
Nationale Forschungsdateninfrastruktur (NFDI)

Sophie Kraft¹, Hendrik Seitz-Moskaliuk¹, York Sure-Vetter¹, Elena Wössner¹, Nils Bohmer², Jan Eufinger³, Juliane Fluck⁴, Oliver Koepler⁵, Jan Korbel⁶, Bernhard Miller⁷, Sarah Pittroff⁸, Cristina Martins Rodrigues⁹, Thorsten Schwetje¹⁰, Dirk von Suchodoletz⁹ und Judith Sophie Weber¹¹

¹Nationale Forschungsdateninfrastruktur (NFDI) e.V., Karlsruhe

²DECHEMA e.V., Frankfurt am Main (NFDI4Cat)

³ Deutsches Krebsforschungszentrum Heidelberg (DKFZ) (GHGA)

⁴ZB MED - Informationszentrum Lebenswissenschaften, Köln (NFDI4Health)

⁵TIB - Leibniz Informationszentrum Technik und Naturwissenschaften, Hannover
(NFDI4Chem)

⁶European Molecular Biology Laboratory (EMBL), Genome Biology Unit, Heidelberg (GHGA)

⁷GESIS - Leibniz-Institut für Sozialwissenschaften. Mannheim und Köln (KonsortSWD)

⁸Akademie der Wissenschaften und der Literatur | Mainz (NFDI4Culture)

⁹Albert-Ludwigs-Universität Freiburg (DataPLANT)

¹⁰Technische Universität Darmstadt (NFDI4Ing)

¹¹MARUM - Center for Marine Environmental Sciences University of Bremen
(NFDI4BioDiversity)

Im Oktober 2020 gingen die ersten 9 Konsortien sowie der Verein Nationale Forschungsdateninfrastruktur (NFDI) e.V. an den Start. Ziel von NFDI ist die Schaffung übergreifender Strukturen für das Forschungsdatenmanagement in Deutschland in enger Anbindung an die Fachcommunities. Dieser Artikel fasst die Kurzvorstellungen der 9 Konsortien und des NFDI-Direktorats aus dem gemeinsamen Workshop der E-Science-Tage 2021 zusammen.

1 Einleitung

In der Nationalen Forschungsdateninfrastruktur (NFDI) werden die wertvollen Datenbestände von Wissenschaft und Forschung für das gesamte deutsche Wissenschaftssystem systematisch erschlossen, vernetzt und nachhaltig sowie qualitativ nutzbar gemacht. Bislang sind sie oftmals projektbezogen oder auf Zeit verfügbar und meist nicht miteinander verknüpft.

Bund und alle 16 Länder fördern NFDI gemeinsam auf Grundlage des §91b Grundgesetz. Der NFDI-Verein hat als wesentlichen Zweck die Förderung von Wissenschaft und Forschung durch eine Nationale Forschungsdateninfrastruktur, die ein übergreifendes Forschungsdatenmanagement in Deutschland etabliert und fortentwickelt und die Effizienz

des gesamten deutschen Wissenschaftssystems steigert. Eine ausführliche Beschreibung des Vereins findet sich in [1].

Zur Erfüllung des Vereinszwecks werden in einem wissenschaftsgeleiteten Verfahren, durchgeführt von der DFG, zwischen 2020 und 2022 bis zu 30 Konsortien ausgewählt, die Service-Portfolios für einen definierten Nutzerkreis entwickeln. Konsortien bilden sich entlang von Fachcommunities oder gemeinsam genutzter Datentypen. Stand 2021 werden bereits neun Konsortien gefördert.

