

Das „Audimax Foyer (Gebäude 30.95)“, Straße am Forum 1, ist der Eingangsbereich des Audimax. Hier fanden zur CoRDI-Konferenz sowohl die Registrierung, viele Vorträge und Workshops, das Mittagessen, Kaffeepausen und Aktivitäten, wie der Marktplatz der Konsortien („Marketplace of the Consortia“), auf dem Campus Süd des Karlsruher Institut für Technologie (KIT),<sup>1</sup> in Karlsruhe, statt.



## Connecting Communities

Bericht über die „Conference on Research Data Infrastructure“, 1. Forschungsdateninfrastruktur-Konferenz der NFDI, Karlsruhe, 12.–14. September 2023

Annette Strauch-Davey

### FDM, NFDI, Konsortien, Base4NFDI und die CoRDI

Forschungsdatenmanagement (FDM) ist längst notwendig, um die Forschungsergebnisse, meistens von Projekten, aus Forschungskollaborationen reproduzierbar zu halten. In Forschungsteams müssen deshalb die Infrastrukturen immer neu gedacht und weiterentwickelt werden. Das Motto der ersten Konferenz *Conference for Research Data Infrastructures, CoRDI*,<sup>2</sup> welche vom Verein Nationale Forschungsdateninfrastruktur (NFDI) e.V.<sup>3</sup> initiiert wurde, lautete deshalb aus gutem Grunde „Connecting Communities“. Das Event war in der Tat ein Mekka für alle, die sich für einen guten Umgang mit Daten aus der Forschung interessieren oder die sich mit ihrer Arbeit dazu beschäftigen. Es gab sehr viele Gelegenheiten

zum Austausch an drei vollgepackten Tagen bis in die Abende hinein. Die CoRDI-Konferenz wurde vom 12. bis 14. September von der NFDI, zusammen mit dem KIT, in mehreren Gebäuden und Hörsälen des KIT ausgerichtet. Die NFDI erschließt wichtige Forschungsdaten für das gesamte deutsche Wissenschaftssystem systematisch. Es wird immer weiter und intensiver daran gearbeitet, dass die Forschungsdaten vernetzt, nachhaltig und qualitativ nutzbar werden, damit sie nicht einfach nur dezentral, projektbezogen oder auf zu kurze Zeit verfügbar sind. Die Deutsche Forschungsgemeinschaft (DFG) beschreibt die NFDI-Ziele, an die einleitend noch einmal erinnert werden soll, so:

„Die nationale Forschungsdateninfrastruktur (NFDI) soll die Datenbestände von Wissenschaft und Forschung systema-

<sup>1</sup> KIT - Karlsruher Institut für Technologie <https://www.kit.edu> (Zugriff am 14.09.2023)

<sup>2</sup> CoRDI 2023 | NFDI (Offizielle Webseite der Konferenz, <https://www.nfdi.de/cordi-2023> (Zugriff am 12.09.2023)

<sup>3</sup> NFDI | Nationale Forschungsdateninfrastruktur e. V., <https://nfdi.de> (Zugriff am 14.09.2023)

Vortragsfolien, Abstracts und Poster von der 1st Conference on Research Data Infrastructure (CoRDI) sind, soweit von den Vortragenden bereitgestellt, über die Plattform Zenodo verfügbar <https://zenodo.org/communities/cordi-2023/?page=1&size=20>

tisch erschließen, nachhaltig sichern und zugänglich machen sowie (inter-)national vernetzen.“<sup>4</sup>

Einrichtungen aus unterschiedlichen Forschungsbereichen, zum Beispiel aus dem Forschungszentrum Deutsches Elektronen-Synchrotron (DESY),<sup>5</sup> sind an den Arbeiten der NFDI beteiligt. Sie beschäftigen sich mit Entwicklungen von wichtigen Standards, Dienstleistungen und Schulungsangeboten für die Forschung.

### Generisches, disziplinspezifisches und interdisziplinäres FDM in der NFDI

Forschungsdatenmanagement (FDM) ist längst ein wichtiger Aspekt guter wissenschaftlicher Praxis.

„Unter Forschungsdaten sind (...) Daten zu verstehen, die im Zuge eines wissenschaftlichen Vorhabens z.B. durch Quellenforschungen, Experimente, Messungen, Erhebungen oder Befragungen entstehen.“

Letzteres ist die von der DFG 2009 formulierte Definition, die immer noch aktuell ist, auch wenn inzwischen die Forschungssoftware im FDM mehr berücksichtigt wird. Erkenntnisse und Innovationen aus Daten entstehen in allen wissenschaftlichen Disziplinen. Gut aufbereitete Forschungsdaten sind somit sehr wichtig für den Fortschritt der gesamten Gesellschaft. Damit sie benutzt werden können, ist eine starke Infrastruktur essentiell, weil sonst, wenn Forschungsdaten, wie so oft heutzutage noch der Fall, nur lokal und nur vorübergehend gespeichert werden, viele Forschungsdaten in der Forschungslandschaft leider verloren gehen. Der Verlust ist teuer. Alle Forschungsdaten aus Deutschland sollen nach den FAIR-Prinzipien (‘‘Findable, Accessible, Interoperable und Reusable’’ – auffindbar, zugänglich, interoperabel und wiederverwendbar) aufbereitet werden.<sup>6</sup> Bei den FAIR-Prinzipien liegt der Fokus auf der maschinellen Verarbeitbarkeit. So sollen Computersysteme die Forschungsdaten finden können, mit ihnen interagieren und sie ohne Eingriffe von Menschen wiederverwenden. In die NFDI war nun im Frühjahr 2023 so richtig Bewegung gekommen, erstens durch die neuen bewilligten NFDI-Konsortien<sup>7</sup> und zweitens durch die Vorbereitungen zur CoRDI-Veranstaltung, da hier im September 2023 der Austausch aller Konsortien untereinander mit Interessierten, dem seit März 2023 arbeitenden Verbund Base4NFDI,<sup>8</sup> anderen Forschungseinrichtungen und den Wissenschaftlern stattfinden konnte. NFDI-Konsortien



