

DAM KERNBEREICH „DATENMANAGEMENT UND DIGITALISIERUNG“ FORSCHUNGSDATENMANAGEMENTKONZEPT FÜR DIE DEUTSCHE MEERESFORSCHUNG

DAM-AG „Datenmanagement und Digitalisierung“

Stand September 2022

1. Ziel	3
2. Hintergrund	3
3. Ausgangslage	5
4. Konzeptionelle Überlegungen für den Aufbau eines kollaborativen Datenökosystems	6
4.1. Voraussetzungen bei den DAM-Mitgliedseinrichtungen:	6
4.2. Datenmanagement-Helpdesk	8
4.3. Gemeinsames Nutzungs- und Betriebsmodell für Dateninfrastrukturen	9
4.4. DAM als Koordinationsplattform	10
5. Datenmanagementkonzept in der Praxis	11
5.1. Datenmanagement in Projekten von Personen	11
5.2. Datenmanagement in Verbundprojekten von Organisationen	11
6. "Unterwegs"-Forschungsdaten	12
7. Referenzen	13

1. Ziel

Die deutsche Meeresforschung benötigt ein kollaboratives "Datenökosystem": Eine gemeinsam nutzbare, verteilte, leistungsstarke und stetig betriebene Forschungsdateninfrastruktur, um Information und Wissen zu sichern und Forschungsdaten für Nutzer:innen aus Wissenschaft, Behörden, Wirtschaft und Öffentlichkeit frei zugänglich zu machen.

Eines der zentralen Ziele dieses Konzeptes besteht darin, die Vereinbarung zur Verknüpfung bewährter Forschungsdateninfrastrukturen einzelner DAM-Mitgliedseinrichtungen, zur Öffnung für die gesamte Deutsche Meeresforschung und zu einem gemeinsamen Betriebskonzept zu treffen. Dafür sollten sich die DAM-Mitgliedseinrichtungen auf ein Konzept verständigen, das ihre institutionelle Souveränität im Umgang mit Forschungsdaten wahrt und gleichzeitig auf die Vereinheitlichung von Prozessen und Mindeststandards hinwirkt.

2. Hintergrund

Hinter dem Begriff **Forschungsdatenmanagement (FDM)** steckt **eine explizite Anforderung an die Wissenschaft**, die unter anderem in einer Leitlinie der Deutschen Forschungsgemeinschaft [1] (DFG) und den Förderrichtlinien der Programme Horizon 2020 und Horizon Europe beschrieben ist. Die verstärkte Aufmerksamkeit für FDM ist dem gesteigerten Interesse an der Professionalisierung und Standardisierung der administrativen und wissenschaftlichen Aufgaben im Umgang mit Daten geschuldet. FDM bringt der wissenschaftlichen Arbeit einen Mehrwert und fördert den Wissenschaftsbetrieb als Ganzes durch Erhöhung der Sichtbarkeit von Forschungsdaten, der Effizienz im Alltag der Wissenschaftsarbeit, der Datensicherheit sowie der Chancen auf Forschungsförderung:

Sichtbarkeit und Nachnutzung von Forschungsdaten

Die Veröffentlichung von Forschungsdaten erhöht die Sichtbarkeit von Projekten und gewährleistet gleichzeitig die Transparenz der Forschung. Durch die Verwendung nachnutzbarer Dateiformate und umfassender Beschreibungen (Metadaten) können Forschungsdaten langfristig in Folgeprojekten sowie von externen Forschenden wieder- und weiterverwendet werden. Durch die zitierfähige Publikation von Forschungsdaten wird der Grundstein für aufbauende Arbeiten gelegt und ein Beitrag zu einer insgesamt effizienteren Wissenschaft geleistet. Nach den FAIR-Prinzipien [2] (Findable, Accessible, Interoperable, Reusable. Auffindbar, Zugänglich, Interoperabel, Wiederverwendbar)

veröffentlichte und offen zugängliche Forschungsdaten können unter Betrachtung anderer Forschungsfragen die Grundlage neuer wissenschaftlicher Erkenntnisse sein.

