

Better Software, Better Research



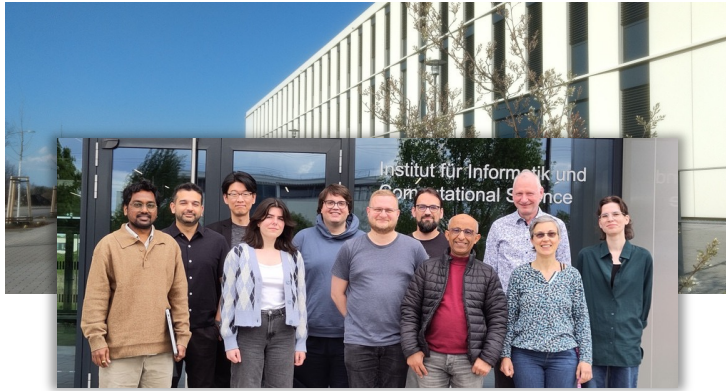
Picture from <https://pixabay.com/de/illustrations/labor-wissenschaft-forschung-8895975/>

~~How Research Software Engineering (RSE) Improves Scientific Practices~~

Let's Advance Research Software Engineering (RSE) Together!

Anna-Lena Lamprecht, BB Research Software Day 2026

About me



Prof. Dr. Anna-Lena Lamprecht

Chair of Software Engineering

Institute of Informatics and Computational Science

University of Potsdam (Campus Golm)

<https://www.uni-potsdam.de/en/cs-se/index>

- Studied “Applied Computer Science” in Göttingen 2002-2007
- PhD in Computer Science (Dortmund 2012)
- PostDoc in Potsdam 2012-2015
- Research Fellow in Limerick 2015-2017
- Assistant Professor in Utrecht 2017-2022
- Member of de-RSE, NL-RSE and the International Council of RSE Associations
- Speaker of the „Fachgruppe RSE“ of Gesellschaft für Informatik (GI)
- Open-Science-“Beauftragte“ of the Science Faculty in Potsdam
- Member of the DFG “Ausschuss für Wissenschaftliche Bibliotheken und Informationssysteme (AWBI)”
- ...



Research needs Software

92%

of researchers
use software

67%

say its fundamental
for their research

56%

develop their own
research software

Brett et al.: *Research Software Engineers: State of the Nation Report*, 2017, <https://doi.org/10.5281/zenodo.495360>

Research Software



*"Research Software includes source code files, algorithms, scripts, computational workflows and executables that were **created during the research process or for a research purpose.**"*

(Gruenpeter et al., *Defining Research Software: a controversial discussion*, 2021, <https://doi.org/10.5281/zenodo.5504016>)

Research Software Engineering

Term coined ~2012.

Emerging, cross-cutting discipline.

Combines:

- Computer science
- (Computational) science disciplines
- Open science practices

Diverse, international community.

Organized in national networks and an international council.



RSE Community



<https://de-rse.org/>

(mailing list, online community on Matrix, working groups, events, ...)



<https://fg-rse.gi.de/>

Joint special interest group of German RSE and Computer Science (Software Engineering) communities



<https://researchsoftware.org/>

(links to national communities, coordination for International RSE Day, International Council, ...)



<https://www.researchsoft.org/>

(policy activities, infrastructure perspective, thematic forums and task forces, monthly newsletter ...)

Research Software Engineering

Shares many commonalities with “mainstream” software engineering.

Specific challenges through:

- Nature of scientific research
- Cultural environment of scientific software development

Increasingly receives separate attention.

Software Engineering for Science Workshop Series

Workshop Workshops and Reports

- 2021

- [2021 International Workshop on Software Engineering for Science \(SE4Sci\)](#)
Held in conjunction with SC21

C. Goble, **Better Software, Better Research**, in IEEE Internet Computing, vol. 18, no. 5, pp. 4-8, 2014, doi: 10.1109/MIC.2014.88.

- 2020

- [2020 International Workshop on Software Engineering for Science \(SE4Sci\)](#)
Held in conjunction with SC20

A. Johanson and W. Hasselbring, **Software Engineering for Computational Science: Past, Present, Future**, in Computing in Science & Engineering, vol. 20, no. 2, pp. 90-109, 2018, doi: 10.1109/MCSE.2018.021651343.

