

Nationale Forschungsdateninfrastruktur für Mikroskopie, Biophotonik und Bildanalyse NFDI4BIOIMAGE

*Sprecher/in: Dr. Stefanie Weidtkamp-Peters, Heinrich-Heine-Universität Düsseldorf,
stefanie.weidtkamp-peter@hhu.de*

Mikroskopie und verwandte bildgebende spektroskopische Technologien stellen wichtige Schlüsseltechnologien für alle naturwissenschaftlichen und biomedizinischen Fachgebiete dar. Die Weiterentwicklung dieser Methoden und die Auswertung der digitalen Daten mithilfe von Bildanalyse-Verfahren bilden zudem ein weitreichendes Forschungsfeld ab. Die Initiative NFDI4BIOIMAGE setzt sich zum Ziel als methodenspezifisches Konsortium für biologische und biomedizinische Bilddaten die Bedarfe im Forschungsdatenmanagement disziplinübergreifend zu erfassen und sowohl existierende, und bereits international vernetzte, als auch neue Lösungsansätze für erfolgreiches Forschungsdatenmanagement zu harmonisieren. Als fachspezifisches Konsortium vertreten wir zugleich Forschende im Bereich biophotonische Technologien und Bildanalyse (*Bioimage informatics*).

Mikroskopie und verwandte bildgebende Technologien haben sich in den vergangenen zwei Jahrzehnten von überwiegend qualitativen hin zu quantitativen „*big data*“ Ansätzen entwickelt. Die heute fast ausschließlich digitalen Bilddaten erlauben detaillierte Einblicke in Gewebe, Zellen und subzelluläre Bestandteile, und auch unbelebte Materialien, mit hoher räumlicher und zeitlicher Auflösung und unter Einbeziehung zusätzlicher Modalitäten wie z.B. Fluoreszenzlebenszeit oder Anisotropie des Fluoreszenzsignals. Dabei mangelt es an einheitlichen Standards beim Management dieser Daten, die oft nur lokal gespeichert und archiviert werden und mangels hinreichender Metadatenannotation oft nicht über das eigene Forschungsprojekt hinaus wiederverwendbar sind. Ein Grund ist, dass biologische Bilddaten zu den besonders großen und hochdimensionalen Datentypen gehören, deren Komplexität die Schaffung von methodenspezifischen und fachgebietsübergreifenden Standards bisher verhindert hat und die geeignete technische Voraussetzungen für die Umsetzung eines Forschungsdatenmanagements erfordern. Dies gilt insbesondere in den Bereichen der Hochdurchsatzmikroskopie und der multispektralen Bildgebung. Neben Datenmodellen und frei verfügbarer, nachhaltiger Software für die Metadatenannotation ist daher auch die Erarbeitung nicht-proprietärer und einheitlicher Dateiformate, die auch für Cloud-Umgebungen und Objektspeichersysteme geeignet sind, sog. *next-generation file formats*, (NGFF), ein hochaktuelles Anliegen im Forschungsdatenmanagement für biologische Bildgebung. Die abgestimmte Entwicklung neuer Formate, Datenmodelle und technischer Infrastrukturen schafft dabei für das Forschungsfeld der Bilddatenanalyse die notwendigen Voraussetzungen, um moderne Analysemethoden einer breiten wissenschaftlichen Gemeinschaft zugänglich zu machen, sowie Datenprovenienz und Reproduzierbarkeit auszubauen und nachhaltig zu sichern. Die rasante Entwicklung der auf künstlicher Intelligenz und neuronalen Netzwerken basierenden Bildanalysemethoden erfordern verbesserte Standards der begleitenden Metadaten inkl. der eindeutigen Identifizierung von Prozessierungsschritten und Software-Versionen. Aus diesen Gründen ergeben sich für biologische Bilddaten bisher nicht gelöste Herausforderungen bei der Implementation der FAIR-Prinzipien, die wir mit der Initiative NFDI4BIOIMAGE adressieren möchten. Wesentliche Maßnahmen zur Verbesserung des Datenmanagements beinhalten:

- Harmonisierte Metadatenstandards (insb. *controlled vocabularies and ontologies*)
- Vereinfachte Richtlinien und Hilfsmittel für die disziplin- und methodenspezifische Annotation der Metadaten

- Weiterentwicklung nicht-proprietärer Cloud-fähiger Dateiformate (NGFF)
- Integration multimodaler Datentypen mit Bilddaten(banken)
- Einfache Anbindung an und Datentransferlösungen für bilddatenspezifische Repositorien (z. B. *BioImage Archive*, *Image Data Resource*)
- Entwicklung von Datenprovenienzstandards für Prozessierung und Analyse von biologischen Bilddaten
- Verbesserung und Verknüpfung von Datenmanagement-Plattformen (z. B. OMERO) und elektronischen Laborbüchern
- Entwicklung modularer Datenmanagementplan-Vorlagen für biologische Bilddaten
- Umfangreiche, frei zugängliche Bildungsmaterialien sowie Weiterbildungsangebote im Bereich Forschungsdatenmanagement für biologische Bilddaten
- Exzellente Kommunikationskanäle zur iterativen und langfristigen Einbindung der fachwissenschaftlichen Gemeinschaft sowie der interessierten Öffentlichkeit

