

Cite.Software

Chris Erdmann @scilifelab.se / @michaeljfox.org, Malvika Sharan @turing.ac.uk, Maaïke de Jong @esciencecenter.nl, Jason Maassen @esciencecenter.nl, Lauren Cadwallader @plos.org, Fares Dhane @alliancecan.ca, Maria Cruz @nwo.nl, Carlos Martinez @esciencecenter.nl, Mark Leggott @alliancecan.ca, Rachael Lammey @crossref.org

Shared notes: bit.ly/2023-cite-software

Agenda

1 - Cite.Software Overview (Chris) - ~30 min / Q&A

2 - Software at PLoS/Survey (Lauren) - ~15 min / Q&A

Break 10 min

3 - Turing Way Software Content/Approach (Malvika) - ~ 15 min / Q&A

4 - eScience Center & the Research Software Directory (Maaike/Jason) - ~ 15 min / Q&A

5 - Software Management Plans (Maaike/Jason for Maria) - ~ 15 min / Q&A

6 - Discussion & Next Steps

Shared notes: bit.ly/2023-cite-software

Software Policies and the Game of Telephone

SOFTWARE CITATION POLICIES INDEX

Over the last few years, publishers have been making their Software Citation Policies known either at the publisher level or at the journal level. These policies range in their mandate but many require authors to cite the software central to their findings, following [published guidance](#) from the [FORCE11 Software Citation Implementation Working Group](#).

CHORUS has created a centralized index of these policies with links to the publisher's site. This chart will be updated at least annually.

Please let us know about any updates or requested enhancements by contacting us at info@chorusaccess.org.



Publisher	Policy URL	Policy Type	Date
AAAS*	https://www.sciencemag.org/authors/science-journals-editorial-policies#research-standards	Publisher	
AAS*	https://journals.aas.org/news/policy-statement-on-software/	Publisher	Jan 1, 2016
AGU*	https://www.agu.org/Publish-with-AGU/Publish/Author-Resources/Data-and-Software-for-Authors		June 11, 2021
AMetSociety*	https://www.ametsoc.org/index.cfm/ams/about-ams/ams-statements/statements-of-the-ams-in-force/software-preservation-stewardship-and-reuse/	Publisher	2019
eLife	https://reviewer.elifesciences.org/author-guide/journal-policies		
Elsevier*	https://www.elsevier.com/about/policies/research-data	Publisher	
F1000	https://f1000research.com/for-authors/article-guidelines/research-articles	Publisher	
GigaScience Press	https://academic.oup.com/gigascience/pages/editorial_policies_and_reporting_standards#Availability%20of%20Data%20and%20Materials https://gigabytejournal.com/editorial-policies#availability-of-data-and-materials	Journal	
IOP Publishing*	https://publishingsupport.iopscience.iop.org/questions/research-data/	Publisher	
OUP*	https://academic.oup.com/journals/pages/authors/preparing_your_manuscript/research-data-policy	Publisher	2021
PNAS*	https://www.pnas.org/authors/editorial-and-journal-policies#materials-and-data-availability		

Community Resource for Sharing Software



RESEARCH DATA ALLIANCE

netherlands
eScience center



Digital Research
Alliance of Canada

Alliance de recherche
numérique du Canada

Software Citation



Koen Hufkens, PhD
@koen_hufkens



Recurring gripe: every so often I check citations on my [#rstats](#) packages. Results are always appalling (e.g. {ecmwfr} 11K downloads, 0 citations). So, I'll stress this again until researchers get it in their thick skull. Properly cite the software you use! [#AcademicTwitter](#) 1/

ecmwfr

build	passing	codecov	80%	CRAN	1.2.2	repo status	Active	downloads	11K	DOI	10.5281/zenodo.2647541
-------	---------	---------	-----	------	-------	-------------	--------	-----------	-----	-----	------------------------

Programmatic interface to the two [European Centre for Medium-Range Weather Forecasts](#) API services. The package provides easy access to the 'ECMWF' [web API services](#) and [Copernicus Climate Data Store](#) or 'CDS' from within R, matching and expanding upon the ECMWF python tools.

8:20 AM · May 15, 2020 · Twitter Web App

Credit for Software



Lisa DeBruine 🏳️‍🌈

@LisaDeBruine



Replying to [@andrewang91](#)

I have the opportunity to suggest things to the APA editors in my role in the APA Open Science committee. This is one of the issues I think is important. It's often ECRs creating research software and the lack of credit discourages investment in this.