Abb. 1:  
Registrierung bei  
der CoRDI 2023

haben ja bereits im Oktober 2020, Oktober 2021 und im März 2023 ihre Arbeit aufgenommen. Sie sind alle beschäftigt innerhalb einer Forschungsdisziplin, aus unterschiedlichen Einrichtungen, und sie kommen sehr oft interdisziplinär und fächerübergreifend zusammen, um schließlich Forschungsdaten FAIR für andere zur Verfügung zu stellen. Die DFG führte das wissenschaftsgeleitete Verfahren zur Aufnahme von Konsortien durch.<sup>9</sup> Die Gemeinsame Wissenschaftskonferenz (GWK)<sup>10</sup> entschied über die Aufnahme aufgrund der Begutachtungen. Konsortien vernetzen Strukturen und Ansätze zum FDM und den Infrastrukturen in den einzelnen Fächern. Ein wichtiges Ziel in der nahen Zukunft ist die Konsolidierung der Konsortien. Der Reifegrad der Konsortien ist aktuell recht unterschiedlich, wie wir im Austausch bei der CoRDI sehen konnten. Im Eckpunktepapier der zweiten Förderphase (Erläuterungen weiter unten) heißt es, dass die Konsortien neu entstehenden Anforderungen helfen müssen. Ein Eckpunkt ist die Umsetzung eines nachhaltigen Organisationsmodells, um die Handlungsfähigkeit des jeweiligen Konsortiums in partizipativen Prozessen und Personalressourcen abzusichern.

### Veranstaltungen der CoRDI

Die über 700 Teilnehmenden der CoRDI hatten von Dienstag bis Donnerstag eine große Auswahl an Workshops und Präsentationen zu den Entwicklungen für ein qualitätsgerechtes, nachhaltiges Forschungsdatenmanagement. Am Dienstag, dem 12. September eröffneten ab 11:00 Uhr Mario Brandenburg, Parlamentarischer Staatssekretär der Bundesministerin für Bildung und

4 DFG - Deutsche Forschungsgemeinschaft – Nationale Forschungsdateninfrastruktur <https://www.dfg.de/foerderung/foerderinitiativen/nfdi/> (Zugriff am 15.09.2023)

5 Deutschlands größtes Beschleunigerzentrum - Deutsches Elektronen-Synchrotron DESY <https://www.desy.de/> (Zugriff am 16.09.2023)

6 FAIR Principles – GO FAIR (go-fair.org) <https://www.go-fair.org/fair-principles/> (Zugriff am 15.09.2023)

7 DFG - Deutsche Forschungsgemeinschaft – Geförderte NFDI-Konsortien [https://www.dfg.de/foerderung/foerderinitiativen/nfdi/gefoerderte\\_konsortien/](https://www.dfg.de/foerderung/foerderinitiativen/nfdi/gefoerderte_konsortien/) (Zugriff am 14.09.2023)

8 Base4NFDI – Home <https://base4nfdi.de/> (Zugriff am 16.09.2023)

9 DFG - Deutsche Forschungsgemeinschaft – Nationale Forschungsdateninfrastruktur <https://www.dfg.de/> (Zugriff am 14.09.2023)

10 Die Gemeinsame Wissenschaftskonferenz (gwkk-bonn.de) <https://www.gwk-bonn.de/> (Zugriff am 14.09.2023)



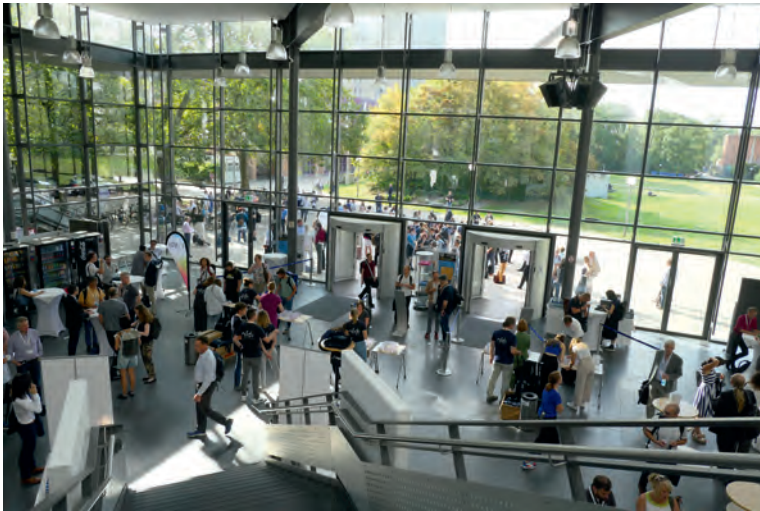


Abb. 2: Erster Vormittag der CoRDI



Abb. 3: Begrüßung York Sure-Vetter



Abb. 4: Eröffnung im Audimax

Forschung,<sup>11</sup> Kerstin Schill, DFG-Vizepräsidentin,<sup>12</sup> Ute Gunsenheimer, Secretary General EOSC,<sup>13</sup> Kora Kristof, Vizepräsidentin Digitalisierung und Nachhaltigkeit, KIT,<sup>14</sup>

sowie CoRDI Programme Chair Carole Goble von der University of Manchester & ELIXIR-UK<sup>15</sup> mit York Sure-Vetter, dem CoRDI General Chair, NFDI-Direktor<sup>16</sup> & KIT im Audimax die Konferenz.