- 2019

- [2019 International Workshop on Software Engineering for Science \(SE4Sci\)](#)
Held in conjunction with SC19

Anzt H, Bach F, Druskat S et al. **An environment for sustainable research software in Germany and beyond: current state, open challenges, and call for action**. F1000Research 2021, 9:295, doi: 10.12688/f1000research.23224.2

- 2018

- [The 2018 International Workshop on Software Engineering for Science \(SE4Sci\)](#)
Held in conjunction with Juniper

- 2017

- [2017 International Workshop on Software Engineering for Science \(SE4Sci\)](#)
Held in conjunction with Juniper

RSE Themes

RSE Practice

Best practices and tools for day-to-day research software development

RSE Training

Developing (R)SE skills in researchers and R(SE) skills in software developers

RSE Infrastructure

Support for developing, executing and maintaining research software

RSE Community

RSE Careers

Developing RSE as an own professional profile, and career paths for RSEs.

RSE Advocacy

Working towards institutional support, funding and recognition of RSE and RSEs

RSE Research

Analyzing and improving (the development process of) research software

RSE Research



*"Research software engineering research (RSE Research) aims at **understanding** and **improving** how software is developed for research."*

(Felderer et al., *Investigating Research Software Engineering: Toward RSE Research*, 2025, <https://doi.org/10.1145/3685265>)

RSE Research Example

"How do researchers develop data analysis workflows, and how can they be supported better?"

```
1 # load required packages
2 library("Biobase")
3 library("affy")
4 library("AffyExpress")
5 library("limma")
6 library("hgu95av2.db")
7 library("XML")
8 library("annotate")
9
10 # load input data
11 celpath <- "spikein/"
12 phenodatafile <- "spikein.pdata.txt"
13 phenodata <- read.table(phenodatafile, row.names = 1, header
14 = TRUE, sep = "\t")
15 affybatch <- ReadAffy(celfile.path=celpath, phenoData =
16 phenodata)
17
18 # AffyExpress: preprocessing
19 exprset <- pre.process(method = "rma", raw = affybatch, plot
20 = FALSE)
21
22 # AffyExpress: filtering
23 filteredexprset <- Filter(exprset, numChip = 3, bg = 7.0)
24
25 # differential expression analysis using limma
26 fit <- lmFit(exprset)
27 fit <- eBayes(fit)
28 toptable <- topTable(fit, adjust = "fdr")
29
30 # get PubMed abstracts for top genes
31 genenames <- as.character(toptable$ID)
32 abstracts <- pm.getabst(genenames, "hgu95av2")
33 abstracts <- unlist(abstracts, recursive = FALSE)
34 pmAbst2HTML(abstracts, filename = "spikein-abstracts.html",
35 frames = FALSE)
```



COMMON
WORKFLOW
LANGUAGE

nextflow

Workflows: orchestrations ("glue code") of calls to different tools, carrying out a computerized scientific (data analysis) process.

Analytical Variability

Article | [Published: 20 May 2020](#)

Variability in the analysis of a single neuroimaging dataset by many teams

[Rotem Botvinik-Nezer](#), [Felix Holzmeister](#), ... [Tom Schonberg](#)  [+ Show authors](#)

[Nature](#) **582**, 84–88 (2020) | [Cite this article](#)

46k Accesses | **269** Citations | **2061** Altmetric | [Metrics](#)

Advances in Methods and Practices in Psychological Science
Volume 1, Issue 3, September 2018, Pages 337-356
© The Author(s) 2018, Article Reuse Guidelines
<https://doi.org/10.1177/2515245917747646>



