

Positionspapier

NFDI4ING: Eine nationale Forschungsdateninfrastruktur für die Ingenieurwissenschaften

Robert Schmitt, Matthias Müller, RWTH Aachen

Peter Pelz, Thomas Stäcker, TU Darmstadt

Torsten Bronger, FZ Jülich

Irina Sens, TIB

Achim Streit, KIT

Ingenieurinnen und Ingenieure sind stark in der Produktion, Nutzung und Verarbeitung von Daten. Zugleich bestehen dringende Bedarfe und Wünsche bei der Schaffung übergreifender Plattformen, Etablierung verbindlicher Standards und einheitlicher Dienste, die es erlauben, Forschungsdaten zu sichern, optimal nachzunutzen und zu vernetzen. Die geplante Nationale Forschungsdateninfrastruktur (NFDI), die der Rat für Informationsinfrastrukturen (RfII) für die politische Umsetzung konkretisiertⁱ und aktuell noch einmal spezifiziert hatⁱⁱ, findet daher große Resonanz und Zuspruch in den Ingenieurwissenschaften. Die NFDI stellt eine einzigartige Chance für die Ingenieurwissenschaften dar, die für die Weiterentwicklung der Disziplinen notwendigen Infrastrukturen zu schaffen und auch Vernetzungsprozesse zu forcieren, um an internationale Entwicklungen wie die European Open Science Cloud (EOSC) in vollem Umfang anschlussfähig zu sein. Auch für die Fortführung der Zusammenarbeit mit der Industrie im Kontext Industrie 4.0 gibt es Bedarf an Methoden zum strukturierten Datenaustausch und –management, die durch eine NFDI wirksam unterstützt werden können.

Eine NFDI für die Ingenieurwissenschaften hat einen zeitlichen Nutzen, indem kurz- (<3 Jahre) und längerfristig die Weiterverwendung von Forschungsdaten innerhalb der Disziplin gefördert wird, darüber hinaus einen institutionsübergreifenden und interdisziplinären Nutzen, indem das Teilen von Forschungsergebnissen über Institutions- und Disziplinengrenzen hinweg erleichtert wird. So wird eine NFDI die Ingenieurwissenschaften unterstützen, sie wird Kooperationen fördern und mit den

Interessen der Wirtschaft und Gesellschaft verbinden. Die Ingenieurwissenschaften sehen großes Potential in der Bildung eines „polyzentrischen Netzwerks“ (RfII, März 2018), das maßgeblich von den wichtigen ingenieurwissenschaftlichen Akteuren definiert, mitgestaltet und mitgetragen wird. Es wird zugleich die Möglichkeit geschaffen, relevante Forschungsinfrastruktureinrichtungen, in der Regel Rechenzentren und Bibliotheken mit ingenieurwissenschaftlichen Schwerpunkten, in einem mit der Forschung koordinierten Prozess zu fördern, um ihre fachspezifischen Dienste einer Vielzahl von Wissenschaftlerinnen und Wissenschaftlern bestmöglich zu Gute kommen zu lassen.

Die Ingenieurwissenschaften unterstützen den NFDI-Prozess, weil erhebliche Effizienzgewinne für die Forschungsarbeiten zu erwarten sind. Rohdaten aus Experimenten zum Beispiel sollten zitierfähig, nachprüfbar und vor allem auch nachnutzbar sein, um kontinuierliche Forschungsprozesse zu ermöglichen. Durch die Umsetzung des Konzepts eines polyzentrischen Netzwerks lassen sich exzellente Rahmenbedingungen für innovative datenbasierte Forschung schaffen. Dies folgt der Einsicht, dass neue Erkenntnisse und Innovationen zunehmend aus der Verfügbarkeit, Integration und Analyse großer Datenmengen entstehen („von Daten zu Wissen“). Um diese Kernressource der Zukunft bereitzustellen, wollen Ingenieurinnen und Ingenieure ihre Daten teilen, vernetzen und nachnutzbar machen. Die Produktionstechnik liefert dafür mit dem Konzept des „Lab of Labs“ ein prägnantes Beispiel. Es steht für vernetztes Forschen und Lernen, wo etwa Prüfstände nicht mehr isoliert nebeneinander betrieben werden, sondern ihre Vernetzung den Zugriff auf die Messdaten der Partner erlaubt.

