20.04.2021 | Software Sustainability Forum

Nachhaltige Softwareentwicklung in NFDI4Culture

Ulrike Henny-Krahmer
Universität zu Köln

Daniel Jettka
Universität Paderborn

ÜBERSICHT

TEIL I: Einführung

- Was meint Nachhaltigkeit für die Entwicklung von Forschungssoftware?
- Welche Bereiche betrifft nachhaltige Softwareentwicklung in der Forschung?
- Welche Relevanz hat nachhaltige Softwareentwicklung für NFDI4Culture?

TEIL II: Einblicke

- Dokumentation von Software
- Versionierung von Software
- Zitation von Software

TEIL III: Fazit

Einführung

EINFÜHRUNG

Was meint Nachhaltigkeit für die Entwicklung von Forschungssoftware?

Transparenz: Software sollte so entwickelt werden, dass transparent und nachvollziehbar bleibt, wie Forschungsergebnisse entstanden sind.

Effektivität: Ressourcen für die Entwicklung von Forschungssoftware sollten effektiv und nachhaltig eingesetzt werden.

Langfristigkeit: Transparenz und Effektivität sollten möglichst langfristig sichergestellt werden.

allgemein: positive ökologische, soziale oder ökonomische Wirkung

vgl. u.a. Wilkinson, M. D. et al (2016): “The FAIR Guiding Principles for scientific data management and stewardship.” *Scientific Data* 3, https://doi.org/10.1038/sdata.2016.18

Stürmer, Matthias (2017): “Digitale Nachhaltigkeit: Digitale Gemeingüter für die Wissensgesellschaft und Zukunft.” *IT business* 2: 9-11. https://www.parldigi.ch/wp-content/uploads/2017/07/DigitaleNachhaltigkeit_ITbusiness2017.pdf

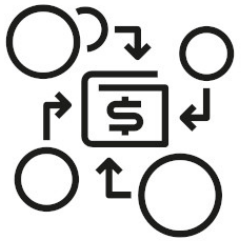
Icons: https://www.parldigi.ch/de/icons-voraussetzungen/

EINFÜHRUNG

Welche Bereiche betrifft nachhaltige Softwareentwicklung in der Forschung?

- technisch:
  - Code-Qualität, Standards (Programmiersprache, APIs, ...)
  - Modularisierung / Containerisierung / Virtualisierung

- organisatorisch:
  - Identifikation / Versionierung / Lizenzierung / Publikation / Zitation / Replizierbarkeit
  - Dokumentation / Kollaboration / Softwaremanagementpläne



- politisch/sozial:
  - Förderung / Finanzierung von Entwicklung und Wartung
  - Anerkennung für EntwicklerInnen
  - Code-Literacy in der Lehre/Ausbildung
  - Openness (Open Science, Open Source, Free Software)

vgl. u.a. Lamprecht, Anna-Lena et al (2020): “Towards FAIR principles for research software.” *Data Science* 3 (1): 37-59. https://doi.org/10/gg66q6

EINFÜHRUNG

Stakeholder: Wer hat Interesse an nachhaltiger Softwareentwicklung in der Forschung?

- Allgemeinheit
- FachwissenschaftlerInnen
- EntwicklerInnen (RSEs)
- ProjektleiterInnen / Forschungsorganisationen
- Förderinstitutionen
- Infrastruktureinheiten (Bibliotheken, Rechenzentren, ...)
- Industrie
- Geopolitische Einheiten

vgl. Anzt, Hartwig et al. (2021): “An environment for sustainable research software in Germany and beyond: current state, open challenges, and call for action [version 2; peer review: 2 approved].” F1000Research 9:295. https://doi.org/10.12688/f1000research.23224.2

Katerbow, Matthias und Georg Feulner (2018): *Handreichung zum Umgang mit Forschungssoftware.* Hrsg. von der Schwerpunktinitiative Digital Information der Allianz der deutschen Wissenschaftsorganisationen. https://doi.org/10.5281/zenodo.1172970

EINFÜHRUNG

Welche Relevanz hat nachhaltige Softwareentwicklung für NFDI4Culture?

Status quo:

- viele Aspekte von Software-Nachhaltigkeit sind allgemein
- dennoch:
  - besondere Voraussetzungen für die Entwicklung von Forschungssoftware in den Geisteswissenschaften
  - besondere Arten von Software in NFDI4Culture
  - besondere Bedeutung von nachhaltiger Software in NFDI4Culture

EINFÜHRUNG

Welche Relevanz hat nachhaltige Softwareentwicklung für NFDI4Culture?