Effizienz und Zeitersparnis im Alltag der Wissenschaft

Besonders in Projekten und Verbundvorhaben mit einer langen Laufzeit, mit vielen Kooperationspartnern, die an verteilten Standorten sind und häufig mit personeller Fluktuation einhergehen (als Beispiele seien die DAM-Forschungsmissionen und deren Verbünde genannt), behalten Wissenschaftler:innen durch FDM den Überblick über die eigenen sowie die kollaborativ zu bearbeitenden Forschungsdaten. FDM spart mittelfristig Zeit im Arbeitsalltag, bei der ortsunabhängigen Zusammenarbeit und bei der Einarbeitung neuer Kolleg:innen.

Wissenschaftler:innen verbringen viel Zeit mit der Kontrolle, Aufbereitung und Homogenisierung von einzelnen Datensätzen, bevor sie diese für Analysen nachnutzen können. Umso mehr Zeit wird benötigt, wenn große heterogene und mangelhaft beschriebene Datenmengen mit modernen Methoden (Künstliche Intelligenz [KI], Maschinelles Lernen) analysiert werden sollen. FDM ist die Basis für effiziente, interdisziplinäre und in der Breite durchgeführte Datenwissenschaft. Zusätzlich ermöglicht eine gute Organisation von Daten, Datenlücken aufzudecken die im Rahmen zukünftiger Projekte oder Messkampagnen geschlossen werden können.

Datensicherheit und Nachvollziehbarkeit

Die systematische Archivierung von Forschungsdaten minimiert das Risiko des Datenverlusts. Sensible Daten können im Rahmen des FDM geschützt werden. Zudem ermöglicht die Organisation von Forschungsdaten Projektbeteiligten und Dritten, Forschungsergebnisse besser nachzuvollziehen, diese zu reproduzieren oder eigene Ergebnisse zu validieren.

Bessere Chancen auf Forschungsförderung

Die Empfehlungen der DFG sehen vor, dass Daten, die im Rahmen von Forschungsvorhaben erhoben, bearbeitet oder erschlossen werden, mindestens zehn Jahre zugänglich sein sollen. Umso wichtiger ist es, bereits bei der Antragstellung ein vorausschauendes FDM zu erwägen und einen Datenmanagementplan zu erstellen. Mit einem durchdachten Forschungsdatenmanagementplan erhöhen sich die Chancen auf Förderung, da Förderorganisationen ein Konzept für das Management von Forschungsdaten erwarten. So soll sichergestellt werden, dass Wissenschaft und Gesellschaft gemäß den Ideen hinter Open Access bzw. Open Data von den

eingesetzten Projektmitteln bestmöglich profitieren, indem Forschungsdaten frei und ohne Einschränkungen für jeden zugänglich sind.

Das vorliegende Konzept dient der Unterstützung der Wissenschaftler:innen der DAM-Mitgliedseinrichtungen zur Erfüllung der expliziten Anforderungen, die im Forschungsdatenmanagement insbesondere in Großprojekten wie den DAM-Forschungsmissionen an sie gestellt werden.

3. Ausgangslage

Die Deutsche Allianz Meeresforschung (DAM) will Forschungsdatenmanagement im Bereich Meeresforschung, unter Berücksichtigung bestehender und in Entwicklung befindlicher Forschungsdaten-Infrastrukturen, koordinieren, um so Innovation und Datenwissenschaften zu befördern. Forschungsdatenmanagement muss als Basis zur Anwendung moderner Analysemethoden (KI, Maschinelles Lernen) verstanden werden.

Forschungsdatenmanagement geht einher mit der Nutzung von Forschungsdateninfrastruktur, die die Wissenschaft in ihren Forschungsdatenmanagementaufgaben unterstützen soll. Um eine vollständige und nachvollziehbare Dokumentation der vielfältigen Forschungsdaten in der Zukunft zu erreichen, sollten Wissenschaftler:innen in allen Phasen des Datenlebenszyklus (Projektplanung, Datenaufnahme, -aufbereitung, -analyse, -archivierung, -publikation) durch Datenmanagementexpert:innen unterstützt werden. Hierfür sind bestenfalls an den jeweiligen Mitgliedseinrichtungen institutionelle Datenmanagementsysteme etabliert, die die notwendigen Infrastrukturlösungen betreiben. Sie unterstützen beispielsweise bei der Bereitstellung von Speicherlösungen, der Definition und Durchführung von Dataflow- und Workflows (z.B. für DOI) oder der Erstellung standardisierter Dienste.