Data Sharing setzt Vereinbarungen und Regeln für die Dokumentation der Prozesse sowie gepflegte Metadaten voraus. Das Messsystem und die Validität der erhobenen Daten müssen überprüfbar sein. Neben den Daten muss der Softwarecode nachvollziehbar dokumentiert werden, um die Reproduzierbarkeit der Ergebnisse sicherzustellen. Große Potenziale bietet auch die Verfügbarkeit großer Datenmengen (Big Data) für Anwendungen der Künstlichen Intelligenz (KI), da sie die notwendige Grundlage für selbstständig lernende Verfahren (Machine Learning) bilden. Datenbestände dieser Dimension entstehen in Forschungseinrichtungen durch systematische Datenaufbereitung und Verfahren des Teilens und Vernetzens. Forschungsergebnisse lassen sich auf die Produktion übertragen, die Produktionssysteme lernen und passen sich an. Auf der anderen Seite erzeugen Techniken der Industrie 4.0 und der zunehmenden Digitalisierung extrem große Datenmengen (z.B. predictive maintenance), die für wissenschaftliche Fragestellungen nutzbar gemacht werden können. Diese spartenübergreifenden Szenarien brauchen Regelungen und Dienste, die in koordinierten Prozessen von Wissenschaftlerinnen und Wissenschaftlern, Forschungsinfrastruktureinrichtungen und weiteren Expertinnen und Experten zu erarbeiten sind.

Vor diesem Hintergrund vermag eine Nationale Forschungsdateninfrastruktur (NFDI) als zentrale Dachorganisation und ein in ihr organisiertes Konsortium „Ingenieurwissenschaften“ (NFDI4ING) die zukünftige Forschung nachhaltig zu befördern, indem die Datennutzung und -zugänglichkeit entscheidend verbessert werden. Das gilt umso mehr angesichts der disziplinspezifischen Besonderheiten der

Ingenieurwissenschaften. Deren verschiedene Sparten und Teildisziplinen insgesamt zu vertreten, ist zugleich Herausforderung und Chance.

Eines der wesentlichen Kennzeichen der Ingenieurwissenschaften ist ihre methodische Vielfalt. Sie arbeiten mit qualitativ und quantitativ erhobenen Daten, mit Experimenten, Simulationen und Feldstudien. Darin besteht die eigentliche Herausforderung an ein für die Ingenieurwissenschaften adäquates Forschungsdatenmanagement. Eine wissenschaftliche Methode der Datenbeschreibung und -dokumentation sichert die Qualität der Datensätze. Metadaten bspw. Parametereinstellungen und Umgebungsinformationen sind erforderlich, um automatisiert erzeugte Daten in ihren Kontext einzuordnen. Datensätze zu beschreiben und aufzubereiten erfordert besondere Kompetenzen, die im Rahmen der Aus- und Weiterbildung vermittelt werden müssen. Mit Forschungsdaten zu arbeiten, um Synergien zu schaffen und Effizienz zu fördern, verlangt auch langfristig einen Prozess des Umdenkens. Diese Leistung muss sich lohnen und gerade im akademischen Bereich honoriert werden. Die Ingenieurinnen und Ingenieure und ihre Communities werden sich im Rahmen des NFDI-Prozesses dafür einsetzen, neue Kriterien der Reputation und Bewertung zu entwickeln und einzuführen.

Beschreibung, Nachweis und Speicherung von Forschungsdaten haben die Forschungsinfrastruktureinrichtungen für die Wissenschaft schon zu ihrem Thema gemacht. Eine enge und partnerschaftliche Zusammenarbeit ist notwendig, um den Ingenieurwissenschaften leistungsstarke und benutzerfreundliche Arbeitsinstrumente bereitzustellen. Ansätze der Datenbeschreibung müssen fortentwickelt werden, aber flexibel handhabbar bleiben. Notwendig sind hier agile Methoden, wie sie auch in der Softwareentwicklung zum Einsatz kommen, bei denen die Nutzerinnen und Nutzer, d. h. die Wissenschaftlerinnen und Wissenschaftler, von Anbeginn in die Ausgestaltung mit einbezogen werden.