Ein wesentliches Merkmal eines jeden Konsortiums ist die enge Zusammenarbeit von Wissenschaft und Wissenschaftsinfrastruktur. Die Konsortien arbeiten gemäß der in ihren Communities ermittelten Bedarfe und beteiligen sich darüber hinaus am Aufbau von NFDI insgesamt. Ziele dieser Partnerschaften sind die nachhaltige, qualitative und systematische Sicherung, Erschließung und Nutzbarmachung von Forschungsdaten über regionale und vernetzte Wissensspeicher, die Etablierung eines Forschungsdatenmanagements nach den FAIR-Prinzipien sowie die Anbindung und Vernetzung zu internationalen Initiativen wie der European Open Science Cloud (EOSC).

2 NFDI Konsortien

Im Folgenden stellen sich die neun Konsortien der ersten Auswahlrunde kurz vor.

DataPLANT

Das Hauptziel von DataPLANT besteht in der Unterstützung von Pflanzen-Grundlagenforschenden in Fragen des Forschungsdatenmanagements (FDM) [2]. DataPLANT als ein eher kompaktes Konsortium adressiert die wichtigsten Kernpunkte eines fachorientierten FDMs in drei zentralen Task Areas zu Standardisierung, zu technischen Diensten und Infrastruktur sowie zur persönlichen Unterstützung vor Ort. Letzteres geschieht über die sogenannten Data Stewards als Kernelement der DataPLANT FDM-Strategie [3]. Durch das in DataPLANT vereinbarte Präsenz-Modell der Data Stewards in den Forschungsgruppen profitieren diese von direkter und bedarfsgerechter Unterstützung. Data Stewards bilden die entscheidende Brücke zwischen den technischen Lösungen, der Infrastruktur und den Forschenden und fördern durch Unterstützung und Beratung im Daten- und Workflow-Management den Prozess der Standardisierung von Metadaten, Nutzung von Ontologien und Provenienz in der Datenverarbeitung. Bereits in der Antragsphase haben sich 32 Forschungsgruppen aus dem Bereich der Pflanzenforschung für eine Zusammenarbeit entschieden. Um sukzessive die gesamte Forschungslandschaft abzudecken, strebt DataPLANT an, zunächst große Verbünde, dann kleinere Forschungsgruppen und anschließend einzelne Personen einzugliedern.

DataPLANT arbeitet datenzentriert und baut auf bestehenden Strukturen auf. Ein zentrales Moment zur Zielerreichung ist der **A**nnotated **R**esearch **C**ontext (ARC), welcher

als Einstiegspunkt fungiert und zukünftig die Struktur einer Datenpublikation im Fachgebiet definiert. Der ARC wird den gesamten Forschungszyklus abdecken, vom Experiment über die rechnerischen Aspekte bis hin zu den eigentlichen Daten und Metadaten sowie die daraus resultierenden Publikationen. Es basiert auf bestehenden Formaten, Terminologien und Richtlinien, sodass die Integration bestehender Daten in bereits existierende Repositorien möglichst reibungslos erfolgen kann. Mit einem initialen Fokus auf Proteomik- und Transkriptomik-Datensätzen stehen bereits eine Reihe von Tools und Services [4], um von Beginn an einen schnellen Einstieg in ein verbessertes Datenmanagement zu gewährleisten. Gleichzeitig wird die technische Infrastruktur für die gemeinsame Nutzung und Versionieren von ARCs aufgebaut [5]. Dementsprechend stellt DataPLANT die zentrale Anlaufstelle für Forschende im Bereich der Pflanzen-Grundlagenforschung dar, um ein entsprechendes Forschungsdatenmanagement einzurichten.

