

Forschungssoftware

Nachhaltige Entwicklung und Bereitstellung

| HARTWIG ANZT | AXEL LOEWE | In der Wissenschaft

werden viele digitale Daten erzeugt. Ihre Analyse und Nachnutzung setzt voraus, dass die Daten über eine Forschungssoftware aufbewahrt und aufbereitet werden. Wie nachhaltig geschieht das? Wo liegen beim Thema „Forschungssoftware“ in Deutschland die Probleme?

Wie wichtig ist Forschungssoftware für die moderne Wissenschaft? Um diese Frage zu beantworten, wurden bereits 2014 im Vereinigten Königreich und 2017 in den USA großangelegte Umfragen unter promovierten Wissenschaftlerinnen und Wissenschaftlern (Post-Docs) durchgeführt. Die Umfrageergebnisse spiegeln auf unmissverständliche Weise die Relevanz von Forschungssoftware für den wissenschaftlichen Erfolg wider: Neun von zehn Befragten geben an, dass sie in ihrer Tätigkeit Forschungssoftware benutzen (95 Prozent UK, 92 Prozent USA). Des Weiteren gaben zwei Drittel der Befragten (67 Prozent UK, 63 Prozent US) an, in solch einem Grad von Software abhängig zu sein, dass die Forschung ohne diese nicht weitergeführt werden kön-

te. Nur ein geringer Anteil der Wissenschaftlerinnen und Wissenschaftler könnte die Forschung auch ohne Software durchführen. Diese Zahlen sollten aufhorchen lassen: Es gibt wenig Anzeichen, dass die Situation in Deutschland (wo keine so groß angelegte Untersuchung durchgeführt wurde) anders ist; zudem kann man davon ausgehen,

»Nur ein geringer Anteil der Wissenschaftlerinnen und Wissenschaftler könnte die Forschung auch ohne Software durchführen.«

dass durch die verstärkte Verwendung von Methoden der künstlichen Intelligenz Software eine noch wichtigere Rolle im wissenschaftlichen Alltag eingenommen hat.

Mangelnde akademische Anerkennung

Neben der Abhängigkeit von Forschungssoftware haben die Umfragen aber auch gezeigt, dass ein signifikanter Anteil der Wissenschaftlerinnen und Wissenschaftler Software selbst entwickelt – in den meisten Fällen als in-house Softwarelösung, um die eigene Forschung voranzutreiben. Das Spektrum erstreckt sich von kleinen Skripten zur Datenauswertung bis hin zum komplexen Simulationsworkflow. Da sich Forschungssoftware nicht in den traditionellen Metriken wie Hirsch-Index oder Drittmitteleinwerbungen widerspiegelt, gilt bei der Entwicklung meist das Credo „Aufwand minimieren, Output maximieren“, was häufig zu Lasten

der Nachhaltigkeit geht. Die Folge ist, dass die meisten wissenschaftlichen Softwareprojekte nur eine sehr beschränkte Lebensdauer haben, oftmals limitiert durch das Ende eines Forschungsprojekts oder den Weggang der Hauptentwicklerin oder des Hauptentwicklers, z.B. nach erfolgreicher Promotion. Das führt dazu, dass viele Wissenschaftlerinnen und Wissenschaftler Software entwickeln, die bereits zuvor in wenig nachhaltiger Form und meist nicht öffentlich zugänglich existierte. Wenngleich die Entwicklung von Forschungssoftware in sich einen pädagogischen Nutzen haben mag, so ist diese Strategie global gesehen doch höchst ineffizient, da sie signifikant Kapazitäten bindet, die nicht zur Realisierung des eigentlichen Projektziels zur Verfügung stehen.

Zwar gibt es neben den Inhouse-Softwareprojekten auch etablierte Forschungssoftware, die von einer wissenschaftlichen Community gemeinsam entwickelt wird, aber auch diese Projekte leiden darunter, dass die langfristige Finanzierung der Entwicklung und Bereitstellung nur in seltenen Fällen gesichert ist. Der Kern des Problems besteht darin, dass wissenschaftliche Software und deren Entwicklung im Gegensatz zu Großgeräten nach wie vor nicht als „First-Class Citizen“ im Forschungsportfolio anerkannt wird.