2:14 PM · Apr 5, 2021 · Twitter Web App

Searching for Software



Konrad Förstner  
@konradfoerstner



One of the core conclusions of the "finding and publishing research software" session is that there is a painful lack of search engines for research software.


[#oscibar](#) [#OpenScience](#) [#ResearchSoftware](#)

8:02 AM · Mar 12, 2018 · Twitter Web Client

Software Citation & Discovery





Software Citation Checklist

Upload Communities Log in Sign up

Zenodo.org will be unavailable for 2 hours on September 29th from 06:00-08:00 UTC. See announcement.

April 21, 2021 Presentation Open Access

Software Citation Checklist

 Erdmann, Christopher;  Stall, Shelley

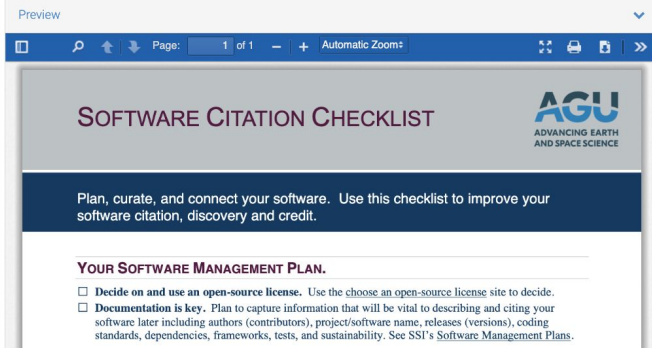
Plan, curate, and connect your software. Use this checklist to improve your software citation, discovery and credit.


This work is part of the **Building New Tools for Data Sharing and Re-use through a Transnational Investigation of the Socioeconomic Impacts of Protected Areas (PARSEC)** project with funding provided by the Belmont Forum through the National Science Foundation, Grant 1929464 as well as the **Accelerating Open and FAIR Data Practices Across the Earth, Space, and Environmental Sciences: A Pilot with the NSF to Support Public Access to Research Data** project funded by the National Science Foundation, Grant 2025364.

Special thank you to Solange Santos, Francine Curivil, and Jennifer Goncalves of SciELO who are members of the PARSEC project and provided recommendations on improvements.

Preview

Page: 1 of 1 Automatic Zoom




SOFTWARE CITATION CHECKLIST 

Plan, curate, and connect your software. Use this checklist to improve your software citation, discovery and credit.

YOUR SOFTWARE MANAGEMENT PLAN.

- Decide on and use an open-source license.** Use the [choose an open-source license](#) site to decide.
- Documentation is key.** Plan to capture information that will be vital to describing and citing your software later including authors (contributors), project/software name, releases (versions), coding standards, dependencies, frameworks, tests, and sustainability. See SSI's [Software Management Plans](#).

417 views 206 downloads [See more details...](#)

Indexed in 

Publication date:
April 21, 2021

DOI:
[DOI 10.5281/zenodo.4706164](https://doi.org/10.5281/zenodo.4706164)

Keyword(s):
[Software Citation](#) [Availability Statement](#)
[Citation File Format](#) [Software Management Plan](#)
[Software Development Platform](#) [Repository](#)

Communities:
AGU Open Science Leadership
PARSEC: Building New Tools for Data Sharing and Re-use through a Transnational Investigation of the Socioeconomic Impacts of Protected Areas

License (for files):
[Creative Commons Attribution 4.0 International](#)

Example: Availability Statements/References



Vides, E. G., Adhikari, A., Chiang, C. Y., Lis, P., Purlyte, E., Limouse, C., Shumate, J. L., Spínola-Lasso, E., Dhekne, H. S., Alessi, D. R., & Pfeffer, S. R. (2022). A feed-forward pathway drives LRRK2 kinase membrane recruitment and activation. In eLife (Vol. 11). eLife Sciences Publications, Ltd. <https://doi.org/10.7554/elife.79771>

Data availability

All primary data associated with each figure has been deposited in a repository; most can be found at <https://doi.org/10.5061/dryad.3tx95x6j7>. Quantitation data of the blots in Figure 3--figure supplement 4 (for the bar graphs in Figures 3C and 3D) can be found at doi (10.5281/zenodo.7057419). Analysis presented in Figure 8--figure supplement 1 can be found at <https://doi.org/10.5281/zenodo.7108943>. All code is available at https://github.com/PfefferLab/Vides_et_al_2022 (copy archived at swh:1:rev:2b50525ee1d48790466d35222956f16615ae96e8).

