Handling research software: recommendations to users, developers and research managers

Jürgen Fuhrmann, Weierstrass Institute Berlin

https://doi.org/10.5281/zenodo.1172969

Leading authors: Dr. Matthias Katerbow (DFG), Dr. Georg Feulner (PIK Potsdam)

Contributing authors: Mathias Bornschein, Prof. Dr. Björn Brembs, Dr. Michael Erben-Russ, Dr. Konrad Förstner, Michael Franke, Dr. Bernadette Fritzsch, Dr. Jürgen Fuhrmann, Prof. Dr. Michael Goedicke, Stephan Janosch, Dr. Uwe Konrad, Dennis Zielke.

Contents

Background

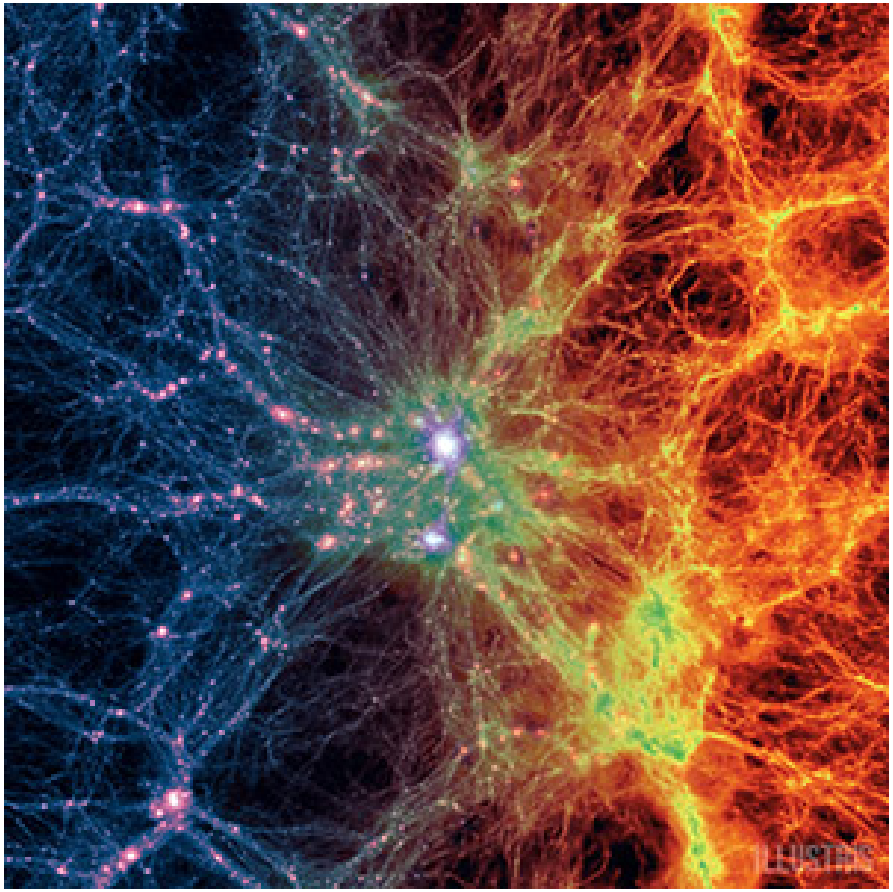
"Allianz der Deutschen Forschungsorganisationen"
represents universities, Max-Planck-, Fraunhofer-, Leibniz- and Helmholtz institutes, DFG and others

Priority initiative Digital Information

Ad-hoc working group "Scientific software" (2016-2017)
→ Working group "Digital tools, software and services" (2018-2022)

Leibniz Association represented via MMS network representatives
Georg Feulner (PIK), JF (WIAS)

Software and Research



Illustris simulation

Research software: central component of scientific work

- Simulation: models + algorithm development
- Generation, processing, analysis and visualization of research data
- Control of devices and experiments.