GHGA

Ziel des NFDI-Konsortiums GHGA (German Human Genome-Phenome Archive) ist der Aufbau eines nationalen Genomarchivs für die sichere Speicherung, den Zugriff und die Analyse menschlicher Omics-Daten (z.B. Genome, Transkriptome) in einem einheitlichen ethisch-rechtlichen Rahmen. Solche Genomdaten und andere verwandte sogenannte Omics-Daten, die mithilfe moderner Sequenzierverfahren gewonnen werden, sind integraler Bestandteil der biomedizinischen Forschung. In Zukunft werden diese Daten auch die klinische Versorgung immer stärker prägen. Im Rahmen der Datennutzung für die Forschung muss das Bedürfnis, Daten offen und FAIR zu teilen immer mit dem Schutz der Privatsphäre der Patient:innen ausbalanciert und gegeneinander abgewogen werden. Zugriff kann dabei nur unter Einhaltung der notwendigen technischen und organisatorischen Schutzmaßnahmen und für legitime Forschungszwecke gewährt werden. Auf europäischer Ebene gibt es für diesen Zweck bereits das Europäische Genom-Phänom-Archiv (EGA). Da die zentrale EGA Infrastruktur die spezifischen nationalen Regelungen zum Datenschutz nur ungenügend abbilden kann, bereitet das EGA aktuell eine Umwandlung in eine föderierte Infrastruktur aus nationalen Knoten ("föderiertes EGA") vor. GHGA wird hierzu den deutschen Knoten innerhalb des föderierten EGA aufbauen [6].

Bei der Entwicklung von GHGA werden auch die Wünsche der Forschungsgemeinschaften nach effizienten, benutzerfreundlichen Analysen im großen Maßstab und zur Replikation von Ergebnissen auf anderen Kohorten berücksichtigt. GHGA setzt dabei auf existierenden, nationalen Omics-Datenlieferanten und deren IT-Infrastrukturen auf, um eine harmonisierte, interoperable Infrastruktur zu schaffen. Ziel ist es, Forschende in Deutschland in die Lage zu versetzen, humane Genomdaten rechtssicher entsprechend der FAIR-Richtlinien austauschen zu können und dabei internationale Standards zum Datenaustausch stärker mitzugestalten. GHGA ist dabei eingebunden in flankierende internationale Forschungsnetzwerke wie etwa die europäische *Beyond One Million Genomes Initiative*. Eine ausführliche Beschreibung von GHGA ist hier zu finden [7].

KonsortSWD

Mit dem Konsortium für die Sozial-, Bildungs-, Verhaltens- und Wirtschaftswissenschaften (KonsortSWD) wird innerhalb der NFDI ein bereits erfolgreich etabliertes Kooperationsnetzwerk zu einer integrierten Dateninfrastruktur weiterentwickelt [8]. Das Netzwerk hat das geteilte Verständnis, dass viele Daten nur dann überhaupt für die (Sekundär-)Nutzung bereitgestellt werden können, wenn sie bei den Datenanbieter:innen verbleiben. Dass diese Dezentralität gleichzeitig auch Effizienz bedeutet, hat mit der Vielfalt von Datentypen, vor allem aber mit der Vielfalt an rechtlichen und forschungsethischen Einschränkungen in den Disziplinen von KonsortSWD zu tun. KonsortSWD gründet wesentlich auf der Arbeit des Rates für Sozial- und Wirtschaftsdaten (RatSWD), der seit 2004 dem forschungsfreundlichen Zugang zu administrativen und später auch anderen Daten im Rahmen von Forschungsdatenzentren (FDZ) zum Erfolg verholfen hat. Heute stellen 39 FDZ Daten deutlich über die namensgebenden Sozial- und Wirtschaftswissenschaften hinaus zur Verfügung.

KonsortSWD setzt sich für möglichst leicht nutzbare, qualitativ hochwertige Daten für die Gesellschaftsforschung ein. Unser Forschungsdatenmanagement (FDM) wird daher so angelegt sein, dass es Forschende und FDZ technisch und inhaltlich bei Sicherung und Nachnutzung (neuer) sensibler und nicht sensibler Daten unterstützt. Eine Grundlage dabei sind die Arbeiten der Data Documentation Initiative (DDI) aufzubauen, einem international verwendeten Metadatenstandard, der maßgeblich von der KonsortSWD Community entwickelt wurde und wird. Insbesondere baut KonsortSWD die Unterstützung für die qualitative Sozialforschung aus: Forschende sollen u.a. durch Anonymisierungstools, Handreichungen zu Informed Consent über den gesamten Forschungsprozess beim FDM unterstützt werden. Auch werden die Zugangspunkte zu den Daten stärker vernetzt, um den Aufwand für die Nutzenden, z.B. beim Zugang zu sensiblen Daten in den FDZ, zu reduzieren.