Vides EG, Pfeffer SR (2022) **Dryad Digital Repository** Data from: A feed-forward pathway drives LRRK2 kinase membrane recruitment and activation.

<https://doi.org/10.5061/dryad.3tx95x6j7>

Limouse C, Vides EG, Adhikari A, Pfeffer SR (2022) **Zenodo** PfefferLab/Vides_et_al_2022: v1.0.

<https://doi.org/10.5281/zenodo.7108943>

Lis P, Alessi DR (2022) **Zenodo** Figure 3--Figure Supplement 4 of the paper 'A Feed-forward Pathway Drives LRRK2 kinase Membrane Recruitment and Activation'.

<https://doi.org/10.5281/zenodo.7057419>

Software Citation Challenges

☰ Menu

🔍 Search entire site



[Home](#) > [Get involved](#) > **Data citation**

Why data citation is important

Data sharing and citation are important for scientific progress. The three key reasons for this are:

- **Transparency and reproducibility:** Most scientific results that are shared today are just a summary of what researchers did and found. The underlying data are not available, making it difficult to verify and replicate results. If data would always be made available with publications, transparency of research would be greatly improved.
- **Reuse:** The availability of raw data allows other researchers to reuse the data. Not just for replication purposes, but to answer new research questions.
- **Credit:** When researchers cite the data they used, this forms the basis for a data credit system. Right now researchers are not really incentivized to share their data, because nobody is looking at data metrics and measuring their impact. Data citation is a first step towards changing that.

How to cite data in your Crossref metadata

Journal Guidance

DOWNLOAD PDF



346 2

Journal Production Guidance for Software and Data Citations

DATA CITATION

PUBLISHER WORKFLOW

SOFTWARE CITATION



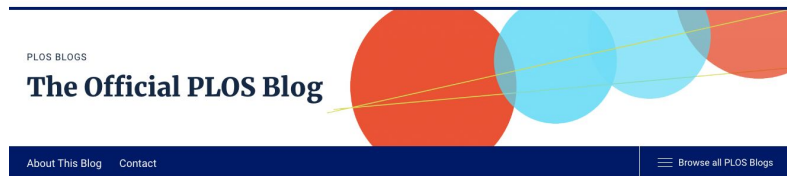
Shelley Stall , Geoffrey Bilder, Matthew Cannon , Neil Chue Hong , Scott Edmunds, Christopher C. Erdmann , Michael Evans , Rosemary Farmer, Patricia Feeny, Michael Friedman , Matthew Giampoala , R. Brooks Hanson , Melissa Harrison , Dimitris Karaiskos, Daniel S. Katz , Viviana Letizia , Vincent Lizzi , Catriona MacCallum, August Muench , Kate Perry , Howard Ratner , Uwe Schindler , Brian Sedora , Martina Stockhause , Randy Townsend, Jake Yeston, Timothy Clark

Abstract

Software and data citation are emerging best practices in scholarly communication.

This article provides structured guidance to the academic publishing community on how to implement software and data citation in publishing workflows. These best practices support the verifiability and reproducibility of scientific results; sharing and reuse of valuable data and software tools, and attribution to the creators of the software and data.

PLoS Open Science Indicators & Badges



Explore the first Open Science Indicators dataset—and share your thoughts

December 12, 2022 / PLOS / [Open Code](#) [Open Data](#) [Open Science](#) [Open Science Indicators](#) [Preprints](#)