An diesen Aufgaben arbeitet seit kurzem eine Gruppe von Ingenieurinnen und Ingenieuren aus Technischen Universitäten mit ausgewiesenen Schwerpunkten in den Ingenieurwissenschaften: Die RWTH Aachen (mit dem Forschungszentrum Jülich/JARA), die TU Darmstadt und das KIT in Karlsruhe mit ihren Forschungsinfrastruktureinrichtungen, Rechenzentren und Bibliotheken sowie die TIB Hannover. Alle Partner sind einschlägig und vielfältig qualifiziert. Die TIB und das KIT verfügen als Mitglieder großer Wissenschaftsorganisationen (Leibniz- und Helmholtz-Gemeinschaft) über hochkompetente und leistungsstarke Forschungsinfrastrukturen und Dienste. Die Serviceeinrichtungen an den Universitäten zeichnen sich durch wissenschaftsnahe Angebote, vor allem im Bereich Beratung und Weiterbildung, Erfahrungen in Kooperationen und durch ihre internationale Vernetzung aus. Die Forschungsinfrastrukturangebote umfassen Repositorien für Publikationen und Forschungsdaten mit national und international anschlussfähigen Metadaten systemen sowie Forschungsumgebungen und Datenspeicher für unterschiedliche Anwendungen.

Um die Bedarfe und Interessen der ingenieurwissenschaftlichen Communities zu artikulieren und zu definieren, haben die Partner seit Herbst 2017

Koordinierungstreffen und Workshops organisiert, die sich aus einem kleineren Ansatz heraus entwickelt und aufeinander aufgebaut haben, siehe <https://www.ulb.tu-darmstadt.de/nfdi4ing>. Vertreten waren bisher vor allem die Disziplinen, in denen die genannten Universitäten und Forschungseinrichtungen führend sind: Materialwissenschaften, Produktionstechnik und Computational Engineering. Die Aufgaben und Ziele, die durch das Konsortium erkannt und weiter verfolgt werden sollen, sind aber für alle ingenieurwissenschaftlichen Disziplinen relevant: Prozessanalyse und -beschreibung (Transformationsprozess), Forschungsdatenpublikation und -zitationen als Kriterien wissenschaftlicher Reputation, formalisierte Qualifizierungs- und Weiterbildungskonzepte für Wissenschaftlerinnen und Wissenschaftler bzw. Expertinnen und Experten in den Infrastruktureinrichtungen sowie vor allem eine disziplinconforme Ausgestaltung der FAIR (Findable, Accessible, Interoperable, Reusable)-Prinzipienⁱⁱⁱ zur Verzeichnung und Nachnutzung von Forschungsdaten, die größtenteils bereits international anerkannt sind.

Das Konsortium, das sich hier konstituiert hat, versteht sich selbst als Nukleus einer weiteren Entwicklung. Es formuliert mit diesem Papier nicht nur die feste Absicht, einen konstruktiven inhaltlichen Beitrag zum Aufbau einer NFDI durch Vernetzung und Koordination der verschiedenen Stakeholder, Vertreterinnen und Vertreter der jeweiligen Fachdisziplinen und der Forschungsdateninfrastrukturen leisten zu wollen, sondern möchte auch insgesamt Ingenieurinnen und Ingenieure sowie Infrastrukturen mit ingenieurwissenschaftlichen Schwerpunkten einladen, sich dem Konsortium anzuschließen, um eigene Expertise einzubringen und den für die Etablierung einer forschunggetriebenen Dateninfrastruktur notwendigen Prozess des Community Building zu unterstützen. Über eine an der RWTH Aachen eingerichtete Kontaktstelle können sie sich weiter informieren und auch Interessensbekundungen zur Teilnahme aussprechen (www.nfdi4ing.de, contact@nfdi4ing.de).

ⁱ RfII – Rat für Informationsinfrastrukturen: Leistung aus Vielfalt. Empfehlungen zu Strukturen, Prozessen und Finanzierung des Forschungsdatenmanagements in Deutschland, Göttingen 2016. RfII – Rat für Informationsinfrastrukturen: Schritt für Schritt – oder: Was bringt wer mit? Ein Diskussionsimpuls zu Zielstellung und Voraussetzungen für den Einstieg in die Nationale Forschungsdateninfrastruktur (NFDI), April 2017. Online über: <http://www.rfii.de/de/category/dokumente/> (letzter Aufruf 5.4.2018).

ⁱⁱ RfII – Rat für Informationsinfrastrukturen: Zusammenarbeit als Chance. Zweiter Diskussionsimpuls zur Ausgestaltung einer Nationalen Forschungsdateninfrastruktur (NFDI) für die Wissenschaft in Deutschland, März 2018. Online über: <http://www.rfii.de/de/category/dokumente/> (letzter Aufruf 5.4.2018).

ⁱⁱⁱ <https://www.force11.org/group/fairgroup/fairprinciples>