AgEcon SEARCH
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search
http://ageconsearch.umn.edu
aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

Leibniz-Institut für
Agrarentwicklung
in Transformations-
ökonomien (IAMO)
www.iamo.de

IAMO
Policy Brief

Nº 42

Ausgabe Nr. 42
August 2021

Forschungssoftware – Nachhaltige Entwicklung und Unterstützung

Franziska Appel
Axel Loewe

Forschungssoftware ist zu einem zentralen Gut der akademischen Forschung geworden. Sie ist oft das Rückgrat bestehender Forschungsmethoden, ermöglicht neue Forschungsmethoden, implementiert und bettet Forschungswissen ein und stellt selbst ein wesentliches Forschungsprodukt dar. Forschungssoftware muss nachhaltig sein, um bestehende Wissenschaft zu verstehen, zu replizieren, zu reproduzieren und darauf aufzubauen, um neue Forschung effektiv durchzuführen. Daher ist eine Änderung der Art und Weise erforderlich, wie die Entwicklung und Wartung von Forschungssoftware derzeit motiviert, incentiviert, finanziert, strukturell und infrastrukturell unterstützt und rechtlich behandelt wird. Geschieht dies nicht, gefährdet dies die Qualität und Validität der Forschung. Vor diesem Hintergrund trafen sich im November 2019 mehr als 50 Wissenschaftler:innen aus verschiedenen Institutionen zu einem DFG-Rundgespräch am Robert-Koch-Institut, um Nachhaltigkeitsaspekte von Forschungssoftware zu diskutieren und auf die Notwendigkeit von Förderinstrumenten für die nachhaltige Entwicklung und Bereitstellung von Forschungssoftware aufmerksam zu machen. Das anschließend veröffentlichte Positionspapier (Anzt et al., 2021) wendet sich gezielt an politische und akademische Entscheidungsträger, um das Bewusstsein für die Bedeutung und den Bedarf einer nachhaltigen Forschungssoftwarepraxis zu schärfen. Insbesondere empfiehlt es Strategien und Maßnahmen, um ein Umfeld für nachhaltige Forschungssoftware zu schaffen.

Das Rückgrat der Forschung des 21. Jahrhunderts: Aktueller Stand und offene Herausforderungen

Die computergestützte Analyse großer Datensätze, computergestützte Simulationen und Softwaretechnologie im Allgemeinen spielen eine zentrale Rolle für praktisch alle wissenschaftlichen Durchbrüche des 21. Jahrhunderts. Zur gleichen Zeit stützen sich Forschende oft auf Software, die hausintern oder von Freiwilligen entwickelt wurde. Viele dieser Softwareumgebungen können nicht aufrechterhalten werden – oft, weil die Forschungssoftware kein benanntes und primäres Ergebnis eines Forschungsprojekts war und daher im Prototyp-Zustand verblieb oder wegen fehlender Anreize und Ressourcen die Software nach dem Ende der Projektfinanzierung weiter zu pflegen. Dies führt zu einem höchst ineffizienten System, in dem jedes Jahr Millionen von Codezeilen erzeugt werden, die nach dem Ende eines Projekts oder dem Auslaufen der Stelle des:der Entwickler:in nicht wiederverwendet werden.

In Deutschland erkennen Fördereinrichtungen zunehmend die Bedeutung und den Wert nachhaltiger Forschungssoftware und damit verbundener Infrastrukturen an. Die Deutsche Forschungsgemeinschaft (DFG), die größte Fördereinrichtung für Grundlagenforschung in Deutschland, hat zum Beispiel Ausschreibungen für nachhaltige Forschungssoftwareentwicklung und Qualitätsmanagement im Bereich Forschungssoftware eröffnet. Als weiteres Beispiel führt der Kodex „Leitlinien zur Sicherung guter wissenschaftlicher Praxis“ der DFG aus 2019

Software nun explizit neben anderen Forschungsergebnissen und -daten auf.

Trotz dieser positiven Entwicklungen fehlen leider immer noch Leitlinien und Konzepte für eine nachhaltige Entwicklung von Forschungssoftware in Deutschland. Langfristige Förderstrategien fehlen ebenfalls. Die Schlussfolgerungen und Empfehlungen des DFG-Rundgesprächs sind im Folgenden in Anlehnung an Anzt et al. 2021 zusammengefasst.

Wie wird entschieden, welche Software erhalten werden sollte?