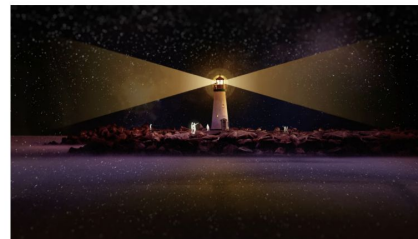


Written by Lauren Cadwallader, Lindsay Morton, and Iain Hrynaskiewicz

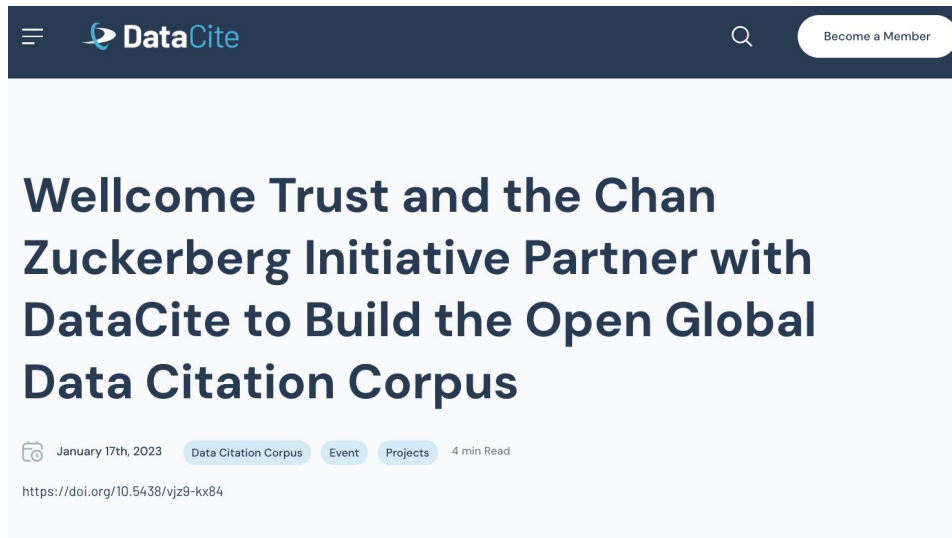


PLOS launches new feature to promote data sharing and access

March 29, 2022 / PLOS / [In the News](#) [Open Data](#) [Open Science](#)



Data Citation Corpus



The screenshot shows the top navigation bar of the DataCite website, which is dark blue with the DataCite logo and a search icon. Below the navigation bar is a large white area containing the article title, metadata, and a brief introductory paragraph. The article title is 'Wellcome Trust and the Chan Zuckerberg Initiative Partner with DataCite to Build the Open Global Data Citation Corpus'. The metadata includes the date 'January 17th, 2023', tags for 'Data Citation Corpus', 'Event', and 'Projects', and a '4 min Read' indicator. A DOI link is provided below the metadata. The introductory paragraph is partially visible at the bottom of the screenshot.

☰ DataCite 🔍 Become a Member

Wellcome Trust and the Chan Zuckerberg Initiative Partner with DataCite to Build the Open Global Data Citation Corpus

📅 January 17th, 2023 Data Citation Corpus Event Projects 4 min Read

<https://doi.org/10.5438/vjz9-kx84>

Aggregated references to data across outputs will help the community monitor impact, inform future funding, and improve the dissemination of research



Paul Vierkant

Share this:

Mining for Software

New data reveals the hidden impact of open source in science

Understanding software used by scientists by mining the biomedical literature



Chan Zuckerberg Initiative Science · Follow

Published in CZI Technology · 8 min read · Sep 28, 2022



66



At the Chan Zuckerberg Initiative (CZI), we believe open source tools are critical to accelerating scientific discovery. In an effort to improve our understanding of the impact of software (and scientific open source in particular) in biomedical science, we're releasing the CZ Software Mentions Dataset — a dataset entirely composed of software mentions mined from

OA Report

Find an organization by name, acronym, or ROR...

About

PUT OA POLICY INTO PRACTICE AT
Michael J. Fox Foundation

GET DATA IN CSV BY EMAIL

For January 1, 2023 — September 10, 2023

you@example.com

Send CSV

2023 2022 2021 All time

INSIGHTS ON ARTICLES

Total published

383

Free-to-read

88%

336 of 383 articles

OA policy-compliant

3%

1 of 35 articles covered

Open Access

60%

231 of 383 articles

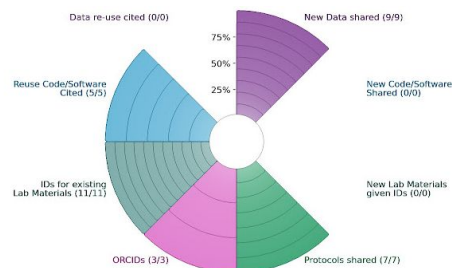
With data availability statements

60%

154 of 258 articles checked



DataSeer



DataCite Commons

Software Sharing & Context

scite_

Search by title, author, key... 