Zur Umsetzung dieser und weiterer Maßnahmen streben wir die Rekrutierung von Expert*innen in den Bereichen Data Science, Softwareentwicklung, Bildanalyse und verwandten Themenbereichen an, die in interdisziplinären Teams in den Bereichen 1) (Meta-)Datenstandards, 2) Technische Infrastrukturen für Bilddatenmanagement, 3) Integration multimodaler Daten, 4) Bilddatenanalyse und 5) Etablierung von Trainings- und Schulungskonzepten zusammenarbeiten werden. Ein signifikanter Teil der Mittel soll für Personal zur Unterstützung der wissenschaftlichen Gemeinschaft im Sinne von *Data Stewards* für biologische Bilddaten und Bildanalyse verwendet werden. Hierbei spielen auch Fallbeispiele und *Best-Practice*-Lösungen eine wichtige Rolle.

Die Mitglieder der Initiative NFDI4BIOIMAGE setzen sich aus Fachwissenschaftler*innen, Leiter*innen von Mikroskopie-Serviceeinrichtungen, Bioinformatiker*innen aus dem Bereich Bildanalyse, Expert*innen für Datenmanagement und Forschungssoftware für Bilddaten, sowie Infrastruktureinrichtungen für Informationstechnologien zusammen. Der Großteil in Deutschland erhobener Bioimaging-Daten wird in Kooperation mit oder durch direkte Nutzung von Mikroskopie-Serviceeinrichtungen (*Core Facilities*) generiert. German BioImaging – Gesellschaft für Mikroskopie und Bildanalyse (GerBI-GMB) ist ein Zusammenschluss von über 70 Mikroskopie *Core Facilities* in Deutschland und seit 2017 in Fortführung des DFG-geförderten German Bioimaging Netzwerks als gemeinnütziger Verein zur Förderung der Wissenschaft eingetragen. Seither wurden die Aktivitäten zur gemeinsamen Erarbeitung von Forschungsdatenmanagementstandards intensiviert und mündeten 2019 in der Gründung der *Research Data Management for Microscopy* (RDM4Mic) Gruppe. In regelmäßigen, informellen Treffen findet ein offener Erfahrungsaustausch über standortspezifische Lösungen sowie Vernetzung mit internationalen Partnerinitiativen statt. Die Initiative NFDI4BIOIMAGE hat sich in diesem Umfeld aus Mitgliedern von GerBI-GMB, RDM4Mic, der Leibniz Research Alliance Health Technologies – Medical Photonics, sowie weiteren Fachexpert*innen aus dem Forschungsbereich (inkl. Exzellenzcluster und Sonderforschungsbereiche) konstituiert. Es besteht eine etablierte Zusammenarbeit mit Mitgliedern des *Open Microscopy Environment* (OME) Teams, Euro-BioImaging Bio-Hub (EMBL-Heidelberg), Global BioImaging, European Bioinformatics Institute (EMBL-EBI) und anderen internationalen Partnerinitiativen (z. B. *4D-Nucleome project*; *Quality Assessment and Reproducibility for Instruments and Images in Light Microscopy*, QUAREP-LiMi). Mitglieder von NFDI4BIOIMAGE sind zudem an Aktivitäten internationaler Forschungsdatenmanagement-Initiativen (z. B. der Research Data Alliance, RDA) beteiligt. Mehrere Mitglieder sind bereits in bestehenden NFDI-Konsortien oder Konsortialinitiativen aktiv. Trotz der sehr verbreiteten Verwendung und strukturellen

Bedeutung von biologischer Bildgebung (vor allem in den Naturwissenschaften) werden biologische Bilddaten von bestehenden fachspezifischen Konsortien und Konsortialinitiativen aufgrund der dargelegten Komplexität und Größe von Bilddaten nur in geringem Umfang adressiert. Daher bilden die Datentypen per se wichtige Schnittstellen zu anderen Konsortien. Zur Vorbereitung unseres Arbeitsprogramms streben wir neben dem bestehenden Austausch mit DataPLANT, NFDI4Health, GHGA, DeBioData, NFDI-Neuro, NFDI4Immuno, NFDI4Patho, NFDI4Chem und NFDI4Phys eine enge Abstimmung mit allen interessierten Konsortien und Konsortialinitiativen an. Dies betrifft sowohl die Einbindung fachspezifischer Besonderheiten für Bilddaten und Integration multimodaler Daten als auch die Zusammenarbeit an Querschnittsthemen gemäß der Berlin-Leipzig-Erklärung zu NFDI Querschnittsthemen der Infrastrukturentwicklung¹ und der Zweiten Stellungnahme des NFDI-Expertengremiums im November 2020². Relevant sind für uns u. a. die abgestimmte Vision und Strategie der NFDI, Nachhaltigkeit von technischen Infrastrukturen, Personalentwicklung und Berufsperspektiven, *Authentication and Authorization Infrastructure* (AAI), Transfer von großen Datentypen, Weiterbildungs- und Trainingsangebote. Die Weiterentwicklung neuer Dateiformate, die Implementierung von Datenmanagementplattformen für Bilddaten inkl. deren Anbindung an Fachrepositorien, und die Integration multimodaler Datenformate stellen von uns adressierte und konsortienübergreifend relevante Querschnittsthemen in der NFDI dar.