Empirical Article

**Many Analysts, One Data Set: Making Transparent
How Variations in Analytic Choices Affect Results**

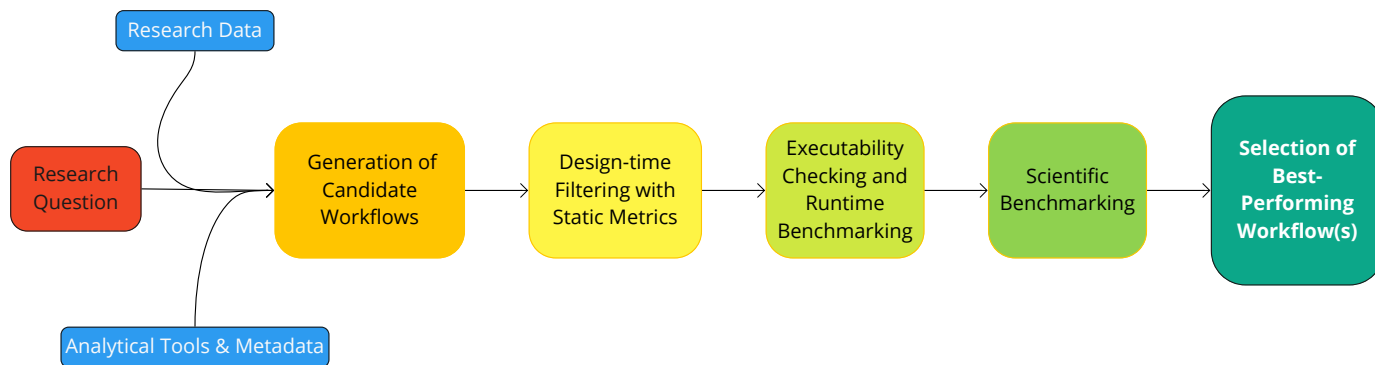
Researchers arrive at disparate analysis results for the same dataset.

Why? Some reasons:

- Enormous variability of possible data analysis workflows.
- Identifying the best workflows for a research question is challenging.
- Lack of tool support leads researchers to rely on familiar tools, intuition, and personal preferences.

Workflomics

(Kasalica et al., *Tackling Analytical Variability with Workflomics*, 2026, <https://doi.org/10.14279/depositonce-25819>)

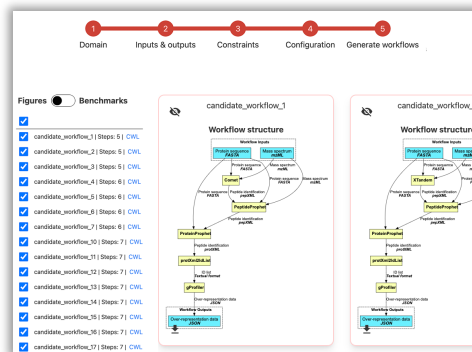


WORKFLOWMICS

Generate and Benchmark workflows
Explore bioinformatics workflows generated according to your description
EXPLORE

Visualize benchmarks
Upload and visualize the benchmarks performed locally
BENCHMARK

How to use Workflomics?
Explore the Workflomics documentation
USER GUIDE



Workflow	Domain	Inputs & outputs	Constraints	Configuration	Generate workflows	Execution time	Memory usage	Network usage	Storage usage	Overall score
candidate_workflow_1	1	1	1	1	1	1	1	1	1	1
candidate_workflow_2	1	1	1	1	1	1	1	1	1	1
candidate_workflow_3	1	1	1	1	1	1	1	1	1	1
candidate_workflow_4	1	1	1	1	1	1	1	1	1	1
candidate_workflow_5	1	1	1	1	1	1	1	1	1	1
candidate_workflow_6	1	1	1	1	1	1	1	1	1	1
candidate_workflow_7	1	1	1	1	1	1	1	1	1	1
candidate_workflow_8	1	1	1	1	1	1	1	1	1	1
candidate_workflow_9	1	1	1	1	1	1	1	1	1	1
candidate_workflow_10	1	1	1	1	1	1	1	1	1	1
candidate_workflow_11	1	1	1	1	1	1	1	1	1	1
candidate_workflow_12	1	1	1	1	1	1	1	1	1	1
candidate_workflow_13	1	1	1	1	1	1	1	1	1	1
candidate_workflow_14	1	1	1	1	1	1	1	1	1	1
candidate_workflow_15	1	1	1	1	1	1	1	1	1	1
candidate_workflow_16	1	1	1	1	1	1	1	1	1	1
candidate_workflow_17	1	1	1	1	1	1	1	1	1	1

COMMON
WORKFLOW
LANGUAGE

<https://workflomics.org>

FONDA C2: Understanding Workflow Design Better

Ongoing work (with Mendling group, HU Berlin):

- Interview study with workflow developers
- Focusing on (re)use of research software for data analysis workflows

Preliminary results:

- **Environment, personal positioning and project origin** factors seem to influence software use more than technical factors
- Researchers do consider **consequences** of using particular software



