

Nationale Forschungsdateninfrastruktur für wissenschaftliche Software (NFDI4RSE)

Sprecher: Frank Löffler, Friedrich-Schiller-Universität Jena, nfdi4rse-orga@de-rse.org

Software ist eine kritische Komponente aller datengetriebenen Forschung und damit auch **ein Grundbaustein der NFDI**. Wissenschaftler/-innen nahezu jeder Fachrichtung benutzen und entwickeln Software, um ihre Forschung zu untermauern und innovativ voranzutreiben. Mit Hilfe von Software werden in der NFDI Daten automatisiert erhoben, gespeichert, verarbeitet, integriert, analysiert, visualisiert und letztendlich durch Veröffentlichung mit anderen geteilt werden. Ziel von NFDI4RSE ist es, **interdisziplinär** die Nachhaltigkeit der Software im Rahmen der NFDI sicherzustellen. Ähnliche, schon vorhandene Ansätze in Fachkonsortien sollen unterstützt und vernetzt werden, sowohl untereinander, als auch mit Partnern außerhalb der NFDI. Letztere schließen Wissenschaftler/-innen in Deutschland mit starkem Softwarebezug ein, die eventuell im dann bestehenden NFDI-Gefüge nicht gleich ein eindeutiges Zuhause finden, aber auch ausländische Organisationen mit ähnlich gelagerten Zielen.

Es gibt in der deutschen Wissenschaftslandschaft bereits vielfältige Initiativen zum Thema Forschungssoftware. Diese umfassen unter anderem die Gesellschaft für Forschungssoftware (de-RSE e.V.), die Arbeitsgruppen des DFG-Programms „Nachhaltigkeit von Forschungssoftware“, organisationsspezifische Plattformen (z. B. das HIFIS Projekt der HGF) oder auch einzelne Gruppen an wissenschaftlichen Instituten und Universitäten (z. B. HU RSE (HU Berlin), die Software Engineering Gruppe des DLR, Research Software Engineering for Transformation Pathways des PIK (Potsdam)), und viele mehr. Auch wenn diese Organisationen die treibende Kraft von NFDI4RSE darstellen, so richtet sich NFDI4RSE jedoch gleichermaßen an einzelne Wissenschaftler/-innen oder kleine Gruppen mit Interesse oder Problemen im Zusammenhang mit Forschungssoftware. Daten, mit denen sich NFDI4RSE beschäftigt wird, sind einerseits Forschungssoftware selbst (z. B. Anwendung von FAIR-Prinzipien auf Software, Nachhaltigkeit von Software, Software als Publikation), als auch Metadaten über z. B. den Softwarelebenszyklus im praktischen, wissenschaftlichen Umfeld, sowie Analysen der Karrierelaufbahnen von hauptsächlich mit Forschungssoftware beschäftigten Wissenschaftlern und Wissenschaftlerinnen sowie deren Mitarbeitern und Mitarbeiterinnen.

Internationale Partner

Eine Nationale Forschungsdateninfrastruktur in Deutschland wird immer in das internationale Umfeld eingebettet sein müssen. Für das Querschnittsthema Forschungssoftware sind hier unsere engen Beziehungen zum Beispiel zum Software Sustainability Institute (SSI, UK), dem US Research Software Sustainability Institute (URSSI, USA), dem eScience Center (Netherlands), der UK Research Software Engineers Association (UKRSE, UK) und der Netherlands Research Software Engineer community (nl-RSE, NL) von Bedeutung. Weitere relevante Partner stellen ReSA, RDA und Software Heritage dar.