NFDI4BioDiversity

So mannigfaltig die Fragestellungen und Zusammenhänge innerhalb der Biodiversität sind, so komplex und heterogen sind auch die Daten, die erhoben und analysiert werden, um diese zu durchdringen. Das Konsortium NFDI4BioDiversity arbeitet innerhalb der NFDI daran, diese Daten besser verfügbar zu machen, diese zu harmonisieren, zu integrieren und eine breitere Basis für Analysen zu schaffen. Und nicht nur die Daten selber, sondern auch die Akteure der Biodiversitätslandschaft sind vielfältig und reichen von wissenschaftlichen Institutionen zu Behörden, Bürgerwissenschaften, Museen, Sammlungen und Fachgesellschaften. Vertreter dieser Akteure bündelt NFDI4BioDiversity mit seinen 49 Partnern.

NFDI4BioDiversity baut dabei auf die Vorarbeiten der German Federation for Biological Data (GFBio)¹ auf, die bereits erprobte Services und Beratungsangebote zur Unterstützung von Forschenden in allen Phasen des Datenlebenszyklus aufgebaut hat. Das Angebot wird jetzt auf den NFDI4BioDiversity Nutzerkreis erweitert und um weitere Services und Tools ausgebaut. Nutzereinbindung und Praxistauglichkeit der Angebote werden durch insgesamt 23 Use Cases gewährleistet, die von der Anbindung einzelner institutioneller Datenressourcen über die Verbindung internationaler Netzwerke bis zum Ausrollen neuer Software-Tools für das Management und die Darstellung von Daten reichen.

Auf der technischen Ebene ist das zentrale Arbeitsziel die Entwicklung einer cloud-basierten Infrastruktur, der NFDI Research Data Commons. Dort sollen Daten aus verschiedenen Quellen zusammengebracht, integriert und harmonisiert werden. Dies beinhaltet auch standardisierte Workflows der jeweiligen Datenprovider, die ausgehandelt und weiterentwickelt werden müssen, sowie die Verwendung gängiger (Meta)datenstandards und Ontologien. Diese Schritte bilden die Grundlage für z.B. die Anwendung von Analysetools, das Bereitstellen von Trainingsdatensätzen, die Visualisierung der Daten in Dashboards oder die Darstellung von Datenpaketen für bestimmte Zielgruppen. Eine ausführliche Darstellung von NFDI4BioDiversity findet sich in [9].

NFDI4Cat

Um das volle Potenzial der in der Katalyseforschung generierten Daten auszunutzen, ist ein grundlegender, digitaler Wandel in diesem Wissenschaftsbereich sowie in der Prozess- und Verfahrenstechnik erforderlich. Aus dieser Notwendigkeit heraus wurde im Rahmen der Initiative zur Nationalen Forschungsdateninfrastruktur (NFDI) das NFDI4Cat-Konsortium gebildet, das seit Oktober 2020 von der DFG gefördert wird. Das NFDI4Cat-Konsortium zielt darauf ab, eine Transformation vom derzeitigen Status zur "digitalen Katalyseforschung" zu ermöglichen und zu beschleunigen. Die entscheidenden Elemente, die für eine "digitale Katalyseforschung" benötigt werden, sind die Vereinheitlichung von Konzepten, Vokabularen und Datenformaten sowie die Schaffung von vernetzten Informationsarchitekturen, die die Speicherung und den Austausch von semantisch reichen Daten ermöglichen. Diese Dateninfrastrukturen sollen den Einsatz moderner Analyseansätze ermöglichen, insbesondere mit Werkzeugen, die auf künstlicher Intelligenz basieren.