Beispiel: meico MEI Converter https://github.com/cemfi/meico

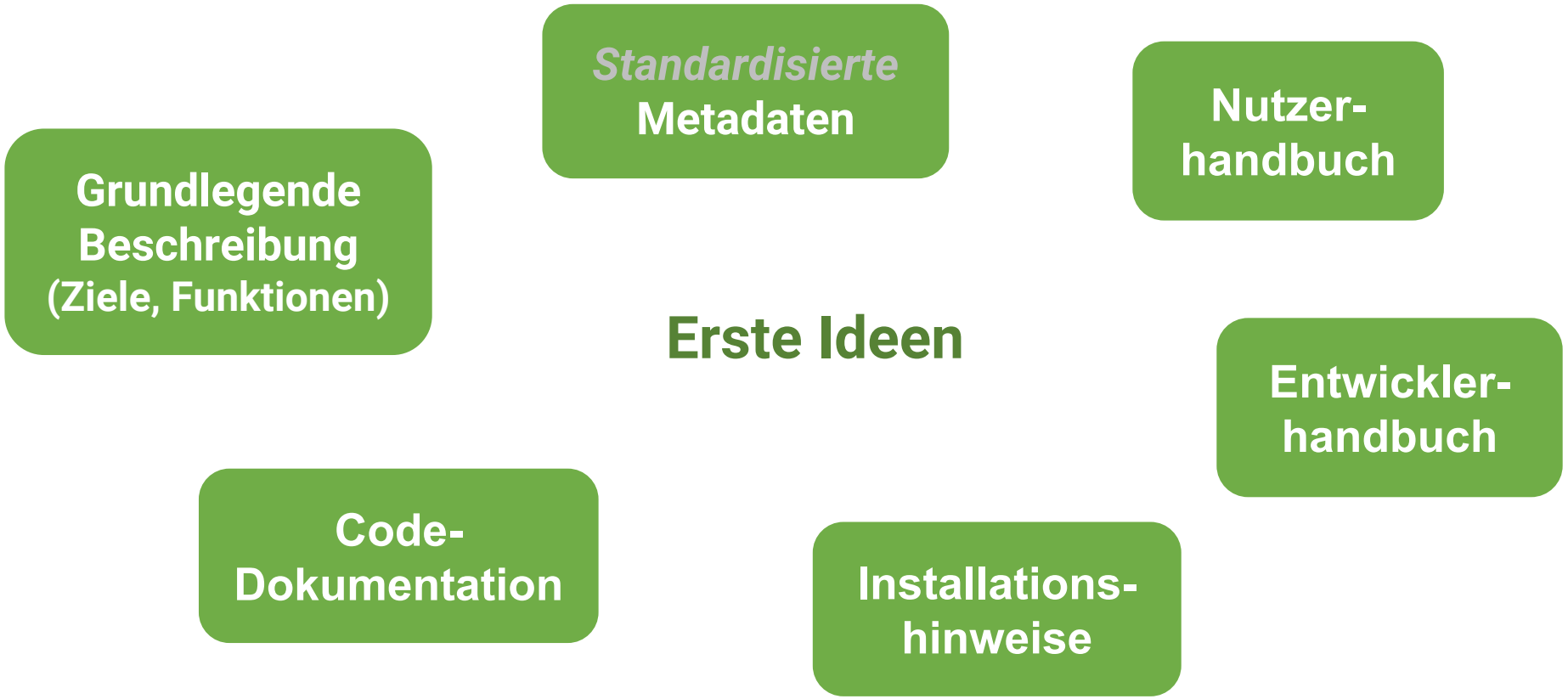
- Ein Converter-Framework für MEI-Dateien (*Music Encoding Initiative*), mit Unterstützung für MEI, MSM, MPM, MIDI, WAV, MP3, chroma, XSLT, geschrieben in Java
- Lizenz: GPL-3.0; Versionierung: v0.8.22
- Identifikation? Zitation? - Referenzpublikation: Bernd, Axel, Simon Waloschek und Aristoteles Hadjakos (2018): "Meico: A Converter Framework for Bridging the Gap between Digital Music Editions and its Applications." In: *AM'18: Proceedings of the Audio Mostly 2018 on Sound in Immersion and Emotion.* https://doi.org/10.1145/3243274.3243282
- Dokumentation: README auf GitHub, Referenzpublikation, Javadoc (Installation, Nutzung, Deployment, Dependencies, Code)
- Software management / Finanzierung? ZenMEM-Projekt (BMBF 2015-2019); Fritz Thyssen-Stiftung (2019-2022)