Wissenschaftliche Vorträge gab es an allen drei Tagen auf Englisch, außerdem am Mittwoch eine Podiumsdiskussion und weitere Vorträge, daneben Poster Sessions und viele Gelegenheiten sich weiter mit den Konsortien zu vernetzen, z.B. auf dem „Market Place“ der Konsortien und der Base4NFDI.

### Vier disziplinspezifische Konferenz-Tracks

Es gab vier disziplinspezifische Konferenz-Tracks, die alle Infrastrukturaspekte behandelten, welche für die Disziplinen relevant sind:

1. Geistes- und Sozialwissenschaften
2. Naturwissenschaften
3. Lebenswissenschaften
4. Ingenieurwissenschaften.

### Sechs disziplinübergreifende Tracks (Bereitstellung von Infrastruktur)

Sechs Tracks waren interdisziplinär:

1. „**FDM ermöglichen, inkl. Software**“ (Ziel der Beiträge war die Entwicklung von Infrastruktur und Softwarekomponenten, die von der gesamten Gemeinschaft genutzt werden können, um Interoperabilität zu gewährleisten, damit Netzwerkeffekte systematisch genutzt werden können.),
2. „**Harmonisierung von FDM**“ mit Hilfe von (Meta-)Daten, Terminologien und Provenienz,
3. „**Sicherstellung von FDM**“ (ethische, rechtliche und soziale Fragen im Zusammenhang mit Forschungsdaten),
4. „**Verbreitung von FDM**“ durch Datenkompetenz,
5. Der Track „**Linking RDM**“ diskutierte wie zwischen Wissenschaft, Wirtschaft und Gesellschaft Forschungsdaten ausgetauscht werden können,
6. „**Connecting RDM**“ konzentrierte sich darauf, wie internationale und nationale Forschungsdateninfrastrukturen miteinander verbunden werden könnten.

### Keynotes<sup>17</sup>

Die erste Keynote präsentierte Prof. Christine Borgman von der University of California, Los Angeles, zum Thema

11 Mario Brandenburg – BMBF, [https://www.bmbf.de/bmbf/de/ueber-uns/die-leitung-des-hauses/mario-brandenburg/mario-brandenburg\\_node.html](https://www.bmbf.de/bmbf/de/ueber-uns/die-leitung-des-hauses/mario-brandenburg/mario-brandenburg_node.html) (Zugriff am 12.09.2023)

12 DFG - Deutsche Forschungsgemeinschaft – Professorin Dr. Kerstin Schill, [https://www.dfg.de/dfg\\_profil/gremien/presidium/mitglieder\\_praesidium/schill/index.html](https://www.dfg.de/dfg_profil/gremien/presidium/mitglieder_praesidium/schill/index.html) (Zugriff am 12.09.2023)

13 Ute Gunsenheimer - EOSC Association, <https://eosc.eu/staff/ute-gunsenheimer> (Zugriff am 12.09.2023)

14 KIT - Das KIT - Organisation - Leitung, <https://www.kit.edu/kit/organisation.php> (Zugriff am 12.09.2023)

15 Home (manchester.ac.uk), <https://staff.cs.manchester.ac.uk/~carole/> (Zugriff am 12.09.2023)

16 NFDI-Direktor im Amt: York Sure-Vetter wird Direktor der Nationalen Forschungsdateninfrastruktur (NFDI), <https://forschungsdaten.info/nachrichten/nachricht-anzeige/nfdi-direktor/> (Zugriff am 12.09.2023)

17 CoRDI 2023: Keynotes | NFDI, <https://www.nfdi.de/cordi-2023/> (Zugriff am 12.09.2023)

Wissensinfrastrukturen am Dienstag, 12. September, 11.30–12.30 Uhr. Der Titel ihres Beitrags lautete *“Knowledge Infrastructures: The Invisible Foundation of Research Data or, How Infrastructure Connects and Disconnects Research Communities”*.

In ihrer Präsentation ging es Borgman um die Investitionen in die Forschungsdateninfrastrukturen, die auf der Annahme beruhen, dass Daten kostbare Einheiten sind, d.h. die es wert sind, bewahrt, verwaltet, gemeinsam genutzt und wiederverwendet zu werden. Allerdings seien die Datenpraktiken immer lokal, so Borgman, und sie variieren von Forschungsbereich zu Forschungsbereich, von Person zu Person und von Land zu Land. Je größer die Zahl und Vielfalt der Forschungspartner würden, umso größer seien die Schwierigkeiten bei der gemeinsamen Nutzung der Daten und Wiederverwendung sowie für den dauerhaften Zugang zu Forschungsdaten.

*“Efforts to develop global research infrastructures are hindered by communities’ lack of agreement on data management practices – or on what constitutes ‘research data.’”*

Die Bemühungen zur Entwicklung globaler Forschungs- und Forschungsinfrastrukturen würden dadurch behindert, dass sich die jeweiligen Gemeinschaften in der Forschung immer noch nicht über die Praktiken des Datenmanagements einig seien. In ihrem Vortrag plädierte sie deshalb für einen breiteren Fokus auf die Wissensinfrastrukturen. Ihrer Ansicht nach seien die robusten Netzwerke von Menschen, Artefakten und Institutionen zur Produktion und zum Austausch und Aufrechterhaltung von Wissen elementar. Technische Aspekte der Infrastrukturen, d.h. von dauerhaften Identifikatoren bis hin zu Rechnerkapazitäten und Speicherplatz, seien leichter zu behandeln als die sozialen Aspekte, wie Datenverantwortung, Vertrauen, Governance, Wirtschaft, Infrastruktur, Normen und Wissenschaftspolitik.