Das NFDI4Cat-Konsortium besteht aus 16 erfahrenen Partnern aus dem Bereich der homogenen, heterogenen, Photo-, Bio- und Elektrokatalyse und wird von der DECHEMA Gesellschaft für Chemische Technik und Biotechnologie e.V. koordiniert. Das Konsortium besteht außerdem aus Datenproduzent:innen und -nutzer:innen aus dem akademischen Bereich und aus anderen Kooperationsverbänden innerhalb der NFDI. Die Besonderheit innerhalb dieses Konsortiums ist die Beteiligung von industriellen Fachkolleg:innen mit beratender Funktion.

¹<http://www.gfbio.org>

NFDI4Chem

NFDI4Chem ist das Fachkonsortium Chemie innerhalb der NFDI, mit der Vision alle Prozesse im Umgang mit Forschungsdaten aus der chemischen Forschung zu digitalisieren und zu vernetzen [10]. Beginnend bei der Datenerzeugung, über deren Verarbeitung und Analyse bis hin zur Publikation entwickelt NFDI4Chem eine modular aufgebaute, vernetzte Infrastruktur. Zum Beispiel stellt sie Software-Tools wie das elektronische Laborjournal und Daten-Repository „Chemotion“ Forschenden zur Unterstützung im Laboralltag zur Verfügung. NFDI4Chem fördert das Forschungsdatenmanagement und Open Science gemäß der FAIR-Prinzipien (Findable, Accessible, Interoperable, Reusable). Der Digitalisierungsprozess wird weiterhin durch die Entwicklung von Ontologien zur semantischen Beschreibung von Daten, Standards für Daten- und Metadatenformate sowie Minimalinformationen unterstützt. NFDI4Chem möchte gemeinsam mit der wissenschaftlichen Community einen kulturellen Wandel zur Etablierung und Akzeptanz eines FAIRen Datenumgangs gestalten.

Die Gründung des Konsortiums begann im Jahr 2018 als Graswurzelbewegung durch den Zusammenschluss von Vertreter:innen aus universitärer und außeruniversitärer Forschung, Infrastruktureinrichtungen, Rechenzentren und nationalen Fachgesellschaften, wie der Gesellschaft Deutscher Chemiker (GDCh), der Deutschen Bunsen-Gesellschaft (DBG) und der Deutschen Pharmazeutischen Gesellschaft (DPhG). Aus dieser Initiative bildete sich das Fachkonsortium Chemie NFDI4Chem unter der Leitung von Christoph Steinbeck (Friedrich-Schiller-Universität, Jena) und Oliver Koepler (TIB - Leibniz Informationszentrum Technik und Naturwissenschaften, Hannover), welches damit mehr als 40.000 Mitglieder der chemischen Gemeinschaft repräsentiert.

NFDI4Culture

Ziel von NFDI4Culture ist der Aufbau einer bedarfsorientierten, forschungsgetriebenen Infrastruktur für Forschungsdaten des materiellen und immateriellen Kulturerbes innerhalb der NFDI. Mit seinem Fokus auf Kulturgüter deckt das Konsortium ein breites Spektrum an akademischen Disziplinen und ihrer Gegenstände ab: von der Musikwissenschaft, Kunstgeschichte und Architektur bis hin zur Theater-, Film- und Medienwissenschaft. Bisher existiert auf nationaler Ebene keine Struktur, die sich um die fächerübergreifende Auffindbarkeit und Zugänglichkeit sowie um die langfristige Sicherung und kontinuierliche Pflege von Forschungsdaten des kulturellen Erbes bemüht. NFDI4Culture will diese Lücke schließen [11]. Zu den Gegenständen des Konsortiums gehören digitale Repräsentationen (2D-Digitalisate von Gemälden, Fotografien und Zeichnungen, digitale 3D-Modelle kulturhistorisch bedeutender Gebäude, Denkmäler, audiovisuelle Daten von Musik-, Film und Bühnenaufführungen u.a. mehr), genuine Forschungsdaten sind ebenso ihre Metadaten, Annotationen oder andere durch Forschung am Objekt gewonnene Daten. Bei der Sicherung, Standardisierung und Bereitstellung von nachhaltiger Infrastruktur und an den Nutzer:innen orientierten Diensten des Forschungsdatenmanagements ist das Konsortium entlang der FAIR- und CARE-Prinzipien konzipiert.