Die bestehenden Voraussetzungen, Organisationsstrukturen und Infrastrukturen für eine solche Unterstützung sind an den DAM-Mitgliedseinrichtungen sehr heterogen. Die Umsetzung des vorliegenden Konzeptes zielt darauf ab, diese Heterogenität auszugleichen. Das Konzept skizziert den Weg zu einer Homogenisierung des Managements von Forschungsdaten und einer zwischen den DAM-Mitgliedern abgestimmten Vorgehensweise und Infrastruktur, die die Voraussetzungen für die Publikation FAIRer Datensätze schafft, im Einklang mit den Zielen der Nationalen Forschungsdateninfrastruktur (NFDI). Die NFDI soll die nationalen Datenbestände von Wissenschaft und Forschung systematisch erschließen, nachhaltig sichern und zugänglich machen sowie (inter-)national mit der European Open Science Cloud (EOSC)

vernetzen. Sie verfolgt somit kongruente Ziele zur DAM im Kernbereich „Datenmanagement und Digitalisierung“.

Die NFDI wird seit dem 01. Oktober 2020 in einem wissenschaftsgetriebenen Prozess als vernetzte Struktur eigeninitiativ agierender Konsortien aufgebaut. Die DAM übernimmt eine koordinierende Funktion für die Meeresforschung im bewilligten NFDI-Konsortium für Biodiversitätsforschung (NFDI4Biodiversity) und für Erdsystemforschung (NFDI4Earth). Damit unterstützt der DAM-Kernbereich „Datenmanagement und Digitalisierung“ den Aufbau einer übergreifenden, dezentralen und nachhaltig betriebenen Forschungsdateninfrastruktur für die nationale Meeresforschung auch in einem übergeordneten Rahmen. Über die Beteiligung an den Querstrukturen der NFDI, wie der Sektion Common Infrastructures oder Metadata, werden Standards und Architekturen mit der gesamten Gemeinschaft der Forschenden abgestimmt und geteilt.

4. Konzeptionelle Überlegungen für den Aufbau eines kollaborativen Datenökosystems

Als Datenökosystem wird das prägende, ganzheitliche Umfeld verstanden, in dem verschiedene Akteure zusammenkommen, um Daten zu produzieren, anzubieten, zu finden, nachzunutzen, zu verarbeiten, anzureichern, zu archivieren, zu publizieren und Entscheidungen darauf zu fällen. Innerhalb der DAM sollte dafür 1. eine Vereinheitlichung im Umgang mit Forschungsdaten stattfinden und 2. eine gemeinsame Forschungsdateninfrastruktur entstehen, die sich aus vorhandenen und in Entwicklung befindlichen Komponenten zusammensetzt. Dies setzt voraus, dass die DAM-Mitgliedseinrichtungen bestimmte Grundvoraussetzungen an den jeweiligen Einrichtungen schaffen - sofern sie nicht bereits bestehen - (siehe 4.1) und ein Nutzungs- und Betriebskonzept für die Dateninfrastruktur der deutschen Meeresforschung beschlossen wird (siehe 4.3).

Es wird außerdem vorgeschlagen, dass ein „Datenmanagement Help-Desk“ zum Ausgleich der derzeit heterogenen Unterstützungsmöglichkeiten für die Wissenschaft unter den DAM-Mitgliedseinrichtungen und als Bestandteil des „User Support Network (USN)“ der NFDI4Earth eingerichtet wird (siehe 4.2).

Die DAM sollte die dafür notwendigen Prozesse koordinieren (siehe 4.4).

4.1. Voraussetzungen bei den DAM-Mitgliedseinrichtungen:

Formale Basis

Um die Basis für ein zukünftiges gemeinsames Vorgehen im Datenmanagement zu schaffen, ohne dass die DAM in die institutionelle Souveränität ihrer Mitglieder im Umgang mit Forschungsdaten eingreift, sollten die DAM-Mitgliedseinrichtungen -sofern

als sinnvoll und umsetzbar erachtet- jeweils institutionelle Forschungsdaten-Leitlinien sowie Datenmanagement-Konzepte innerhalb ihrer Einrichtungen verabschieden, die mit der bestehenden DAM-Forschungsdaten-Leitlinie konform sind. Ein erster wichtiger Schritt in Richtung kollaboratives Datenökosystem war die Zustimmung der Mitgliedseinrichtungen zur DAM-Forschungsdaten-Leitlinie im Mai 2021. Diese wurde von der DAM-AG „Datenmanagement und Digitalisierung“ erstellt. Sie kann als „kleinster gemeinsamer Nenner“ der DAM-Mitgliedeinrichtungen im Umgang mit Daten betrachtet werden.