Software braucht mehr als nur Datenmanagement

Software wird oft als äquivalent zu Daten angesehen, sprich als statisches digitales Objekt. Jedoch weist Software eine Reihe spezieller Anforderungen und Eigenschaften auf, welche so bei klassischen Datenobjekten nicht auftreten: (1) Software muss gepflegt und kontinuierlich angepasst werden um in einem Umfeld wandelnder Software- und Hardwareumgebungen nutzbar zu bleiben. (2) Aufgrund der notwendigen Anpassungen und weiteren Entwicklung ändert sich Software über die Zeit. (3) Solche Veränderungen werden oft - insbesondere nach der initialen Publikation und "Abnahme" - von Externen vorgenommen. Diese grundsätzlich erstrebenswerten Kollaborationen werfen aber Fragen in

NFDI4RSE, initiiert von: de-RSE e.V. - Gesellschaft für Forschungssoftware

Kleineweg 79 · 12101 Berlin

Email: nfdi4rse-orga@de-rse.org · Web: www.rse4nfdi.de/de/

NFDI4RSE

Bezug auf Anerkennung, Verantwortlichkeit und Urheberrechte auf. (4) Softwareprojekte haben typischerweise niedrige Anforderungen in Bezug auf Speicherplatz, benötigen jedoch detaillierte Änderungshistorien, Versionsverwaltung, kollaborative Onlineplattformen und eine skalierende Recheninfrastruktur für kontinuierliches Testen und Benchmarks. (5) Die meiste Software baut auf existierenden Programmbibliotheken auf, welche auch über die Zeit einem Entwicklungs- und Änderungsprozess unterliegen. Die daraus entstehenden Abhängigkeiten müssen zwingend bei einer Langzeitarchivierung mit abgebildet werden, da ansonsten eine Reproduzierbarkeit von Ergebnissen nicht gewährleistet ist. (6) Software kann sowohl als Quellcode als auch als kompilierte Binärdatei weitergegeben werden, was zu komplexeren Lizenzierungsbedingungen führt.

Diese miteinander verwobenen und oft vernachlässigten Eigenschaften machen einen speziellen Umgang mit Forschungssoftware notwendig, um sie zu einem wirklich nachhaltigen und FAIRen Forschungsergebnis zu machen. Sollten diese Spezifitäten angemessen berücksichtigt werden, würde dies die datengetriebene Forschung auf höchstem Niveau ermöglichen. Eine weitere Vernachlässigung hingegen würde langfristig fundamentale wissenschaftliche Prinzipien von Transparenz und Reproduzierbarkeit verletzen.

Geplante Maßnahmen

Die geplante Arbeit von NFDI4RSE basiert auf den drei folgenden Säulen und ihren jeweiligen Schwerpunkten.

- Vereinbarung von Forschungssoftwarestandards in der NFDI durch Finden einer gemeinsamen Sprache (z. B. Bedeutung des Wortes "Workflow")
 - Erstellen von Handlungsanweisungen/-richtlinien
 - Langzeitarchivieren von Artefakten
 - Betrachten von Software auch als Kulturgut (Software Heritage)
- Softwareintegration über Domänengrenzen hinweg
 - Schaffung oder Unterstützung von Referenzarchitekturen
 - IT-Organisation (wo nötig, aber nicht durch andere möglich)
 - Anwendungsportfolio/Forschungssoftwareverzeichnis
- Unterstützung von Humanaktivitäten
 - Bereitstellen von Personalressourcen (z. B. für Durchführung/Koordination von Software-Carpentry-Kursen)
 - Schaffung einer Gemeinschaft von Praktiken (community of practise)
 - Neubewertung (z. B. Gutachterpool für Forschungssoftware)

Die NFDI soll bereichert werden durch die Bereitstellung von noch nicht existierenden Diensten und von Personen, die bei der Umsetzung und Vermittlung der FAIR-Prinzipien im Bereich der Forschungssoftware agieren.

Schnittstellen NFDI

Als Querschnittsinitiative wird NFDI4RSE themenspezifisch mit allen NFDI Konsortien interagieren. Des Weiteren planen wir auf Grund der inhärenten Nähe zur Informatik eine enge Abstimmung mit NFDIxCS, damit z. B. die im Rahmen der NFDI betriebenen Forschungssoftware-repositorien für wissenschaftliche Analysen genutzt werden können (Prozess der Softwareentwicklung als Forschungsgegenstand).