Nachhaltige Förderung nötig

Andere Wissenschaftsstandorte sind bereits einen Schritt weiter und haben die Bedeutung von Forschungssoftware und die Notwendigkeit der nachhaltigen Förderung im Sinne von Software als Infrastruktur erkannt. Das vielleicht beste Beispiel ist das noch von Barack

AUTOREN



Dr. Hartwig Anzt ist Leiter der Helmholtz-Young-Investigator-Gruppe am Steinbuch Center for Computing am Karlsruher Institut für Technologie (KIT).



Dr.-Ing. Axel Loewe leitet die Arbeitsgruppe Computermodelle des Herzens und ist Mitglied im Young Investigator Network (YIN) am Karlsruher Institut für Technologie (KIT).



Foto: mauritius-images

Obama per Executive Order initiierte *US Exascale Computing Project*, das trotz des Ziels, das Rennen um den ersten Exaflop-Supercomputer zu gewinnen, die Hälfte des 3,4 Mrd. USD Gesamtbudgets in die nachhaltige Softwareentwicklung investiert.

Und Deutschland? – Es ist eher als ein Entwicklungsland einzustufen, was die Förderung von nachhaltiger Forschungssoftware angeht. Zwar gab es bereits Initiativen, die Lebensdauer von Software, die aus DFG-Mitteln entwickelt wurde, über die Projektlaufzeit zu verlängern, aber nach wie vor sind Wissenschaftler meist darauf angewiesen, Ressourcen aus Forschungsprojekten einzusetzen, um eine bestehende Software aktuell und funktionsfähig zu halten. Eine langfristige Finanzierung von Personal im Sinne eines Research Software Engineers (RSEs) als Gegenstück zum wohl etablierten Labmanager für physikalische Infrastruktur ist eine Seltenheit. Dabei hat Forschungssoftware als virtuelles und kollaboratives Ökosystem oft eine viel größere Reichweite, weit über

die Institutionsgrenzen hinaus. Doch genau dieser Aspekt ist vielleicht auch ein Problem in einem klassischen Metriken gefangen Wissenschaftssystem, denn virtuelle Software ist weder

»Ein zukunftsfähiges Wissenschaftssystem benötigt Förderinstrumente für eine nachhaltige Entwicklung und Bereitstellung von Forschungssoftware.«

leicht medial in Szene zu setzen noch erhöht sie den Hirsch-Index. Zudem ist es bei kollaborativer Softwareentwicklung schwierig, einen Forschungserfolg einer einzelnen Gruppe oder Institution zuzuordnen. Dass nachhaltige Forschungssoftware dennoch eine notwendige Infrastruktur für die Zukunfts-

higkeit unseres Wissenschaftssystems bildet, spiegelt sich trotz aller Lippenbekenntnisse noch nicht in den Förderinstrumenten von Bund und Ländern wider.

Positionspapier mit Handlungsempfehlungen

Vor diesem Hintergrund haben sich im November 2019 mehr als 50 Wissenschaftler verschiedener Institutionen zu einem DFG-Rundgespräch im heute bundesweit bekannten Vorlesungssaal des Robert Koch-Instituts versammelt, um Nachhaltigkeitsaspekte von Forschungssoftware zu diskutieren und auf die Notwendigkeit von Förderinstrumenten zur nachhaltigen Entwicklung und Bereitstellung von Forschungssoftware aufmerksam zu machen. Im Nachgang wurde in *F1000Research* ein Positionspapier veröffentlicht, das zusammen mit einem offenen Peer-Review Prozess frei zugänglich ist. In diesem Positionspapier werden Handlungsempfehlungen an Forschungseinrichtungen, Forschungsgruppenleiter und Forschungsförderer vorgeschlagen, um die deutsche Wissenschaftslandschaft in puncto Forschungssoftware fit für die Zukunft zu machen. Es beinhaltet zudem eine Diskussion über Lizenzmodelle bzw. ob aus öffentlichen Mitteln geförderte Forschungssoftware nicht per se im Sinne einer permissiven Lizenz frei verfügbar sein sollte. Auch wenn man über die Handlungsempfehlungen und deren Anwendung unterschiedlicher Meinung sein kann, so sollte Konsens darüber bestehen, dass ein zukunftsfähiges Wissen-

schaftssystem die Etablierung von Research Software Engineers als Berufsbild mit langfristiger Perspektive benötigt sowie Infrastruktur zur Unterstützung nachhaltiger Softwareentwicklung und Förderinstrumente für eine nachhaltige Entwicklung und Bereitstellung von Forschungssoftware.

LITERATURHINWEIS

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

Das Positionspapier ist frei zugänglich unter <https://doi.org/10.12688/f1000research.23224.2>