***“May your problems always be technical, because those are the easier problems to solve.”***

(Prof. Borgman bei der CoRDI 2023)

Christine L. Borgman ist “Distinguished Research Professor” für Information Studies an der University of California, Los Angeles. Sie ist international anerkannt für ihre Forschung in den Bereichen Informations- und Computerwissenschaften, Datenwissenschaften, Kommunikation, digitale Geisteswissenschaften, Datenschutz und Recht. Ihre aktuellen Forschungsarbeiten konzentrieren sich auf Wissensinfrastrukturen, wissenschaftliche Datenpraktiken und dabei auch auf Open Science. Borgman war Gastwissenschaftlerin in Oxford, Harvard, Lund, Budapest Economic Sciences, Eotvos Lorand und Lough-



Abb. 5: Kora Kristof



Abb. 6: Carole Goble und York Sure-Vetter



Abb. 7: Keynote Speaker Prof. Christine Borgman



Abb. 8: *“Data are representations of observations, objects, or other entities.”*



borough sowie am "Dutch National Centre of Expertise and Repository", DANS.<sup>18</sup>

Am Dienstag nach der Eröffnungsveranstaltung gab es Vorträge zu "Enabling RDM" zwischen 11:00 Uhr und 12:30 Uhr im Audimax, zu "Harmonizing RDM" im Großen Hörsaal (Gebäude 10.50), "Connecting RDM" im Kleinen Hörsaal (Gebäude 10.50), "Securing RDM" im Seminarraum Forum A und B (Gebäude 30.95), gefolgt am Mittag zwischen 12:30 Uhr und 14:00 Uhr vom "Lunch Market of the Consortia" im Audimax Foyer. Am Dienstagnachmittag folgten im Audimax zwischen 14:00 Uhr und 15:30 Uhr Sitzungen im Bereich der Geistes- und Sozialwissenschaften.<sup>19</sup> Veranstaltungen zu den Lebenswissenschaften fanden im Kleinen Hörsaal, Gebäude 10.50 statt.<sup>20</sup> Alle Sitzungen können aufgrund der Fülle des Tagungsberichts hier nicht genannt werden. Einzelne Sessiondetails der jeweiligen Tage können auf der Webseite der NFDI zur CoRDI an den jeweiligen Tagen nachgelesen werden,<sup>21</sup> und es sollte an dieser Stelle auch auf die „Proceedings“ der Konferenz verwiesen werden, die bereits am 9. September veröffentlicht wurden: *Vol. 1 (2023): 1st Conference on Research Data Infrastructure (CoRDI) – Connecting Communities*, *Vol. 1 (2023): 1st Conference on Research Data Infrastructure (CoRDI) – Connecting Communities* <https://www.tib-op.org/ojs/index.php/CoRDI>

Die Künstlerin Julia Janssen<sup>22</sup> hielt am Dienstagabend einen Vortrag mit dem Titel "Behind the Click". Sie beschäftigt sich schon länger mit der Frage, wie wir in einer datengetriebenen Gesellschaft leben.

"How do we deal with fairness, equality, autonomy, freedom and democracy in a data-driven society?"

Julia Janssen ist eine Künstlerin, die den Einfluss der Digitalisierung auf unsere Gesellschaft erforscht. Sie richtete ihr Augenmerk auf Algorithmen, informierte Einwilligung und digitale Bürgerrechte. Seit März 2022 ist Janssen Bot-

schafterin für die niederländische gemeinnützige Stiftung Data Bescherming Nederland.<sup>23</sup> 2016 hatte sie ihren Abschluss in Grafikdesign an der ArtEZ School of the Arts gemacht. Mit ihrer Abschlussarbeit gewann sie den Crypto Design Award und widmet sich in ihrer Kunst seither der Datensouveränität.

Am Mittwoch, dem 13. September konnten sich die Teilnehmenden wieder zwischen 8:00 Uhr und 9:00 Uhr im Audimax Foyer registrieren. Zwischen 9:00 Uhr und 10:30 Uhr fand die zweite Poster Session statt, am Montag hatte es bereits die erste gegeben. Alle 53 gezeigten Poster sind auf der Seite der CoRDI Zenodo Community veröffentlicht worden: <https://zenodo.org/communities/cordi-2023/>

Am Mittwochabend fanden das Konferenzdinner, eine Exkursion zum KIT Campus Nord<sup>24</sup> statt oder Teilnehmende konnten die Schlosslichterspiele<sup>25</sup> in Karlsruhe besuchen und sich dort weiter kennenlernen und vernetzen. Es gab leuchtende Armbänder, sodass sich CoRDI-Teilnehmende im Dunkeln wiederfinden konnten. Tolle Angebote, die das NFDI-Vorbereitungsteam für die CoRDI-Teilnehmenden organisiert hatte.

### Eckpunkte der zweiten Förderphase der Konsortien

Am 21. August 2023 wurden die „Eckpunkte der zweiten Förderphase der NFDI-Konsortien“ publiziert, die ab 2025 beginnen wird.<sup>26</sup> Darin heißt es: „Das übergeordnete Ziel der Nationalen Forschungsdateninfrastruktur (NFDI) gemäß der ihr zugrunde liegenden Bund-Länder-Vereinbarung (BLV)1 ist die Etablierung und Fortentwicklung eines übergreifenden Forschungsdatenmanagements. Um dieses Ziel zu erreichen, stellen die NFDI-Konsortien bedarfsgerechte Dienste für die von ihnen adressierten Zielgruppen bereit und legen dadurch das

18 DANS | Centre of expertise & repository for research data (know.nl), <https://dans.know.nl/en/> (Zugriff am 12.09.2023)