Die von NFDI4Culture adressierte Forschungslandschaft ist durch eine hohe institutionelle Diversität gekennzeichnet. Die Forschungseinheiten in der Interessengemeinschaft des Konsortiums reichen von individuellen Forscher:innen bis zu Universitätsinstituten, Kunsthochschulen, Akademien sowie Institutionen des kulturellen Erbes wie Galerien, Bibliotheken, Archiven und Museen (GLAM). Konzept und Struktur des Konsortiums wurden über zwei Jahre in enger Zusammenarbeit zwischen elf Fachgesellschaften, neun Trägerinstitutionen und 52 Partnern entwickelt. Zu den Trägerinstitutionen gehören vier Universitäten (Köln, Heidelberg, Marburg, Paderborn), drei Infrastruktureinrichtungen (FIZ Karlsruhe, TIB Hannover, SLUB Dresden) und die Stiftung Preußischer Kulturbesitz. Sprecherinstitution ist die Akademie der Wissenschaften und der Literatur | Mainz.

NFDI4Health

Das Ziel von NFDI4Health ist es, in Deutschland eine Forschungsdateninfrastruktur für personenbezogene Gesundheitsdaten aufzubauen. Die Integration wichtiger deutscher Forschungsinstitute mit Erfahrung als Datenhalter, -analyst und Methodenentwickler macht NFDI4Health zu einem interdisziplinären Konsortium, das auf etablierten Strukturen, Kompetenzen und Know-How sowie einer zunehmenden Unterstützung und Teilnahme der Forschungsgemeinschaft aufbaut.

Deutschland verfügt über eine Fülle gesundheitsbezogener Daten aus gut strukturierten Langzeitstudien und Datenerhebungen bei gesunden Personen (epidemiologische / Public Health Studien) sowie aus klinischen Studien mit Patient:innen in Krankenhäusern, die eine umfassende Beschreibung der Studienteilnehmenden anhand von Fragebögen, medizinischen Untersuchungen und molekularen oder genetischen Profilen aufweisen. Durch ihren Längsschnittcharakter und ihre hohe Qualität sind diese Daten eine wertvolle Ressource für die Entwicklung von präventiven und therapeutischen Maßnahmen auf Individual- und Populationsebene.

NFDI4Health möchte ein umfassendes Inventar deutscher epidemiologischer, Public Health- und klinischer Studiendaten aufbauen. Diese Daten sollen nach den FAIR-Prinzipien (inter-)national zugänglich gemacht werden.

Die in die NFDI eingebetteten Aufgaben von NFDI4Health sind:

1. Auffindbarkeit von und Zugang zu strukturierten Gesundheitsdaten ermöglichen.
2. Föderalen Rahmen für Datenhaltungsorganisationen erhalten.
3. Austausch und Verknüpfung von personenbezogenen Daten unter Wahrung des Datenschutzes ermöglichen.
4. Automatisierte Dienste (z.B. Suche, Analysetools) etablieren.
5. Interoperabilität und Wiederverwendbarkeit der Daten etablieren und verbessern.
6. Anwendungsfallorientierte Zusammenarbeit zwischen Forschungsgemeinschaften fördern.

Eine ausführliche Beschreibung des Konsortiums findet sich in [12].