Software and Good Scientific Practice

- Transparency, traceability, accessibility and reproducibility of research results

FAIR Principles for software:
Findable,**A**ccessible, **I**nteroperable, **R**eusable

Stakeholders

- Management
- Developers
- Users
- Infrastructure facilities

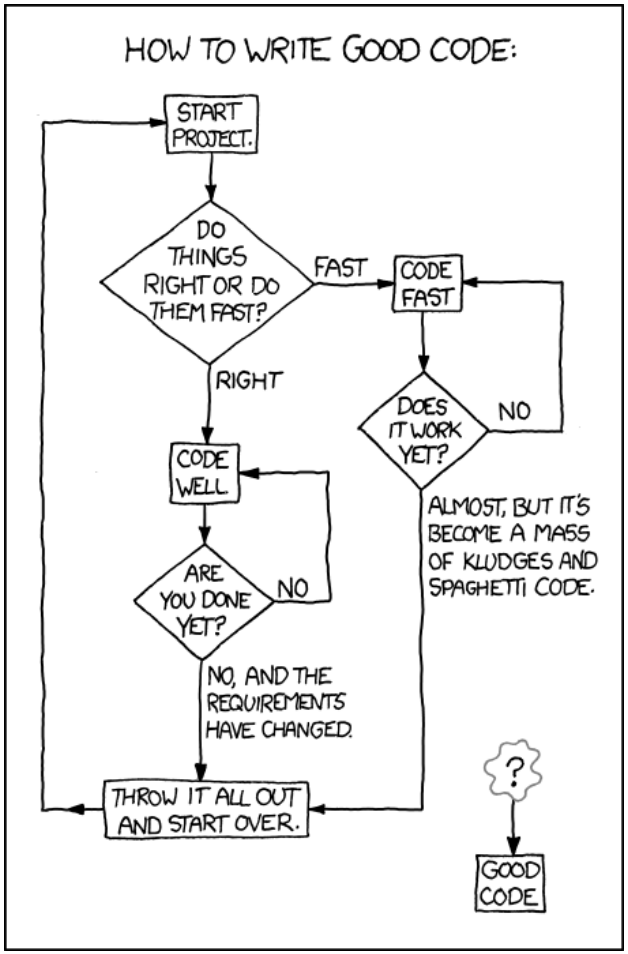
Types of software

- Research codes
- Frameworks used in research, e.g. Matlab, ER
- Services, e.g. Zenodo

Relation to software

- Developing
- Using
- Providing

Developer responsibility



xkcd.com

- Research software engineers
- Academic software developers

Development

- Adhere to best practice standards for coding, packaging, documentation, validation
- Care about interoperability
- Re-use well established components
- Networking, communication, qualification
- Critical self-assessment

User responsibility

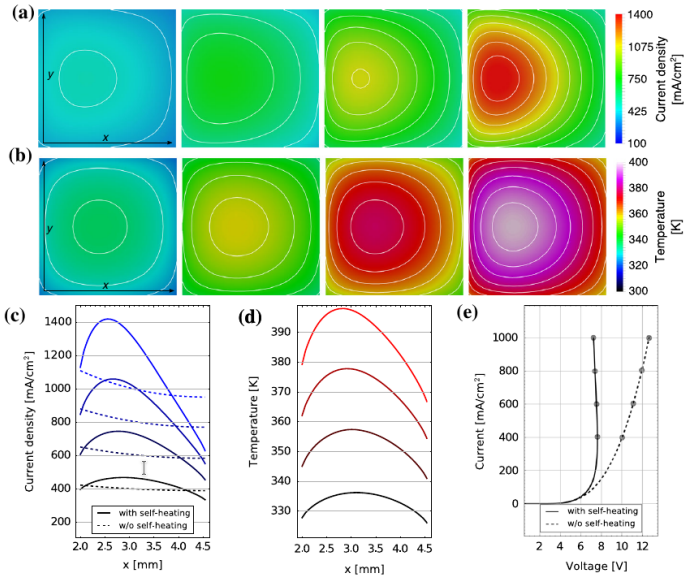


Fig. 2 **a** Simulated inhomogeneous current density and **b** temperature distribution in a horizontal cross section of the OLED stack at values of an average current density (total current over anode devided by anode area) of 400, 600, 800, and 1000 mA/cm^2. **c** Values for current density and temperature at cut along symmetry line of the horizontal cross-section, dashed lines correspond to the isothermal case where self-

DOI 10.1007/s11082-017-1167-4

- Research scientists
- Postdocs, PhD students

Using

- Develop functional understanding of software used
- Care about provenance – cite and document software usage in publications, including version, configuration and parameters
- Consider free and open source alternatives
- Use reviewer power to encourage standards

Infrastructure facilities

Trinity college Dublin © D. Iliff. License: CC-BY-SA 3.0

- IT Departments
- Libraries
- Technology transfer units

Development

- Maintain infrastructure to implement FAIR principles
- Develop know-how on licensing, software publication, long term maintenance
- Offer consulting to developers and users

Providing

- Provide technical infrastructutre supporting collaboration, citation, long term availability
- Create solutions to quantify usage and citations
- Support developers on legal issues, pricing
- Initiate new career paths, e.g. software/data librarians