19 P. Kamocki; E. Hinrichs; S. Springer; P. Leinen; A. Witt; D. Zechmann Open Science and Language Data: Expectations vs. Reality: The Role of Research Data Infrastructures • F. Thiery; A. Mees; B. Weisser; F. Schäfer; S. Baars; S. Nolte; H. Senst; P. von Rummel Object-related Research Data Workflows within NFDI4Objects and beyond • M. Fichtner; R. Nasarek; T. Wiesing WissKI: A Virtual Research Environment based on Drupal • S. Lieber; A. Van Camp; D. De Witte; E. Coudyzer; E. Buelinckx; E. Angenon; H. Lowagie; J. Birkholz; K. Lasaracina MetaBelgica Project: A Linked Data Infrastructure Between Federal Scientific Institutes in Belgium Natural Sciences I (Großer Hörsaal (Building 10.50)) • O. Koepler; C. Steinbeck; F. Bach; S. Herres-Pawlis; N. Jung; J. Liermann; S. Neumann; M. Razum Digitalizing the Chemical Landscape: A Comprehensive Overview and Progress Report of NFDI4Chem • L. Amelung; A. Barty; B. Murphy; C. Schneide; A. Schneidewind; T. Schoerner The DAPHNE4NFDI and PUNCH4NFDI Consortia in the NFDI • H. Weber; S. Brockhauser; C. Koch; L. Rettig; M. Aeschlimann; W. Hetaba; M. Grundmann; M. Kühbach; M. Krieger Research Data Management for Experiments in Solid-State Physics: Concepts • J. Bode; P. Jaeger; S. Schneidewind Integrating Data Literacy into University Curricula: Student Centred Learning in Undergraduate Physics Lab Courses

20 J. Fluck; M. Golebiewski; J. Darms Data publication for personalised health data: A new publication standard introduced by NFDI4Health • Pigeot; J. Fluck; J. Darms; C. Schmidt The NFDI4Health – Task Force COVID-19 • B. Ebert; J. Engel; I. Kostadinov; A. Güntsch; F. Glöckner Connecting National and International Data Infrastructures in Biodiversity Research • C. Goble; F. Bacall; S. Soiland-Reyes; S. Owen; I. Eguinoa; B. Drosesbeke; H. Ménager; L. Rodriguez-Navas; J. Fernández; B. Grüning; S. Leo; L. Pireddu; M. Crusoe; J. Gustafsson; S. Capella-Gutierrez; F. Coppens The EOSC-Life Workflow Collaboratory for the Life Sciences

21 Detaillierte Informationen zu allen Sessions finden sich auch hier auf der Veranstaltungswebseite: CoRDI 2023 | NFDI, <https://www.nfdi.de/cordi-2023/> (Zugriff am 12.09.2023)

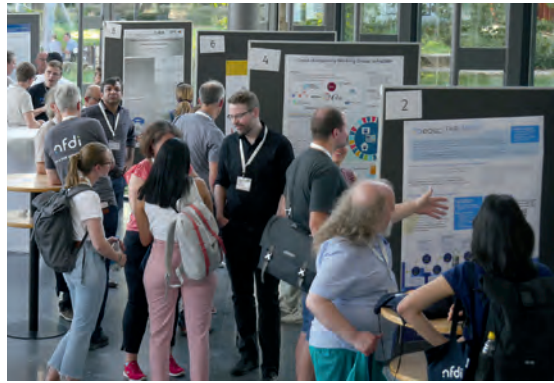
22 About | Studio Julia Janssen, <https://studiojulijanssen.com/about-me/> (Zugriff am 12.09.2023)

23 Over ons - Stichting Data Bescherming Nederland, <https://www.stichtingdatabescherming.nl/> (Zugriff am 12.09.2023)

24 Standort Karlsruhe - KIT Campus Nord - PTKA Projektträger Karlsruhe, <https://www.kit.edu/downloads/campus-nord.pdf> (Zugriff am 12.09.2023)

25 Abendprogramm Karlsruhe / Schlosslichterspiele, <https://www.schlosslichterspiele.info/> (Zugriff am 12.09.2023)

26 [https://www.dfg.de/foerderung/info\\_wissenschaft/2023/info\\_wissenschaft\\_23\\_69/index.html](https://www.dfg.de/foerderung/info_wissenschaft/2023/info_wissenschaft_23_69/index.html) (Zugriff am 14.09.2023)



*Fundament für eine communitygetriebene, vernetzte Informationsinfrastruktur.“*

In der „Leipzig-Berlin-Erklärung zu NFDI-Querschnittsthemen der Infrastrukturentwicklung“ steht, dass der „Erfolg einer Infrastruktur (...) immer am Mehrwert seiner Leistungen für die Nutzerinnen und Nutzer zu messen“ ist. Den Mehrwert erzeugen Infrastrukturen. Die Konsortien in der NFDI stellen bedarfsgerechte Dienste bereit, welche die Forschenden bei ihren Arbeiten unterstützen.<sup>27</sup> Hier im Folgenden sind die 26 NFDI-Konsortien (Stand September 2023)<sup>28</sup>, die nach DFG-Klassifikation eingeordnet sind und sich bei der CoRDI präsentierten, aufgelistet:

#### **Konsortien der Geistes- und Sozialwissenschaften:**

- **BERD@NFDI:** NFDI für Betriebswirtschaftslehre, Volkswirtschaftslehre und verwandte Daten
- **KonsortSWD:** Konsortium für die Sozial-, Bildungs-, Verhaltens- und Wirtschaftswissenschaften
- **NFDI4Culture:** Konsortium für Forschungsdaten zu materiellen und immateriellen Kulturgütern
- **NFDI4Memory:** Konsortium für historisch arbeitende Geisteswissenschaften
- **NFDI4Objects:** Forschungsdateninfrastruktur für die materiellen Hinterlassenschaften der Menschheitsgeschichte
- **Text+:** Sprach- und textbasierte Forschungsdateninfrastruktur

So werden Ergebnisse aus den Geistes- und Sozialwissenschaften beispielsweise computergestützt aufbereitet. Forschungsdaten des Sozio-oekonomischen Panel (SOEP)<sup>29</sup> sind Daten zu Einkommen, Erwerbstätigkeit, Bildung, Gesundheitsdaten und andere, nur um Beispiele zu nennen, die gemanagt werden müssen.