NFDI4Ing

NFDI4Ing repräsentiert die deutsche ingenieurwissenschaftliche Forschungslandschaft.

Das Konsortium vereint Institutionen und Personen aus allen Bereichen ingenieurwissenschaftlicher Tätigkeiten, die ihre jeweils spezifische Expertise, Erfahrung und Netzwerke in das Arbeitsprogramm des Konsortiums einbringen. Damit ermöglicht NFDI4Ing die enge Vernetzung der Perspektiven von Ingenieurwissenschaftler:innen, Anwender:innen und Diensteanbieter:innen, um auf das gemeinsame Ziel hinzuarbeiten, die FAIR-Prinzipien wissenschaftlichen Datenmanagements für die Ingenieurwissenschaften Realität werden zu lassen.

Eine zentrale Aufgabe von NFDI4Ing ist es, die große Vielfalt ingenieurwissenschaftlicher Forschungsansätze und -methoden in einer begrenzten Anzahl gemeinsamer Standards und Anforderungen an Forschungsdatenmanagement (FDM) zu konsolidieren. Dies ist ein fortlaufender Prozess. Seit 2017 konnten systematisch und methodengestützt Bedürfnisse, Erwartungen und Arbeitsabläufe in der ingenieurwissenschaftlichen Forschung identifiziert und in sieben typische ingenieurwissenschaftliche Forschungsprofile klassifiziert werden. Wir nennen diese Forschungsprofile unsere *Archetypen* und verwenden sie in NFDI4Ing zur Strukturierung und stetigen Weiterentwicklung unseres Arbeitsprogramms. In gleichnamigen Arbeitsgruppen werden typische Methoden und Verfahrensweisen harmonisiert und neue Anwendungsfelder eröffnet. Flankiert werden die Archetypen-Arbeitsgruppen durch Teilprojekte, die den Austausch mit den Fachcommunities gewährleisten (*Community Clusters*), und spezialisierte Dienste und Dienstleistungen für Ingenieur:innen entwickeln, standardisieren und bereitstellen (*Base Services*).

Angeleitet werden wir dabei von unseren acht *Kernzielen* [13], die z.B. von der lückenlosen Reproduzierbarkeit aller Teilschritte des Forschungsprozesses über die Automatisierung der Erzeugung und Verwaltung von Metadaten bis zur Etablierung von data literacy in allen Phasen ingenieurwissenschaftlicher Aus- und Weiterbildung reichen, und im Ergebnis ingenieurwissenschaftliches Forschungsdatenmanagement FAIR machen sollen.

3 Fazit

So unterschiedlich die fachlichen Hintergründe der 9 Konsortien sind, so ähnlich sind die Fragestellungen, die sie im Rahmen von NFDI bearbeiten: Wie werden die Bedarfe der Nutzenden von Dateninfrastrukturen ausreichend im Konsortium abgebildet und umgesetzt? Wie können Daten dauerhaft nach den FAIR-Prinzipien gespeichert und zugänglich gemacht werden? Welche rechtlich-ethischen Aspekte sind dabei zu beachten? Was sind geeignete Technologien und Metadatenformate? Wie können Forschende im Datenmanagement geschult werden? Und noch vieles mehr. Innerhalb der Konsortien liegt der Fokus

auf communityspezifischen Antworten. Gleichzeitig bietet der NFDI-Verein den Rahmen, konsortienübergreifende Lösungen zu finden und dabei von den unterschiedlichen Expertisen der Mitglieder zu profitieren. Der Verein nimmt seit Januar 2021 Mitglieder auf und die verschiedenen Vereinsorgane konstituieren sich im Lauf des ersten Halbjahres 2021. Gleichzeitig nimmt der konsortienübergreifende Austausch Fahrt auf, um gemeinsam am Forschungsdatenmanagement der Zukunft zu arbeiten.