Install **extension!** 

Assistant

Product 

Discover 

Pricing

Blog

Sign in


Sign Up


Ask a question, get an answer backed by real research

Start your free 7-day trial

[Book a demo](#) →

assistant by scite_

 How does the structure of a protein affect its function?

 Proteins are essential for the proper functioning of living organisms, and their structure plays a crucial role in their function. The structure of a protein is determined by its amino acid sequence, which in turn affects its folding and stability [6]. The interactions between charged amino acids significantly influence the structure and function of proteins [1]. The number of hydrophobic methylenes linking the backbone to the charged functionality of amino acids affects protein structure stability [1]. Protein secondary structures, such as β -sheet structures, are the fundamental three-dimensional building blocks of protein structures [1]. Intrinsically disordered proteins (IDPs) are a family of proteins that fail to form rigid 3-D structures under physiological conditions, either along their entire lengths or only in localized regions [2]. IDPs exist as dynamic ensembles within which atom positions and backbone Ramachandran angles exhibit extreme temporal fluctuations without specific equilibrium values [2].

What do you need help with? 

References 

Reference #1 ...
Longer charged amino acids favor β -strand formation in hairpin peptides

Jingyuan Chang¹, Jing-Yuan Chang², Nian-Zhi Li³ et al. 20213 Pep Sci

 3  1  7  0

Interactions between charged amino acids significantly influence the structure and function of proteins. The encoded charged amino acids Asp, Glu, Arg, and Lys have different number of hydrophobic methylenes linking the backbone to the charged functionality...

[View full text](#) [Add to Dashboard](#)

Reference #2 ...

Understanding protein non-folding
Vladimir N. Uversky¹, A. Keith Dunker² 2010
Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics

 1017  24  112  0

This review describes the family of intrinsically disordered proteins, members of which fail to form rigid 3-D structures under physiological conditions, either along their entire lengths or only in localized regions...

[View full text](#) [Add to Dashboard](#)

Can we provide researchers with clear instructions based on their goals?

Scenarios:

- Journal requirement
- Reproducibility
- Credit/T&P (software vs paper or both)
- Discovery/use
- Collaboration/community (and learn from others)

Researchers need something that is...

- Easy to use
 - Clear (step by step instructions according to goals)
 - Demonstrates the impact/what will happen by doing X
 - Common solution, shared and supported by stakeholders
- ... and ultimately something that is automated.

The Turing Way - Community Resource



The Turing Way

🔍 Search this book...

Welcome

- Guide for Reproducible Research ▾
- Guide for Project Design ▾
- Guide for Communication ▾
- Guide for Collaboration ▾
- Guide for Ethical Research ▾
- Community Handbook ▾
- Afterword ▾

Visit our [GitHub Repository](#)

This book is powered by [Jupyter Book](#)



☰ Contents

Our Community

History

Citing *The Turing Way*

Welcome

Welcome to The Turing Way handbook to reproducible, ethical and collaborative data science.

The Turing Way project is open source, open collaboration, and community-driven. We involve and support a diverse community of contributors to make data science accessible, comprehensible and effective for everyone. Our goal is to provide all the information that researchers and data scientists in academia, industry and the public sector need to ensure that the projects they work on are easy to reproduce and reuse.

💡 Top Tip

The Turing Way is not meant to be read from start to finish. Start with a concept, tool or method that you need now, in your current work. Browse the different guides that make up the book, or use the search box to search for whatever you would like to learn about first.

All stakeholders, including researchers, software engineers, project leaders and funding teams, are encouraged to use *The Turing Way* to understand their roles and responsibility of reproducibility in data science. You can inspect our resources on [GitHub](#), contribute to the project as described in our [contribution guidelines](#) and re-use all materials ([see the License](#)).

Research Software Directory - eScience Center



Q Search or jump to...

[Software](#) [Projects](#) [Organisations](#)

[Feedback](#) ∨ [Sign in](#)

Show your research software to the world

The [Research Software Directory](#) is designed to show the impact research software has on research and society. We stimulate the reuse of research software and encourage proper citation of research software to ensure researchers and RSEs get credit for their work.