Die dauerhafte Förderung aller bestehenden Softwareprojekte ist nicht nur unmöglich, sondern würde auch das Risiko einer starken Zersplitterung der wissenschaftlichen Fach-Communities mit sich bringen und wäre letztendlich kontraproduktiv für die Effizienz der Forschungsgemeinschaft. Daher ist es wichtig, sich auf eine Liste transparenter Kriterien zu einigen, die ein Softwareprodukt für eine dauerhafte Förderung qualifizieren. Eine solche Liste muss allgemein genug sein, um auf Forschungssoftware aus verschiedenen Forschungsdisziplinen angewandt werden zu können und gleichzeitig die Unterschiede zwischen den Bereichen respektieren.

Hier konzentrieren wir uns auf die obligatorischen Kriterien, die domänenübergreifend erfüllt sein sollten. Weitere zusätzliche wünschenswerte Kriterien, die je nach Anwendungsfall und domänenspezifischen Anforderungen an die Softwareentwicklung in unterschiedlichem Maße umgesetzt werden können, werden in Anzt et al. 2021 genannt.

1) Nutzung und Bedeutung

a) Verwendung über eine einzelne Forschungsgruppe hinaus
b) Mindestens eine begutachtete wissenschaftliche Veröffentlichung basierend auf der Software
c) Marktanalyse (Software ist wichtig für eine Nutzerbasis relevanter Größe und ist entweder einzigartig oder einer der Hauptakteure in einem Bereich mit mehreren bestehenden Lösungen)
d) Geeignete Schulungsunterlagen werden zur Verfügung gestellt

2) Transparenz und Qualität

a) Öffentliche Verfügbarkeit des Quellcodes
b) Versionskontrollsystem
c) Lizenziert
d) Abhängigkeiten von Softwarebibliotheken und Technologien sind definiert

3) Reifegrad

a) Software-Management-Plan
b) Website mit klar definierter Anlaufstelle
c) Gruppe von Entwickler:innen

Wer ist dafür zuständig, Forschungssoftware zu erhalten?

Sowohl Förderorganisationen als auch Forschungseinrichtungen spielen wichtige Rollen für nachhaltige Forschungssoftware. Förderer sollten von den Antragstellenden Überlegungen verlangen, wie die in einem Projekt entwickelte Software über das Ende des geförderten Projekts hinaus aufrechterhalten werden kann. Dies erfordert auch längerfristige Förderbausteine für Softwarewartung, -pflege und Nutzerunterstützung, die von Anfang an beantragt werden können. Eine Weiterverfolgung dieser Pläne während und nach der Projektlaufzeit, d. h. ein dedizierter Software-Management-Plan, ist entscheidend.

Entscheidungsträger:innen sollten die Definition der Forschungswirkung über traditionelle wissenschaftliche Publikationen hinaus erweitern und auch andere wirkungsvolle Ergebnisse (wie Software und Daten) einbeziehen, zum Beispiel bei der Rekrutierung von akademischen Positionen. Diese Erweiterung der Basis für die Bewertung von Forschung und Forscher:innen wird explizit in der „San Francisco Vereinbarung über die Forschungsbewertung" (SF DORA) genannt, die viele deutsche und internationale Förderer und Forschungseinrichtungen jüngst unterzeichnet haben. Darüber hinaus sollten Forschungseinrichtungen attraktive langfristige Karrieremöglichkeiten im akademischen Umfeld schaffen, um den Gewinn von engagierten Forschungssoftwareentwickler:innen (Research Software Engineers; RSE) als Schlüsselkonzept für professionelle Forschungssoftware-Praktiken zu nutzen.

Wie kann Forschungssoftware nachhaltig finanziert werden?

Da die Finanzierung ein entscheidender Faktor für die Nachhaltigkeit von Forschungssoftware ist, müssen die bestehenden projektorientierten Förderinstrumente auf lokaler, nationaler und internationaler Ebene durch Förderinstrumente ergänzt werden, die speziell auf die nachhaltige Pflege und Bereitstellung von Forschungssoftware ausgerichtet sind. Zum Beispiel sollten Softwareprojekte, die die oben genannten Nachhaltigkeitskriterien erfüllen, Zugang zu dauerhafter Förderung haben, solange sie den Standards gerecht werden und weiterhin zentraler Bestandteil der Forschungslandschaft sind.