Querschnittsthemen

Als Querschnittsinitiative begrüßt NFDI4RSE die Unterstützung der Leipzig-Berlin-Erklärung [DOI:10.5281/zenodo.3895208] durch die 21 mitunterzeichnenden Fachkonsortien. Jenseits des in dem Papier dargelegten Querschnittsdienstes "Forschungssoftwaremanagement", plant NFDI4RSE auch substantielle Aktivitäten im Bereich Training, Ausbildung und Beratung (software carpentry). Als relevante Querschnittsthemen zur Erbringung dieser Dienste wurden die folgenden Punkte identifiziert: Personalentwicklung, Credit Systems, Langzeitarchivierung und Identitätsverwaltung.

Konferenzteilnahme

Software ist für die moderne datengetriebene Forschung von entscheidender Bedeutung. Forschungsdaten und Forschungssoftware bilden eine untrennbare Einheit. NFDI4RSE steht für Nachhaltigkeit und Interoperabilität in der NFDI über wissenschaftliche Disziplinen hinweg.

Eine Teilnahme an der NFDI-Konferenz würde uns einen aktuellen Überblick bezüglich der von den Fachkonsortien geplanten Aktivitäten im Bereich von Forschungssoftware geben und damit eine aktualisierte Bedarfsabschätzung ermöglichen. Des Weiteren erhoffen wir uns, das Bewusstsein für die Relevanz von zu schärfen und die Synergien durch ein koordiniertes Vorgehen in Sinn der Leipzig-Berlin-Erklärung unterstreichen zu können.

Erfahrungen und Hintergründe im Datenmanagement der Gruppe

NFDI4RSE setzt sich aus einer sehr heterogenen Personengruppe zusammen. Personen mit Expertise im Bereich Forschungssoftware bzw. dem Wunsch, auf dem Gebiet Fortschritte zu erzielen, kommen hier zusammen. Daher folgt hier eine kurze Auflistung, in welchen Kontexten Erfahrungen gesammelt wurden:

- Einige Personen sind in der Vorbereitung anderer NFDI-Konsortien involviert (NFDI4Earth, NFDI4AIRR, NFDI4Microbiota).
- Es gibt eine personelle Überlappung mit der AG ["Digitale Werkzeuge – Software und Dienste"](<https://www.allianzinitiative.de/handlungsfelder/digitale-werkzeuge-software-und-dienste/>) der Schwerpunktsinitiative „Digitale Information“ der Allianz der deutschen Wissenschaftsorganisationen.
- Andere sind Software-Carpentry-Instruktoren und haben viel Potential in der Weiterbildung von Forschenden erkannt. Egal ob Aus- oder Weiterbildung, die Kurse mit dem Ziel, Forscher fit im Umgang mit Software und Daten zu machen, sind regelmäßig ausgebucht.
- Mitarbeit bei FDM-Landesinitiativen (z. B. saxFDM, FDM Thüringen)
- Es wurden Erfahrungen im Aufbau einer Organisation gesammelt, die das Ziel hat, Forschungssoftwareentwickler aller Disziplinen zusammenzubringen (de-RSE e.V.). Eine deutschlandweite Konferenz zum Thema wurde etabliert, um fächerübergreifend Austausch zu gewährleisten.
- Mitarbeit in der Arbeitsgruppe des DFG-Programms „Nachhaltigkeit von Forschungssoftware“

Vorgesehene Mitglieder des Konsortiums (Co-Sprecherinnen/Co-Sprecher und die weiteren, beteiligten Institutionen):

Hier werden aufgrund der großen Ungewissheit der Antragsmodalitäten für Initiativen mit starkem Querschnittsbezug zum gegenwärtigen Zeitpunkt nur die Co-Sprecher aufgeführt.

Co-Sprecher/in	Zugehörige Institution
Dr. Christian Busse Projektleiter christian.busse@dkfz-heidelberg.de	Deutsches Krebsforschungszentrum (DKFZ) Im Neuenheimer Feld 280 69120 Heidelberg
Daniel Nüst Wissenschaftlicher Mitarbeiter daniel.nuest@uni-muenster.de	Institut für Geoinformatik, WWU Münster Heisenbergstraße 2 48149 Münster