Journal Guidance/Support

PLOS ONE

PUBLISH

ABOUT

BROWSE

SEARCH



advanced search

Sharing Materials

Sharing Software

Sharing Code

Materials, Software and Code Sharing

The following policy applies to all PLOS journals, unless otherwise noted.

 **Related policy:** [Data Availability](#)

PLOS is committed to ensuring the availability of materials that underpin research. Sharing materials encourages reuse and facilitates reproducibility.

PLOS reserves the right to issue a correction, expression of concern, or retraction if unreasonable restrictions on sharing are discovered after publication. PLOS will contact authors' institutions and funders as appropriate.

Funder Guidance/Support



[Contact](#) [Nederlands](#) [Read aloud](#) [ISAAC](#)  


[Find funding](#)  [Research policy NWO](#)  [Research & results](#)  [News & events](#)  [About NWO](#) 

[News & events](#) [News](#) [A national guide to research software management](#)

A national guide to research software management

27 October 2022

The Practical Guide to Software Management Plans, released today by the Dutch Research Council (NWO) and the Netherlands eScience Center, offers guidance for research institutions, research groups, as well as individual researchers, on how to manage research software produced as an output of research projects.

[→ Download guide](#) 

Characteristics

Themes

[Open Access](#)

[Open Science](#)

Type

[Policies](#)



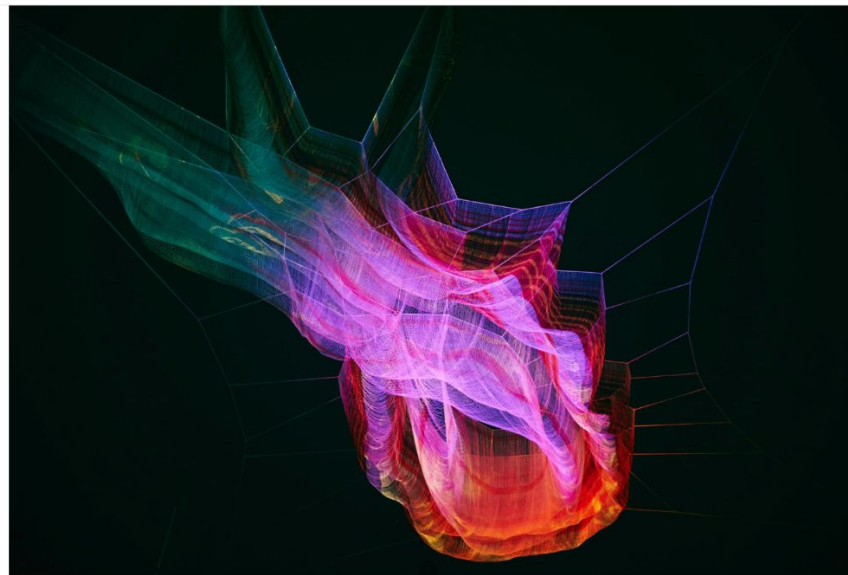
Research Organization Guidance/Support

A data-centric approach to Life Science

Life science research is increasingly becoming not only technology-driven, but also data-driven. SciLifeLab coordinates and supports activities throughout the life cycle of data, from project planning, data production, data analysis, data sharing, to publishing and reuse of data, where researchers are dependent on advanced data analysis and e-infrastructures.

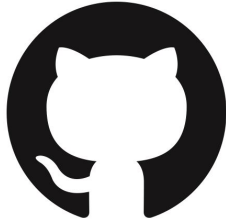
SciLifeLab Data Centre is a central unit within SciLifeLab with responsibility for IT- and data management issues, serving the SciLifeLab and the Data Driven Life Science (DDLs) research program.

At SciLifeLab, we see data as one of the most valuable and long-lasting products of our operations and strive to make our data FAIR, handled according to open science standards and that its long-term value to the scientific community is maximized.



Cite.Software Phases

Phase 1



+ Your Support

Phase 2



+ Your Support

Our proposal

See draft:

https://docs.google.com/document/d/10wyRdqIuJSW557e7jAyu2aeW9hTvHIL0_QgR99U6d8g/edit

Your Support? Your Ideas?

Thank you & discussion

Lead Contact:

Chris Erdmann, Head of Open Science, SciLifeLab

christopher.c.erdmann@gmail.com