Literaturverzeichnis

- [1] S. Kraft, A. Schmalen, H. Seitz-Moskaliuk, Y. Sure-Vetter et al., “Nationale Forschungsdateninfrastruktur (NFDI) e.V.: Aufbau und Ziele”, Bausteine Forschungsdatenmanagement, zur Veröffentlichung akzeptiert (2021).
- [2] D. von Suchodoletz, T. Mühlhaus, J. Krüger, B. Usadel, C. Martins Rodrigues, “DataPLANT – Ein NFDI-Konsortium der Pflanzen-Grundlagenforschung“, Bausteine Forschungsdatenmanagement, in Revision (2021).
- [3] D. von Suchodoletz, T. Mühlhaus, D. Brillhaus, H. Jasbeen, B. Usadel, J. Krüger, H. Gauza, C. Martins Rodrigues, “DataStewards as ambassadors between the NFDI and the community“, E-Science-Tage 2021, zur Veröffentlichung akzeptiert (2021).
- [4] C. Martins Rodrigues, J. Krüger, T. Mühlhaus, B. Usadel, M. Tschöpe, D. von Suchodoletz, “DataPLANT – Tools and Services to structure the Data Jungle of fundamental plant researchers“, E-Science-Tage 2021, zur Veröffentlichung akzeptiert (2021).
- [5] B. Venn, T. Mühlhaus, D. von Suchodoletz, J. Krüger, B. Usadel, C. Garth, “Immutable yet evolving: ARCs for permanent sharing in the research data-time continuum“, E-Science-Tage 2021, zur Veröffentlichung akzeptiert (2021).
- [6] G. Saunders, M.I. Baudis, R. Becker, S. Beltran, C. Bérout, E. Birney, C. Brooksbank, et al. “Leveraging European Infrastructures to Access 1 Million Human Genomes by 2022.” *Nature Reviews. Genetics* 20, no. 11, 693–701 (2019). <https://doi.org/10.1038/s41576-019-0156-9>.
- [7] J. Eufinger, J. Korb, E. Winkler, O. Kohlbacher, O. Stegle, “Genomdaten FAIR und sicher teilen: Das Deutsche Humangenom-Phänom Archiv (GHGA) als Baustein der Nationalen Forschungsdateninfrastruktur” Bausteine Forschungsdatenmanagement, zur Veröffentlichung akzeptiert (2021).
- [8] B. Hollstein, B. Miller, P. Siegers und C. Wolf, “KonsortSWD: Vom Netzwerk zur integrierten Dateninfrastruktur der Gesellschaftsforschung”, Bausteine Forschungsdatenmanagement, zur Veröffentlichung akzeptiert (2021).
- [9] J. S. Weber, B. Ebert, M. Diepenbroek, I. Kostadinov, F. O. Glöckner, “NFDI4BioDiversity - NFDI-Konsortium für Biodiversitäts-, Ökologische und Umweltdaten” Bausteine Forschungsdatenmanagement, in Revision (2021).

- [10] J. Ortmeyer, F. Schön, S. Herres-Pawlis, N. Jung, F. Bach, J. Liermann, S. Neumann, C. Popp, M. Razum, O. Koepler, C. Steinbeck, NFDI4Chem - Fachkonsortium für die Chemie, Bausteine Forschungsdatenmanagement, in Revision (2021).
- [11] R. Altenhöner, T. Schrade et al., NFDI4Culture - Consortium for research data on material and immaterial cultural heritage (2020). <https://doi.org/10.3897/rio.6.e57036>
- [12] J. Fluck und I. Pigeot et al., “NFDI4Health - Nationale Forschungsdateninfrastruktur für personenbezogene Gesundheitsdaten”, Bausteine Forschungsdatenmanagement, zur Veröffentlichung akzeptiert (2021).
- [13] R. H. Schmitt, V. Anthofer et. al. “NFDI4Ing - the National Research Data Infrastructure for Engineering Sciences” (2020). <https://doi.org/10.5281/zenodo.4015201>