Rechenzentren für die Forschung müssen zweckgebundene Mittel für die Unterstützung nachhaltiger Softwareentwicklung erhalten. Diese Mittel sind notwendig, um Continuous-Integration-Systeme, ein Hardwareportfolio für die Entwicklung, das Testen und das Benchmarking von Software sowie Personal für die Schulung von Domänenforscher:innen im Softwaredesign und der richtigen Nutzung der Dienste bereitzustellen.

Die Erstellung und Pflege von Schulungsmaterialien für die allgemeine Ausbildung im Bereich

Software-Engineering in der Forschung sowie software-spezifische Dokumentation und die Erstellung von Tutorials müssen in den Finanzierungsmöglichkeiten berücksichtigt werden. Dies kann entweder dadurch geschehen, dass Module der Forschungs- oder Softwareförderung dem Support und der Erstellung von Schulungsmaterial gewidmet werden oder durch die Schaffung von Förderprogrammen, die sich auf die interdisziplinäre Softwareentwicklungsausbildung konzentrieren.

Welche Infrastruktur wird für nachhaltige Forschungssoftware benötigt?

Projektmanagement Tools

Die Bereitstellung von Werkzeugen zur verteilten Softwareentwicklung und -verwaltung wird als zentrale Forschungsinfrastruktur empfohlen, um eine institutsübergreifende Zusammenarbeit zu ermöglichen. Ein wichtiger Aspekt in diesem Zusammenhang ist die Katalogisierung von Forschungssoftware, um doppelten Entwicklungsaufwand zu reduzieren. Dies kann effizient realisiert werden, indem die Registrierung aller Forschungssoftware mit einem eindeutigen Identifikator gefördert wird und ein Werkzeug entwickelt wird, das die Erkundung der Forschungssoftwarelandschaft ermöglicht. Mitwirkende an Forschungssoftware sollten ihre ORCID iD nutzen, um eindeutig identifizierbar und referenzierbar zu sein.

Entwicklerschulung, Motivation und Wissensaustausch

Workshops und Seminare, die einen einfachen Zugang zu praktischem Training zu softwarebezogenen Fragen bieten, sollten gefördert und unterstützt werden. Eine aktuelle Initiative in diese Richtung sind die Helmholtz-HIFIS-Events. Die langfristige Lösung zur Erlangung ausreichender Softwareentwicklungskompetenzen sollte jedoch eine Ausbildung sein, die früh in Hochschulcurricula einbezogen wird, idealerweise bereits auf Bachelor-Ebene.

Vernetzung und Community-Building können zu einer dauerhaften Etablierung von nachhaltiger Forschungssoftware führen. Software Carpentry und ähnliche Initiativen wie die Gründung der Data Science Academy HIDA in der Helmholtz-Gemeinschaft helfen, Wissen auszutauschen und zu verbreiten. Lokale RSE-Gruppen und (inter-)nationale Konferenzen werden die Gemeinschaftsbildung weiter fördern. Forschungseinrichtungen sowie Förderorganisationen sollten die Bildung solcher Gruppen aktiv fördern und ideell sowie finanziell unterstützen.

Die Schaffung eines nationalen Software-Nachhaltigkeitsinstituts (mit finanzierten Stellen zum Aufbau von Services und zur Erstellung von Schulungsmaterial) ähnlich dem britischen Software Sustainability Institute (SSI), das als nationaler Ansprechpartner für alle Aspekte im Zusammenhang mit Forschungssoftware dient, wird dringend empfohlen.

Forschungssoftware finden, veröffentlichen und archivieren

Ordnungsgemäße Software-Publikation und Möglichkeiten für die Community, bestehende Software-Lösungen für ein bestimmtes Problem zu finden, sind eine Voraussetzung, um Synergien optimal zu nutzen und redundante Entwicklungen zu vermeiden. Das Auffinden erfordert die Veröffentlichung an einem weltweit zugänglichen Ort mit entsprechenden Metadaten, z. B. Citation File Format (CFF). Umfassende Metadaten (z. B. Mitwirkende, Kontakt, Schlagworte, verlinkte Publikationen, etc.) und Publikationsplattformen können persistente Zitierbarkeit ermöglichen, was wiederum der Forschungsevaluation zugutekommt. Insbesondere in interdisziplinären Umgebungen empfiehlt sich die Erstellung eines Meta-Index, der wichtige (disziplinäre) Software-Indizes abdeckt, um das Auffinden relevanter Software-Standorte zu erleichtern.