Management responsibility

- Scientists in leading positions
- Science managers

Development

- Develop and disseminate best practice rules
- Hire and hold qualified developers
- Organize education/qualification of scientific personnel
- Organize recognition, incentives, career paths
- Provide state of the art infrastructure and enforce its use

Using

- Implement rules of Good Scientific Practice with respect to software
- Set citation + documentation rules
- Prefer open source tools and non-commercial services for reasons of economy and data sovereignity
- Prevent use of outdated software & tools

Providing

- Organize sustainable development and maintenance
- Encourage the formation of developer communities
- Clarify licensing models before publication
- Develop business and financing models
- Secure long term availability if provider role is given up
- Organize documentatiom, tutorials etc. for users

Contents

One step forward . . .

DFG Deutsche Forschungsgemeinschaft

Förderung | Geförderte Projekte | DFG im Profil | DFG MAGAZIN

Startseite › Förderung › Ausschreibungen - Informationen für die Wissenschaft › 2016 › Nachhaltigkeit von Forschungssof

Information für die Wissenschaft Nr. 71 | 2. November 2016

Nachhaltigkeit von Forschungssoftware

DFG erbittet Anträge zur Nutzbarmachung und Nachhaltigkeit von Forschungssoftware

In jeder Phase des wissenschaftlichen Arbeitens wird in vielen Disziplinen Forschungssoftware genutzt, zum Beispiel zur Generierung, Verarbeitung, Analyse und Visualisierung von Forschungsdaten. In diesem Sinne sind mit „Forschungssoftware" die eigens zum wissenschaftlichen Erkenntnisgewinn erstellten Software-Anwendungen und Software-Bibliotheken gemeint.

Ein Teil der Forschungssoftware, die aus wissenschaftlichen Projekten hervorgeht, hat enormes Potenzial für eine breite Nutzung, die weit über die ursprüngliche Nutzung in einem einzelnen Forschungsprojekt hinausgeht. In diesem Fall kann ein über den eigenen Standort hinausgehender, (meist) disziplinspezifischer, aber nicht mehr projektspezifischer Bedarf für die nutzerorientierte Weiterentwicklung, Pflege, Nutzbarhaltung, Emulation, Verbreitung und Archivierung dieser Forschungssoftware entstehen.

www.dfg.de

- Significantly increased awareness on this topic on many hierachy levels
- Funding organizations care about best practices
  - Funding projects devoted to software sustainability
  - → Enforcing rules ?
- High level backing + wide dissemination of this document
  ⇒ backing for scientists + research software engineers when rising issues with their management
- Transfer recommendations into institution level rules:
  - Software repository URLS
  - Documention standards
  - Rules for software to be disseminated
  - . . .

. . . but it's a long way

. . . from my own experience

- Research software engineers and dedicated academic software developers try to do the right thing
- Large "grassroot" interest in software issues

But . . .

- Often, scientists (have to) put minimal effort into dealing with software as this is not counted as proper academic output
- Software development as carrier killer ?
- Licensing is learning by doing, takes time
- Tight institutional licensing policies prevent dissemination
- Often little awareness about the real effort behind development and maintenance of scientific software
- Good intentions vs. limited resources

Software complexity vs. transparency

xkcd.com

- Open Source software promises transparency, reusability, productivity
- In reality, systems become more and more complex and difficult to maintain for non-specialists
  - C++: learning curve, late standardization of important features
  - Python: stability of environments (virtualenv anyone ?)
  - Build systems: learning curves
- Software design, choice of the right tools
- Standardized interfaces
- Containers (docker . . .) vs easy to handle source code

. . . it's a journey, we are just starting

Assuming all agree on the issues discussed: what are the next steps ?

- Transfer of recommendations into institutional rules – in many cases this may be a cultural change
- Standardization activities between related open source communities ?
- Exchange forums on concrete experiences in different institutions
  - Past events:
    - MMS Days are on the forefront since 2016 . . .
    - Force 2017 conference https://www.force2017.org/
    - Helmholtz Open Science Workshop "Zugang zu und Nachnutzung von wissenschaftlicher Software" 2016
  - Future events ?
    - Panel discussion this evening
    - (via SSI) CarpentryCon 2018, Dublin
    - (via SSI) Impact of international collaborations in research software, 2018, Manchester
    - (via de-rse.org) Research Software Engineers in the Geosciences, Vienna, part of EGU General Assembly 2018
    - . . .
  - http://www.de-rse.org/