database of these coefficients for all elements up to uranium ( $Z=92$ ) and a dense grid of electron energies varying from 50 to 5000 eV has been created. The database makes possible the following applications: (i) accurate interpolation of coefficients  $A_l$  for any element and any energy from the above range, (ii) fast calculations of the differential and total elastic scattering cross sections, (iii) determination of the angular distribution of electrons after multiple collisions, (iv) calculations of the probability of elastic scattering from solids

Attached data files:  
MULTI\_SCATT.tar.gz (35 MB)  
Licence: GPLv3  
doi:10.17632/v9y29j8.1  
View dataset on Mendeley >

a link to the program also appears next to the article on SD

Elsevier Research Intelligence

# Back-up slides



# How to submit

- Submissions to *SoftwareX* consist of two major parts:
  - a short descriptive paper of max. 3000 words, max. 6 figures
  - an open source software distribution with support material.
- All articles are published Open Access with OA fee of 500\$
- Authors should include in their article:
  1. **Motivation and Significance** – scientific background and motivation for developing software
  2. **Software Description** – architecture and functionalities
  3. **Illustrative Examples** – at least one example to demonstrate major functions
  4. **Impact** - key factor; should include how widespread software is and citations of results obtained using this software
  5. **Conclusions**



{share code} = create impact

<https://www.journals.elsevier.com/softwarex/submit-your-software/you-can-now-submit-your-software-to-softwarex>

# SoftwareX GitHub repository

SoftwareX  
Your home for Open Access Software publication  
<http://www.journals.elsevier.com/softwarex/> [softwarex@elsevier.com](mailto:softwarex@elsevier.com)

Repositories People 5 Teams 1 Settings

Filters Find a repository... [New repository](#)

**SOFTX-D-16-00113** ★ 0 📄 0  
Excel applications developed to aid high school students' spatial vision when solving tasks on position, intersection, angle and distance of points, lines and planes in simple figures, specifically in rectangular cuboid, rectangle-based right pyramid and regular three- and six-sided prism and pyramid. The figures can be resized and rotated. The ...  
Updated 9 hours ago

**SOFTX-D-16-00127** Matlab ★ 0 📄 0  
The software is a Matlab function that performs a repeated measures ANOVA. Matrices of data as input, together with cell arrays with variable names. Within-subject data are organized as follows: Each row is a subject, and within-subject levels are nested in columns.  
Updated 9 hours ago

People 5 >  
[Invite someone](#)

- **SoftwareX** is maintaining a curated GitHub repository
- A dedicated person is in charge of branching or copying submitted software in the SoftwareX repository
- Elsevier takes responsibility to maintain availability of the submitted software, even if the authors retract/delete their own repository
- Versioning is done following GitHub method
- Each deposited software is peer reviewed (for relevance, re-usability and research value) by independent reviewers
- Whenever possible, documentation is included in the repository