GitLab oder GitHub (gängige kollaborative Arbeitsumgebungen) in Verbindung mit Repositorien wie Zenodo sind geeignete Publikationsplattformen, da letzteres DOIs prägt, Versionierung erlaubt und für den langfristigen Zugriff öffentlich gefördert wird. Metadaten und Referenzen sind essentiell für Tools wie PIDgraph, DataCite.org und CrossRef, die persistente Identifikatoren (PIDs) wie DOIs verwenden. Für Deutschland ist davon auszugehen, dass die Nationale Forschungsdateninfrastruktur (NFDI) Anstrengungen unternimmt, an zentraler Stelle Discovery-Plattformen zu schaffen oder zu unterstützen, die das Auffinden von Informationen erleichtern. Leider spielt Forschungssoftware in der NFDI noch nicht die zentrale Rolle, die sie spielen sollte.

Die Veröffentlichung einer bestimmten Softwareversion als Referenz in Forschungsartikeln erfordert einfache Möglichkeiten, Forschungssoftware langfristig zu archivieren. Die Veröffentlichung von Software, ihren Abhängigkeiten und ihrer Umgebung in Containern kann die Auswertung und weitere Wiederverwendung erleichtern. Softwarekonservierung zielt darauf ab, die Lebensdauer von Software, die nicht mehr aktiv gepflegt wird, zu verlängern. Eine Lösung, um die Software in einem lauffähigen Zustand zu halten, indem ihr Kontext und ihre Laufzeitumgebung erhalten bleiben, ist die Verwendung von Containern wie Docker. Eine weitere Gefahr ist der Verlust von Projektrepositorien auf globalen Plattformen wie GitHub oder BitBucket. Hier ernten globale Plattformen wie Software Heritage diese Repositorien und verhindern den Verlust durch Langzeitarchivierung.

Rechtliche Aspekte

Eine wesentliche Frage im Hinblick auf Nachhaltigkeit von Forschung betrifft die freie und offene Verbreitung von Forschungssoftware und anderen Forschungsergebnissen. Es sollte diskutiert werden, ob Forschungsförderorganisationen wie die DFG erwarten sollten, dass alle geförderte Software unter solchen Lizenzen nach dem Paradigma „public money,

public code" veröffentlicht wird. Die Verwendung einer FSF- oder OSI-anerkannten FLOSS-Lizenz würde beispielsweise ein wirklich freies Modell ermöglichen und auch rechtliche Probleme reduzieren. Wenn Lizenzen wie Apache oder MIT verwendet werden, können die Forschungseinrichtungen die Software später, wenn gewünscht, immer noch kommerzialisieren. Eine solche Open-Source-Lizenzierung ist auch für Start-ups von Vorteil, die beabsichtigen, professionelle Dienstleistungen für die Software anzubieten.

Gerade vor dem Hintergrund des neuen DFG-Kodex ist die Installation einer Forschungssoftware-Task-Force für alle (deutschen) Forschungseinrichtungen/-gemeinschaften wichtig. Als Teil ihrer Aufgaben sollte eine solche Gruppe eine lokale Rechtsberatungsstelle und Bildungsangebote organisieren sowie (falls noch nicht implementiert) die Softwarerichtlinie der Einrichtung entwickeln. Mit Hilfe von Onboarding-Prozessen, die von der Forschungssoftware-Task-Force durchgeführt werden, sollten RSEs zum Beispiel in der Lage sein, die Klärung der Rechte von Anfang an so einfach wie möglich zu halten. Dies hilft zu vermeiden, dass – aus Unsicherheit und Angst, einen rechtlichen Fehler zu machen – einige Forschungsgruppen am Ende gar keine Lizenz wählen, was die Wiederverwendung der Software massiv behindert. Wir schlagen vor, dass die lokalen Task-Forces ein Netzwerk mit den anderen forschenden Organisationen aufbauen, um Ideen auszutauschen, aber auch um eine Bottom-up-Strategie zur Organisation von RSE-Standards für Deutschland und darüber hinaus zu generieren und möglicherweise der Ursprung des oben erwähnten Software-Nachhaltigkeitsinstituts zu sein.