Das SOEP ist z.B. eine forschungsbasierte Infrastruktur.

#### **Konsortien der Ingenieurwissenschaften:**

- **NFDI4DataScience:** NFDI für Datenwissenschaften und Künstliche Intelligenz
- **NFDI4Energy:** Nationale Forschungsdateninfrastruktur für die interdisziplinäre Energiesystemforschung
- **NFDI4Ing:** NFDI für die Ingenieurwissenschaften
- **NFDI-MatWerk:** NFDI für Materialwissenschaft & Werkstofftechnik
- **NFDI4CS:** Nationale Forschungsdateninfrastruktur für und mit Computer Science

#### **Konsortien der Lebenswissenschaften:**

- **DataPLANT:** Daten in Pflanzen-Grundlagenforschung
- **FAIRagro:** FAIRe Dateninfrastruktur für die Agrosystemforschung
- **NFDI4Immuno:** Nationale Forschungsdateninfrastruktur für Immunologie
- **GHGA:** Deutsches Humangenom-Phenomarchiv
- **NFDI4Biodiversity:** Biodiversität, Ökologie und Umweltdaten
- **NFDI4BIOIMAGE:** Nationale Forschungsdateninfrastruktur für Mikroskopie und Bildanalyse
- **NFDI4Health:** NFDI für personenbezogene Gesundheitsdaten
- **NFDI4Microbiota:** NFDI für Mikrobiota-Forschung

#### **Konsortien der Naturwissenschaften:**

- **DAPHNE4NFDI:** Daten aus PHotonen- und Neutronen Experimenten
- **FAIRmat:** FAIRe Dateninfrastruktur für die Physik der kondensierten Materie und die chemische Physik fester Stoffe
- **NFDI4Cat:** NFDI für Wissenschaften mit Bezug zur Katalyse
- **MaRDI:** Mathematische Forschungsdateninitiative

*Abb. 9: V.l.n.r.: Poster Session 1, Poster Session der Konsortien 1, Poster Session 2*

<sup>27</sup> Maik Bierwirth, Frank Oliver Glöckner, Christian Grimm, Sonja Schimmler, Franziska Boehm, Christian Busse, Andreas Degkwitz, Oliver Koepler & Heike Neuroth. (2020). Leipzig-Berlin-Erklärung zu NFDI-Querschnittsthemen der Infrastrukturentwicklung. Zenodo, S. 1. DOI: <https://doi.org/10.5281/zenodo.3895209>

<sup>28</sup> Konsortien | NFDI, <https://www.nfdi.de/konsortien/> (Zugriff am 15.09.2023)

<sup>29</sup> DIW Berlin: Forschungsbasierte Infrastruktureinrichtung, Sozio-oekonomisches Panel, [https://www.diw.de/de/diw\\_01.c.615551.de/forschungsbasierte\\_infrastruktureinrichtung\\_\\_sozio-oekonomisches\\_panel\\_\\_soep.html](https://www.diw.de/de/diw_01.c.615551.de/forschungsbasierte_infrastruktureinrichtung__sozio-oekonomisches_panel__soep.html) (SOEP) (Zugriff am 15.09.2023)



Abb.10: Connecting Communities – Kevin Ashley, Daniel Mietchen und Annette Strauch-Davey bei der CoRDI 2023

- **NFDI4Chem:** Fachkonsortium Chemie in der NFDI
- **NFDI4Earth:** NFDI-Konsortium für Erdsystemforschung
- **PUNCH4NFDI:** Teilchen, Universum, Kerne und Hadronen für die NFDI

Neben den Konsortien gibt es den Verbund von Konsortien, **Base4NFDI**.<sup>30</sup> Letzterer ist ein Konsortienverbund der Basisdienste für NFDI, wie bereits erwähnt.<sup>31</sup>

Während der CoRDI stellten sich die Konsortien und der Verbund „Base4NFDI“ bei der Aktion „Marktplatz der Konsortien“ neben ihrem jeweiligen Konferenzposter vor. Perspektivisch, langfristig soll die NFDI an internationale Initiativen wie die der **European Open Science Cloud (EOSC)**<sup>32</sup> angeschlossen werden, um hier gemeinsam zusammenzuarbeiten. Im European Strategy Forum (ESFRI)<sup>33</sup> stimmen sich 27 Mitgliedstaaten in Europa zu europäischen Forschungsinfrastrukturen ab. Bis zu 90 Millionen Euro stellen Bund und Länder der Förderung der NFDI zwischen 2019 und 2028 zur Verfügung. Im März 2023 erklärte Bundesministerin Bettina Stark-Watzinger Ziele der NFDI, mit der Schaffung eines digitalen, disziplinübergreifenden Wissensspeicher für die deutsche Forschungslandschaft, denn die NFDI ist die vom Bundesministerium für Bildung und Forschung (BMBF) offiziell mandatierte Organisation, die zur Umsetzung der European Open Science Cloud (EOSC) auf europäischer Ebene beitragen wird:

*„Forschungsdaten sind ein Schatz, der hierzulande noch*

*nicht richtig gehoben wurde. Mit der Nationalen Forschungsdateninfrastruktur wollen wir das ändern, indem wissenschaftliche Ergebnisse aus unterschiedlichen Quellen, Standorten und Domänen besser zugänglich und nutzbar gemacht werden. Auf diesem Weg haben wir mit der Komplettierung der NFDI nun einen Meilenstein für die Digitalisierung von Wissenschaft und Forschung in Deutschland erreicht. Damit schaffen wir beste Voraussetzungen für neues Wissen und innovative Anwendungen. Die Beteiligung von mehreren Hundert Hochschulen, Forschungseinrichtungen und Datenzentren wird entscheidend dazu beitragen, die NFDI zu einem Vorzeigeprojekt mit internationaler Strahlkraft für den innovativen Umgang mit Forschungsdaten zu entwickeln. Auch Kooperationen mit der Wirtschaft werden dabei eine wichtige Rolle spielen.“*<sup>34</sup> In diesem Zusammenhang wäre die Digitalstrategie des BMBF zu nennen.<sup>35</sup> Eine langfristige Archivierung von Forschungsdaten ist Voraussetzung für die Nachvollziehbarkeit und Überprüfbarkeit wissenschaftlicher Ergebnisse weltweit, die auf der Auswertung dieser Daten beruhen.

### Leistungsfähige Infrastrukturen

Die großen Mengen an Forschungsdaten sollen also immer sicher aufbewahrt werden. Leistungsfähige Computer sind in der Virenforschung, Materialforschung nötig, bei KI-Anwendungen und Simulationen, beispielsweise zum Klimawandel oder auch in der Energieforschung, um nur ein paar Beispiele aus dem Alltag zu nennen. High Performance Computing ist in den genannten Fällen essentiell geworden, nun auch mit wichtigen Exascale-Rechnern (Q-Exa).

*„Der Supercomputer soll als erster Rechner in Europa die Grenze von 1 Trillion Rechenoperationen pro Sekunde – einer „1“ mit 18 Nullen – brechen.“*<sup>36</sup>

Leistungsfähige Informationsinfrastrukturen werden benötigt und diese kamen immer wieder in den Veranstaltungen der CoRDI zur Sprache, so in der Keynote von Christine Borgman (s. weiter oben) und auch die DFG fördert den Aufbau der erforderlichen Strukturen mit zwei Förderprogrammen:

- Informationsinfrastrukturen für Forschungsdaten und

30 Base4NFDI - Home, <https://base4nfdi.de/> (Zugriff am 15.09.2023)

31 „Bei seiner Arbeit baut Base4NFDI auf zwei Säulen auf:

1) organisatorisch auf der partizipativen Mitgestaltung der Basisdienste-Entwicklung durch die Communities im NFDI Verein;  
2) technologisch auf einem gemeinsamen Rahmen, der qualitätsgesicherte und kohärente Modelle für die kontinuierliche Identifizierung, Förderung, Entwicklung, den Betrieb und die Bewertung von NFDI-weiten Basisdiensten schafft.“ Base4NFDI | NFDI (Zugriff am 16.09.2023)

32 EOSC Portal ([eosc-portal.eu](https://eosc-portal.eu)) (Zugriff am 15.09.2023)

33 [www.esfri.eu](https://www.esfri.eu) (Zugriff am 15.09.2023)

34 Nationale Forschungsdateninfrastruktur – BMBF, [https://www.bmbf.de/bmbf/de/forschung/das-wissenschaftssystem/nationale-forschungsdateninfrastruktur/nationale-forschungsdateninfrastruktur\\_node.html](https://www.bmbf.de/bmbf/de/forschung/das-wissenschaftssystem/nationale-forschungsdateninfrastruktur/nationale-forschungsdateninfrastruktur_node.html) (Zugriff am 14.09.2023)

35 Digitale Wirtschaft und Gesellschaft – BMBF, [https://www.bmbf.de/bmbf/de/forschung/digitale-wirtschaft-und-gesellschaft/digitale-wirtschaft-und-gesellschaft\\_node.html](https://www.bmbf.de/bmbf/de/forschung/digitale-wirtschaft-und-gesellschaft/digitale-wirtschaft-und-gesellschaft_node.html) (Zugriff am 14.09.2023)

36 Erster europäischer Exascale-Superrechner kommt nach Jülich ([fz-juelich.de](https://blogs.helmholtz.de/research-field-information/2022/06/20/erster-europaeischer-exascale-superrechner-kommt-nach-juelich/)), <https://blogs.helmholtz.de/research-field-information/2022/06/20/erster-europaeischer-exascale-superrechner-kommt-nach-juelich/> (Zugriff am 15.09.2023)



• e-Reserch-Technologien<sup>37</sup> (Support der Forschung mit Verfahren und Tools).<sup>38</sup>

Es ist sehr komplex! In den NFDI-Konsortien<sup>39</sup> werden für das Managen von Forschungsdaten bereits erste Standards gesetzt, und Vieles wird sich erst entwickeln. Bei technischen Infrastrukturen spielen das „Identity Management“ eine Rolle oder ein „Single Sign-On“ zum Beispiel, außerdem Repositorien, Datenversionskontrollen, Teilen von Forschungsdatensätzen oder Forschungsdatensammlungen, Data Publication mit Metadaten und persistenten Referenzen. Letzteres waren Themen bei der CoRDI. Wie bestimmte Großgeräte verwaltet werden, gehört auch mit in die Arbeit zu den Lösungen von Infrastrukturen für alle.