In diesem Jahr nehmen wir zum dritten Mal an der NFDI-Konferenz teil, die aus unserer Sicht eine wichtige Gelegenheit ist, uns sowohl innerhalb der etablierten und geplanten Konsortien als auch gegenüber der interessierten Öffentlichkeit näher vorzustellen. Dabei möchten wir unsere wesentlichen Ziele und Maßnahmen, unsere geplante Konsortialstruktur und Möglichkeiten zur aktiven Mitwirkung in unserem Konsortium darlegen. Einen besonderen Fokus möchten wir auf den Austausch mit anderen Konsortien für eine enge Abstimmung von Maßnahmen und Zusammenarbeit bei interdisziplinären Anwendungsfällen legen. Eine nähere Erörterung der in der dritten Ausschreibungsrunde anvisierten Basisdienste und anderer Querschnittsthemen erachten wir als wichtigen Inhalt der NFDI-Konferenz.

Wir laden alle Interessierten ein, sich auf unserer Webseite <https://nfdi4bioimage.de> zu informieren, sich in unseren Newsletter einzutragen und mit uns in Kontakt zu treten. Über eine Teilnahme an unserem NFDI4BIOIMAGE Community Survey (<https://nfdi4bioimage.de/en/survey>) würden wir uns zudem sehr freuen. Die Ergebnisse möchten wir nach Auswertung auch für andere Konsortien und die interessierte Öffentlichkeit zugänglich machen.

¹ <https://doi.org/10.5281/zenodo.3895209>

² https://www.dfg.de/download/pdf/foerderung/programme/nfdi/stellungnahme_nfdi_201112_de.pdf

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

Co-Sprecher/in	Zugehörige Institution
(apl.) Prof. Dr. Elisa May	Universität Konstanz Bioimaging Center
Susanne Kunis	Universität Osnabrück Integrated Bioimaging (iBiOs)
Joshua Moore	German Biolmaging e.V. (GerBI)
Pavol Bauer, Ph. D.	Leibniz Institut für Neurobiologie, Magdeburg
PD Dr. Thomas Bocklitz	Leibniz Institut für Photonische Technologien
Dr. Anna Kreshuk	European Molecular Biology Laboratory, Heidelberg
Prof. Dr. Marc Thilo Figge	Leibniz Institut für Naturstoffforschung und Infektionsbiologie – Hans Knöll Institut, Jena
Dr. Robert Haase	Technische Universität Dresden, Exzellenzcluster Physics of Life
Dr. Thomas Zobel	Westfälische Wilhelms Universität Münster Imaging Netzwerk
Dr. Björn Grüning	Universität Freiburg, Head of the Freiburg Galaxy Team
Beteiligte Personen	Zugehörige Institution
Dr. Jan Bumberger	Helmholtz Zentrum für Umweltforschung, Leipzig
Dr. Markus Becker	Leibniz Institut für Plasmaforschung und Technologie, Greifwald
Matthias Landwehr	Universität Konstanz Kommunikations-, Informations- und Medienzentrum
Dr. Antje Keppler	European Molecular Biology Laboratory, Heidelberg, Direktorin des Euro-Biolmaging Bio-Hub
Dr. Roland Nitschke Dr. Dirk von Suchodoletz	Universität Freiburg, Life Imaging Center Rechenzentrum, Leiter Abt. eScience
Dr. Timo Dickscheid	Forschungszentrum Jülich
PD Dr. Claire Chalopin	Innovationszentrum für Computer-assistierte Chirurgie der Universität Leipzig
Torsten Stöter	Leibniz Institut für Neurobiologie, Magdeburg
Dr. Carl-Magnus Svensson	Leibniz Institut für Naturstoffforschung und Infektionsbiologie – Hans Knöll Institut, Jena
Dr. Bastian Hülsmann, Dr. Christian Meesters	Universität Mainz
Dr. Astrid Schauss, Peter Zentis	Universität zu Köln, Exzellenzcluster Cellular Stress Responses in Aging- Associated Diseases (CECAD)
Kyle Harrington, Ph. D.	Helmholtz Imaging Platform
Matthew Hartley, Ph. D.	European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Biolmage Archive
Dr. Janina Hanne	German Biolmaging - Gesellschaft für Mikroskopie und Bildanalyse, Geschäftsführerin