Fazit

- Implementierung von Förderprogrammen und Anreizen zur nachhaltigen Unterstützung der Entwicklung und Pflege von Forschungssoftware auf der Basis klarer und transparenter Kriterien.
- Attraktive Karrierewege für Entwickler:innen von Forschungssoftware.
- Bereitstellung einer Infrastruktur, die nachhaltige Softwareentwicklung ermöglicht, einschließlich Plattformen für Software-Entdeckung, Zusammenarbeit, Continuous-Integration, Testen, und langfristige Bewahrung.
- Etablierung einer landesweiten Institution ähnlich dem Software Sustainability Institute (SSI), die Projektberatung und Code-Review-Dienstleistungen sowie Schulungen zur nachhaltigen Softwareentwicklung anbietet.
- Nachhaltige Softwareentwicklungspraktiken als integraler Bestandteil von Lehrplänen an Hochschulen.
- Entscheidung der Forschungsförderer, ob das „public money, public code"-Paradigma angewendet werden soll, welches verlangt, dass alle öffentlich geförderte Software unter einer Free/Libre Open Source Software (FLOSS)-Lizenz öffentlich verfügbar sein muss.

Weiterführende Informationen

Literatur

Anzt, H.; Bach, F.; Druskat, S.; Löffler, F.; Loewe, A.; Renard, B. Y.; Seemann, G.; Struck, A.; Achhammer, E.; Aggarwal, P.; Appel, F.; Bader, M.; Brusch, L.; Busse, C.; Chourdakis, G.; Dabrowski, P. W.; Ebert, P.; Flemisch, B.; Friedl, S.; Fritzsch, B.; Funk, M. D.; Gast, V.; Goth, F.; Grad, J.-N.; Hegewald, J.; Hermann, S.; Hohmann, F.; Janosch, S.; Kutra, D.; Linxweiler, J.; Muth, T.; Peters-Kottig, W.; Rack, F.; Raters, F. H. C.; Rave, S.; Reina, G.; Reißig, M.; Ropinski, T.; Schaarschmidt, J.; Seibold, H.; Thiele, J. P.; Uekermann, B.; Unger, S.; Weeber, R. (2021): An environment for sustainable research software in Germany and beyond: Current state, open challenges, and call for action. F1000Research 9;295. URL: https://doi.org/10.12688/f1000research.23224.2 (Stand: 30. Juli 2021).

de-RSE e.V. – Gesellschaft für Forschungssoftware (Society for Research Software): Research Software Engineers (RSEs) – verantwortlich für wissenschaftliche Software. URL: https://de-rse.org/de/ (Stand: 30. Juli 2021).

Kontakt

Dr.-Ing. Axel Loewe
axel.loewe@kit.edu
Tel.: +49 721 608-42790

Karlsruhe Institute
of Technology (KIT)
Kaiserstraße 12
76131 Karlsruhe
www.ibt.kit.edu

Dr. Franziska Appel
appel@iamo.de
Tel.: +49 345 2928-117

Leibniz-Institut für
Agrarentwicklung in
Transformations-
ökonomien (IAMO)
Theodor-Lieser-Straße 2
06120 Halle (Saale)
www.iamo.de

Printausgabe: ISSN 23 63-5770
ISBN 978-3-95992-123-7

Online-Ausgabe: ISSN 23 63-5789
ISBN 978-3-95992-124-4

Das IAMO wurde im Rahmen der DFG-Ausschreibung für nachhaltige Forschungssoftwareentwicklung mit dem Projekt „AgriPoliS2020 – Sustainability of the Research Software AgriPoliS“ gefördert.

IAMO

Leibniz-Institut für Agrarentwicklung in Transformationsökonomien (IAMO)

Das Leibniz-Institut für Agrarentwicklung in Transformationsökonomien (IAMO) widmet sich der Analyse von wirtschaftlichen, sozialen und politischen Veränderungsprozessen in der Agrar- und Ernährungswirtschaft sowie in den ländlichen Räumen. Sein Untersuchungsgebiet erstreckt sich von der sich erweiternden EU über die Transformationsregionen Mittel-, Ost- und Südosteuropas bis nach Zentral- und Ostasien. Das IAMO leistet dabei einen Beitrag zum besseren Verständnis des institutionellen, strukturellen und technologischen Wandels. Darüber hinaus untersucht es die daraus resultierenden Auswirkungen auf den Agrar- und Ernährungssektor sowie die Lebensumstände der ländlichen Bevölkerung. Für deren Bewältigung werden Strategien und Optionen für Unternehmen, Agrarmärkte und Politik abgeleitet und analysiert. Seit seiner Gründung im Jahr 1994 gehört das IAMO als außeruniversitäre Forschungseinrichtung der Leibniz-Gemeinschaft an.

Leibniz
Gemeinschaft