Einblicke

Dokumentation, Versionierung, Zitation

DOKUMENTATION

Annahme: Umfangreiche Dokumentation erhöht die Wahrscheinlichkeit langfristiger Nutzbarkeit

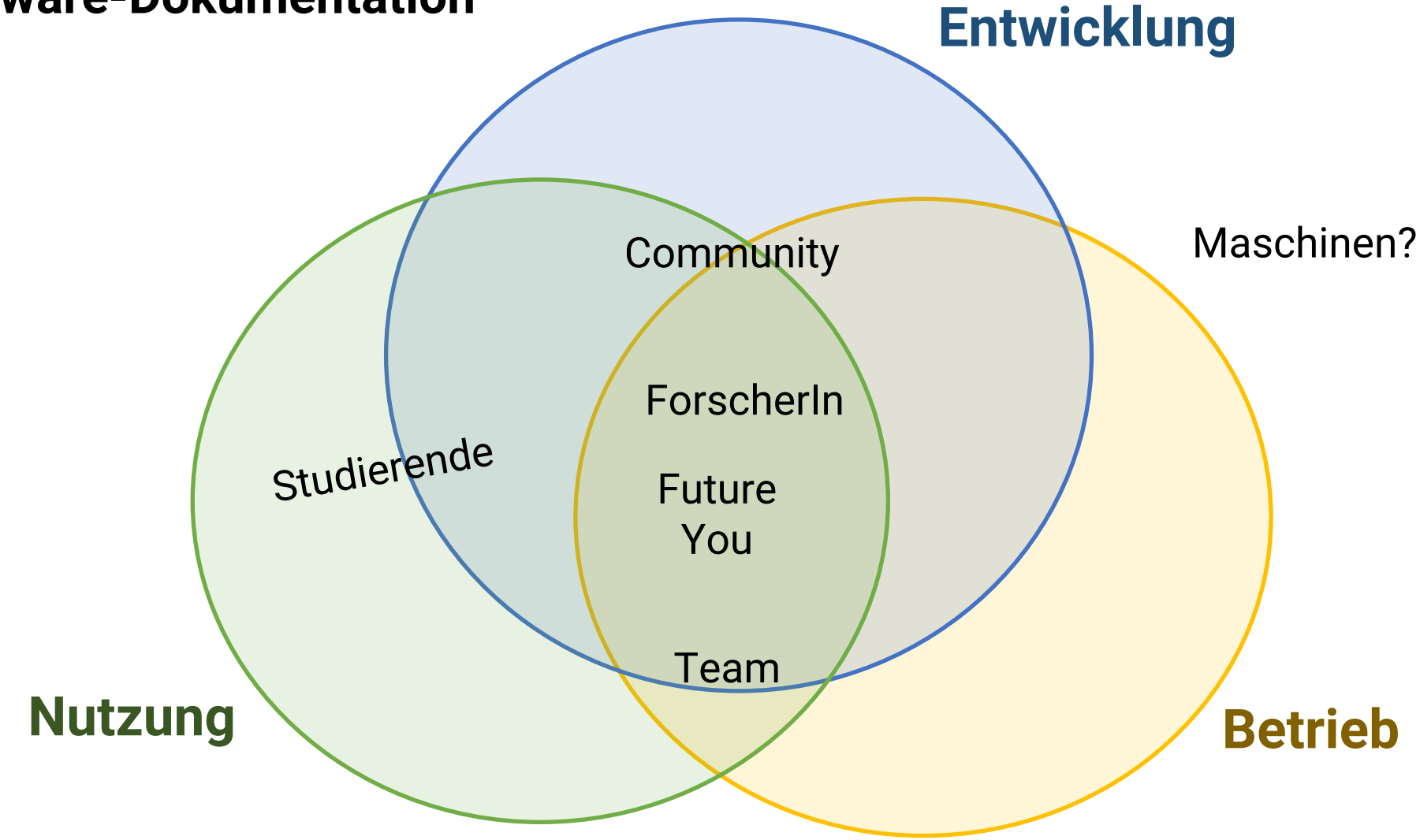


Bestandteile von Software-Dokumentation

Konzeptuelle Dokumentation	Referentielle Dokumentation	Praxis-basierte Dokumentation
• Requirements engineering • Systemmodellierung • Architektur • Objektorientiertes Design	• Code-Kommentierung • API-Dokumentation • README • User guide • Glossar	• How-tos • Getting started • Installationshinweise

vgl. Druskat et al., 2019. https://www.software.ac.uk/blog/2019-06-21-what-are-best-practices-research-software-documentation?mc_cid=7be707599b&mc_eid=8fe099f2f0

