

Forschungsdatenmanagement für die Mikrobiologie

Besser forschen durch offene und FAIRe Daten

LORENZ CHRISTIAN REIMER¹, KONRAD U. FÖRSTNER², JÖRG OVERMANN¹

¹ LEIBNIZ-INSTITUT DSMZ – DEUTSCHE SAMMLUNG VON MIKROORGANISMEN UND ZELLKULTUREN

² ZB MED – INFORMATIONSZENTRUM LEBENSWISSENSCHAFTEN

DOI: 10.1007/s12268-022-1725-6

© Springer-Verlag GmbH 2022

„Ist das Bakterium in der Publikation das gleiche wie das in meinem Labor?“, ist eine gängige Frage, die häufig nicht so einfach zu beantworten ist. Das liegt vor allem daran, dass die notwendigen Metadaten fehlen, wodurch die Daten schlecht mit den eigenen Ergebnissen verknüpfbar und vergleichbar sind. Um den Umgang mit Forschungsdaten zu verbessern, schlug im Jahr 2016 der Rat für Informationsinfrastrukturen (RFII) die Etablierung einer Nationalen Forschungsdateninfrastruktur (NFDI) vor [1]. Ziel ist es, den kulturellen Wandel innerhalb aller Forschungsdisziplinen hin zu einer digitalen und vernetzten Wissenschaft zu fördern und basierend auf den FAIR-Prinzipien (*findable, accessible, interoperable, reproducible*) [2] den offenen Zugang zu wissenschaftlichen Erkenntnissen (Open Science) zu verbessern.

Entgegen dem technisch klingenden Namen soll dabei vornehmlich in Köpfe investiert und eine breite Vermittlung von Datenkompetenz in der Forschung vorangetrieben werden. Dies schließt auch ganz neue Rollen wie Data Stewards ein, die zukünftig Wissenschaftler:innen beim Umgang mit großen Datensätzen und deren Hinterlegung in Repositorien unterstützen sowie auf die Einhaltung von Standards und guter wissenschaftlicher Praxis achten. Die NFDI wird aus etwa 30 fachlichen Konsortien der wissenschaftlichen Gemeinschaften gegründet [3]; so gibt es z. B. NFDI4Chem, NFDI4Ing und eben NFDI4Microbiota. Erklärtes Ziel ist es, dass die verschiedenen Konsortien eine

gemeinsame NFDI entwickeln, die allen Forschenden Angebote bereitstellt und fachübergreifend die Erstellung, Bearbeitung und Speicherung von Daten und Metadaten verbessert.

NFDI4Microbiota will die mikrobiologische Forschung unterstützen. Die besondere Herausforderung liegt hier in der rapiden Zunahme der Forschungsdaten durch den Einsatz von Hochdurchsatztechnologien. Aufgrund fehlender Infrastruktur und mangels entsprechend ausgebildeten Personals werden Daten nicht immer optimal ausgewertet, ausreichend beschrieben und sind daher nicht nutzbar. So fehlen oft die Metadaten, die eine Nachnutzung und Verknüpfung der Daten erst ermöglichen. Gleichzeitig ist für die Analyse von vielen Datentypen ein hohes Maß an bioinformatischer Expertise, spezifische Software und Rechenleistung nötig, die nicht an jedem Forschungsstandort vorhanden sind.

Das Konsortium NFDI4Microbiota (nfdi4microbiota.de) besteht aus zehn etablierten Forschungs- und Infrastrukturinstitutionen aus Deutschland sowie über 50 Partnerinstitutionen. Gemeinsam bieten sie Dienstleistungen an, u. a. ein breites Trainingsangebot zur Handhabung von Daten und ein Helpdesk als Anlaufstelle für Forschende. Darüber hinaus werden analytische Services und Wissensdatenbanken für die Forschung von Mikrobiolog:innen aufgebaut. So soll eine Datenbank die Eingangsfrage beantworten, ob verwendete Bakterien vergleich-

bar sind. Hierfür werden persistente Identifikatoren für mikrobiologische Stämme gesammelt, abgeglichen und mittels einer leicht bedienbaren Weboberfläche zur Verfügung gestellt. Alle Services werden über ein zentrales NFDI4Microbiota-Portal erreichbar sein. Durch enge Zusammenarbeit innerhalb der NFDI sowie mit anderen Infra-

strukturvorhaben, wie z. B. de.NBI oder das EMBL-EBI, werden nachhaltige Lösungen entwickelt, die eine breite Akzeptanz erlangen.

Die Vision des Konsortiums ist, dass Daten von ihrer Generierung bis zur Publikation einen bruchfreien und nachvollziehbaren Workflow durchlaufen. Damit soll (1) der Aufwand für die Forschenden verringert, (2) die Transparenz und die Reproduzierbarkeit verbessert sowie (3) die Qualität und Geschwindigkeit der mikrobiologischen Forschung langfristig erhöht werden. Insbesondere entstehen durch die flächendeckende Verfügbarkeit qualitativ hochwertiger Forschungsdaten ganz neue Möglichkeiten der Vernetzung und Nachnutzbarkeit (z. B. durch KI-Analysen). Momentan befindet sich NFDI4Microbiota in der Startphase und wird in den nächsten Monaten beginnen, die Community weiter zu integrieren. ■

Literatur

- [1] Rat für Informationsinfrastrukturen (2016) Leistung durch Vielfalt. <https://rfii.de/download/rfii-empfehlungen-2016/> (Zugegriffen: 16.02.2022)
- [2] Wilkinson MD, Dumontier M, Aalbersberg JJ et al. (2016) The FAIR guiding principles for scientific data management and stewardship. *Scientific Data* 3: 160018
- [3] DFG (2022) Nationale Forschungsdateninfrastruktur. <https://www.dfg.de/foerderung/programme/nfdi/> (Zugegriffen: 18.02.2022)



Lorenz Christian Reimer, Konrad U. Förstner und Jörg Overmann (v. l. n. r.)

Korrespondenzadresse:

Prof. Dr. Jörg Overmann
Leibniz-Institut DSMZ – Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH
Inhoffenstraße 7 B
D-38124 Braunschweig
joerg.overmann@dsMZ.de
<https://nfdi4microbiota.de>



◀ **Abb. 1:** *Salmonella spec.* (rot) in menschlichem Gewebe, eingefärbte EM-Aufnahme. Bild: Rocky Mountain Laboratories, NIAID.