Die Ziele der NFDI, die bei der CoRDI immer wieder genannt wurden, sind zusammengefasst diese:

- **Zugang zu Forschungsdaten** systematisch und nachhaltig verbessern,
- **Dezentral, projektförmig, temporär gelagerte Datenbestände erschließen**,
- **Gesamtstruktur** untereinander verknüpfter Konsortien schaffen,
- **Anbieter und Nutzer** zusammenbringen,
- **Standards für ein interoperables Forschungsdatenmanagement** entwickeln,
- Gemeinsame **Basis für Datenschutz, Souveränität, Integrität und Qualität** von Daten schaffen,
- **Services** aufbauen,
- **Internationale Anschlussfähigkeit** sicherstellen.

Insgesamt kann so durch neue Lösungen eine nationale Datensouveränität sichergestellt werden. Bei personenbezogenen Forschungsdaten müssen Daten oft anonymisiert werden. Die Interoperabilität innerhalb von Infrastrukturen, also das 'I' in FAIR, zur Nachnutzung von Forschungsdaten, ist essentiell. Schließlich dienen Forschungsinfrastrukturen der Spitzenforschung.

## Connecting Communities

Wir wollen den Klimawandel besser verstehen und Infektionen bekämpfen. Bei der Entwicklung von Arzneimitteln, der Modellierung des Universums, für neue Konzepte im Bereich Solarenergie ist das Hoch- und Höchstleistungsrechnen sehr wichtig geworden. Mit Blick auf die Infrastrukturen dürfen wir nicht das Erheben der großen Datenmengen übersehen, wie in der Erdsystemforschung, bei der Bildgebung in der Medizin oder im Kontext der Astro- und Teilchenphysiksysteme, aber auch in den Sozialwissenschaften. Große Datenmengen fallen auch beim autonomen Fahren an. Forschende fragen nach Zugangsverfahren und in den Disziplinen immer nach passgenauen Lösungen. Die HPC-Systeme sind sehr komplex. In der Teilchenphysik konnte es durch Computersimulationen gelingen, die physikalischen Phänomene der kleinsten Bausteine des Universums aufzuklären. Ein deutschlandweiter föderierter Zugang zu Forschungsdatenressourcen fehlt aktuell. Zur Etablierung eines fächerübergreifenden Forschungsdatenmanagement und zur internationalen Vernetzung waren bei der CoRDI internationale Partner eingeladen, Personen, die sich mit Forschungs- und Forschungsdateninfrastrukturen beschäftigen. Als ein Beispiel wäre die europäische Infrastruktureinheit aus den Lebenswissenschaften Elixir Europe<sup>40</sup> zu nennen. Es wird auch hier global daran gearbeitet, dass z.B. die Metadatenstandards<sup>41</sup> disziplinübergreifend sind. Es werden sich in allen Gemeinschaften und Vernetzungen Regeln zum standardisierten Umgang mit Forschungsdaten etablieren. **I**



### Annette Strauch-Davey

M.A., ist „Section Liaison Officer“ bei Base4NFDI<sup>42</sup> am Deutschen Elektronen-Synchrotron (DESY),<sup>43</sup> Deutschlands größtem Beschleunigerzentrum.  
annette.strauch-davey@desy.de

37 „Da jede Infrastruktur unterschiedliche Phasen von der Bedarfsanalyse bis zum regelhaften Betrieb durchläuft, kann die Förderung beantragt werden, um den Auf- und Ausbau von e-Research-Technologien in drei verschiedenen Phasen funktional und temporär zu unterstützen. Für diese drei Phasen gelten unterschiedliche Bedingungen (s. unter V – Ergänzender Leitfaden). Gefördert werden können Vorhaben – zur anwendungsbezogenen Entwicklung und Erprobung von e-Research-Technologien; – zur Implementierung von e-Research-Technologien; – zur Konsolidierung und Optimierung bestehender e-Research-Technologien. Ausgeschlossen von der Förderung sind durchweg solche Vorhaben, deren Zielsetzung oder Durchführung als Grundaufgabe der Trägereinrichtung zu sehen ist. Dazu gehören z. B. reine technische Erneuerungen, Wartungsarbeiten oder Nutzungen.“ (DFG-Merkblatt - Merkblatt und ergänzender Leitfaden e-Research-Technologien, [https://www.dfg.de/foerderung/programme/infrastruktur/lis/lis\\_foerderangebote/e-research\\_technologien/formulare\\_merkblaetter/index.jsp](https://www.dfg.de/foerderung/programme/infrastruktur/lis/lis_foerderangebote/e-research_technologien/formulare_merkblaetter/index.jsp) (Zugriff am 14.09.2023))

38 DFG - Deutsche Forschungsgemeinschaft - Forschungsdaten in der Infrastrukturförderung, [https://www.dfg.de/foerderung/grundlagen\\_rahmenbedingungen/forschungsdaten/infrastrukturfoerderung/index.html](https://www.dfg.de/foerderung/grundlagen_rahmenbedingungen/forschungsdaten/infrastrukturfoerderung/index.html) (Zugriff am 14.09.2023)

39 DFG - Deutsche Forschungsgemeinschaft - Geförderte NFDI-Konsortien, [https://www.dfg.de/foerderung/foerderinitiativen/nfdi/gefoerderte\\_konsortien/](https://www.dfg.de/foerderung/foerderinitiativen/nfdi/gefoerderte_konsortien/) (Zugriff am 14.09.2023)

40 ELIXIR | A distributed infrastructure for life-science information, <https://elixir-europe.org/> (Zugriff am 16.09.2023)

41 Services: Interoperability and standards | ELIXIR, <https://elixir-europe.org/services> (Zugriff am 16.09.2023)

42 Base4NFDI – People, <https://base4nfdi.de/people> (Zugriff am 17.09.2023)

43 FTX Members, <https://ftx.desy.de/members/> (Zugriff am 17.09.2023)