Beteiligte an Software-Dokumentation



VERSIONIERUNG

Annahme: Software muss sich verändern, um langfristig nutzbar sein

Versionskontrolle
Veröffentlichung von Software Repositories zur Verbesserung der Replizierbarkeit

Verbindung mit Issue/Bug Tracker - Projekt-/Entwicklungsmanagement

Problem: “*computational science software repositories’ lifespan has a distribution with a median lifespan of 15 days. A third of these repositories are live for less than 1 day*” (Hasselbring et al. 2019, https://arxiv.org/abs/1908.05986)

VERSIONIERUNG

Annahme: Software muss sich verändern, um langfristig nutzbar sein

Nachvollziehbares, konsistentes Versionierungsschema (z.B. Semantic Versioning)

Anwendung und Bereitstellung Persistenter Identifikatoren und Metadaten für jede veröffentlichte Version (FAIR - F1)

Problem: Versionierung von Datendiensten? Direkt miteinander verbundene Softwareversionen und Datenversionen, oder gemeinsam versionieren?

Wie Software zitierfähig machen...

Vorbereitung

- Anwendung einer passenden Lizenz
- Angemessene Beschreibung (mit passenden Metadaten/-formaten)
- Konsistente Versionsnummer
- Festlegung und Anerkennung von AutorInnen (und Berücksichtigung in Metadaten)
- Persistenter Identifizierer
- Empfehlung einer zitierfähigen Dokumentation zur Software

(vgl. Chue Hong et al., 2019. Software Citation Checklist for Developers. https://doi.org/10.5281/zenodo.3482769)

Anwendung

Wie Software korrekt zitieren...

- Identifizierung der Software (inkl. der eigenen), die aufgrund ihres Beitrags zitiert werden sollte
- Falls vorhanden, empfohlene Zitation der Software verwenden. Software (zusätzlich) direkt zitieren
- Sollte keine Empfehlung existieren, Zitation so vollständig wie möglich gestalten - AutorInnen, Zeit, Titel, Version, Referenz auf zugreifbare Instanz der Software (PID falls möglich)
- Software im eigenen akademischen Werk angemessen referenzieren, in Übereinstimmung mit Zitationsrichtlinien

(vgl. Chue Hong et al., 2019. Software Citation Checklist for Authors. https://doi.org/10.5281/zenodo.3479199)

ÜBERSICHT: TEILBEREICHE

Fazit

FAZIT

Nachhaltigkeit in der Softwareentwicklung

muss für NFDI4Culture erst etabliert werden!

z.B. durch:

Beratungsagentur	Software Registry	Rezension
Informieren	Finden	Sichten
Empfehlen	Erfassen	Einordnen
Beraten	Zertifizieren	Bewerten
		(“Construction Kit”)

OFFENE FRAGEN UND ZIELE FÜR DAS FORUM

- Gibt es **weitere Bereiche**, die für die nachhaltige Entwicklung von Forschungssoftware relevant sind und die wir bisher übersehen haben?

- Um welche Aspekte der nachhaltigen Softwareentwicklung in NFDI4Culture sollten wir uns **prioritär** kümmern?
  - z. B.: Fokus auf Datendienste?
  - Anwendung/Auslegung allgemeiner Kriterien für 4Culture?

- In welcher Form könnten Ergebnisse des Forums **publiziert** werden?

- ...

KONTAKT

Ulrike Henny-Krahmer
ulrike.henny@uni-koeln.de
https://orcid.org/0000-0003-2852-065X

Daniel Jettka
daniel.jettka@uni-paderborn.de
https://orcid.org/0000-0002-2375-2227

#forum_sustainable_software_development
(Rocket.Chat)

info@nfdi4culture.de
https://nfdi4culture.de

Gefördert durch die Deutsche Forschungsgemeinschaft (DFG) - 441958017