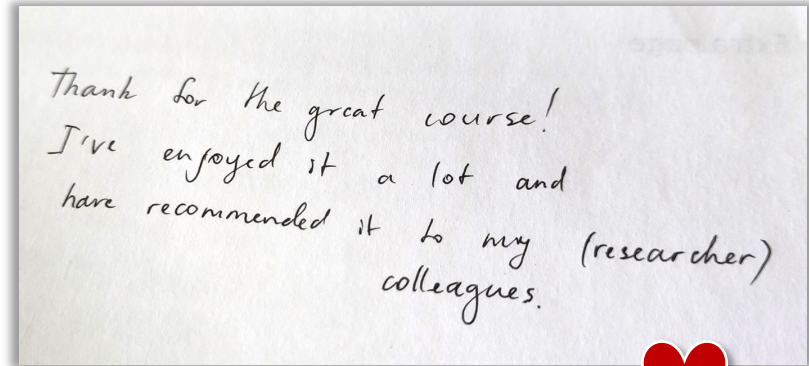
**FONDA - Foundations of
Workflows for Large-Scale
Scientific Data Analysis**

(DFG Collaborative Research
Center 1404 at HU Berlin)

*Subproject C2: Early Workflow
Design: From Collaborative
Scientific Problem-Solving to
DAW Specifications*

RSE Training: University Course

- Taught at University of Potsdam since 2023
- Full semester, worth 6 ECTS credit points
- Assumes only standard proficiency with Python (for data analysis)
- RSE concepts and a lot of practice
- Inspired by Irving et al. “Research Software Engineering with Python” (<https://third-bit.com/py-rse/>)
- Extended by further topics (work in progress, available at <https://github.com/SE-UP/RSE-UP>)
- Each year between 60-90 participants from different disciplines and programs



(at the back of an exam)



(See also: Bertrand et al., *Compared Experiences from Teaching Full-Semester Research Software Engineering Courses at Four German Universities*, 2025, <https://doi.org/10.14279/eceasst.v83.2611>)

RSE Training: University Course

Lecture

(clean code, documentation, Jupyter notebooks, computational narratives, testing and error handling, working in teams, software requirements, software architectures, replication crisis, making tools (re)usable, building command-line tools, configuration and packaging, data analysis workflows with Snakemake, software citation, FAIR, Open Science)

Labs

(Python recap, Shell, Git, CI/CD, project support, presentations, discussions)

Exam

Individual Project

(computational narrative for a self-selected destatis dataset, implemented in Jupyter notebook)

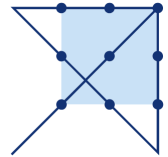
Group Project

(data analysis workflow for self-selected problem involving a weather/climate dataset from the DWD Climate Data Center (CDC), implemented in Snakemake)

Towards an RSE Master Curriculum

- Joint project of deRSE and GI, funded by Klaus Tschira Foundation, coordinated by Dr. Julian Dehne (group lead RSE at GI)
- Goal: development of a reference curriculum for RSE Master education
- Status:
 - Series of thematic community workshops for content definition held
 - Curriculum draft in discussion
 - Piloting new RSE-specific courses
- Working repository:
<https://github.com/the-teachingRSE-project/RSE-Masters>

**Klaus Tschira
Stiftung**



GESELLSCHAFT
FÜR INFORMATIK



SOCIETY FOR
RESEARCH SOFTWARE

Recap

- Research Software and RSE are vital to contemporary research.
- The RSE community works to advance this new discipline:
 - RSE Advocacy
 - RSE Careers
 - RSE Infrastructure
 - RSE Practice
 - RSE Research
 - RSE Training
- **Join and contribute where you can!**



deRSE26 Conference in Stuttgart, [photo](#) by [Adam Pagan](#)

Thank You For Your Attention!

... and engage with the community:

- deRSE e.V. (<https://de-rse.org/>)
- Fachgruppe RSE (<https://fg-rse.gi.de/>)
- ReSA news(letter) (<https://www.researchsoft.org/news/>)
- Conferences (e.g. deRSE Collaborations Workshop 2026 in September in Göttingen, <https://events.hifis.net/event/3249/overview>)

Acknowledgments

The work presented here is the result of collaborations with many great people and a number of funded projects.

To name just some of the great people:

Nikolas Bertrand, Julian Dehne, Mario Frank, Charlotte Grunert, Paul-Julius Hillmann, Vedran Kasalica, Peter Kok, Hyeokjin Kwon, Rob Marissen, Sara Mohammedahmed, Jan Mendling, Sebastian Müller, Magnus Palmblad, Veit Schwämmle, Maja Toebs



netherlands
eScience center

LU
MC Leiden University
Medical Center



FONDA

DFG
Deutsche
Forschungsgemeinschaft

**Klaus Tschira
Stiftung**