Institutionelle Forschungsdaten-Leitlinien können und sollten womöglich spezifischer sein. Beispiele für spezifischere Leitlinien sind die Forschungsdaten-Leitlinie des [AWI](#), [3] [GEOMAR](#) [4] und des [MARUM](#) [5].

Eine Forschungsdatenleitlinie regelt Grundsätze für den einheitlichen Umgang mit Forschungsdaten. Dies betrifft:

- den offenen Zugang zu Daten und die Einhaltung der FAIR-Prinzipien,
- die Anerkennung von qualitätsgesicherten Datenpublikationen als wissenschaftliche Leistung,
- Kriterien der Qualitätssicherung der Daten,
- die Verantwortlichkeiten für die Einhaltung der Kriterien und den sorgfältigen Umgang mit Daten,
- Fristen für die Archivierung von Primärdaten und die Dauer von Moratorien,
- die Dauer der Archivierung der Daten,
- die Bereitschaft zur langfristigen Vernetzung der Dateninfrastrukturen.

Forschungsdatenleitlinien berücksichtigen den Kodex zur Sicherung der guten wissenschaftlichen Praxis der Deutschen Forschungsgemeinschaft und konkretisieren dessen fachspezifische Umsetzung. Zur Steigerung der Reproduzierbarkeit und Datenqualität sollen im Rahmen der Umsetzung der FAIR-Prinzipien, die Provenienz der Daten bei der Datenerhebung, -sammlung, -nutzung und -aufbewahrung besonders berücksichtigt werden. Projektverantwortliche sollten bestärkt und unterstützt werden, im Projekt gewonnene Daten spätestens mit Projektabschluss frei zugänglich zu machen.

Datenmanagementkonzept/Richtlinien: DAM-Mitgliedseinrichtungen sollten Datenmanagementkonzepte bzw. Richtlinien auf institutioneller Ebene erarbeiten, die der Wissenschaft vor Ort helfen, Forschungsdatenleitlinien unter Nutzung der jeweils vorhandenen Infrastrukturen einzuhalten. Die Nutzung von Infrastrukturen anderer Mitgliedseinrichtungen sollte technisch und administrativ ermöglicht werden (siehe Abschnitt „Gemeinsames Nutzungs- und Betriebsmodell für Dateninfrastrukturen“),

damit dies in Datenmanagementkonzepten einzelner Einrichtungen berücksichtigt werden kann.

Datenmanagementkonzepte beschreiben das Regelwerk zur wissenschaftlichen Datenverarbeitung und Datenmanagement. Es umfasst folgende 4 Phasen:

1. Aufnahme der Daten bzw. Produktion von numerischen Simulationsdaten inkl. der Metadaten zur Einhaltung der FAIR-Prinzipien.
2. Die Auswertung, Bearbeitung und Validierung der Daten im Zusammenhang mit der wissenschaftlichen Fragestellung.
3. Die Publikation wissenschaftlicher Arbeiten und die allgemeine Verfügbarkeit erhobenen Daten.
4. Langzeitarchivierung der Daten.

Um eine vollständige und nachvollziehbare Dokumentation der Forschungsdaten zu erzeugen, müssen Wissenschaftler:innen und Datenmanager:innen in diesen Phasen intensiv zusammenarbeiten.

4.2. Datenmanagement-Helpdesk

Zum Ausgleich der derzeit heterogenen Unterstützungsmöglichkeiten für die Wissenschaft unter den DAM-Mitgliedseinrichtungen, sollte ein "Datenmanagement Help-Desk" als Bestandteil des "User Support Network" der NFDI4Earth [6] eingerichtet werden. Das Konsortium NFDI4Earth der Nationalen Forschungsdateninfrastruktur baut ein verteiltes, institutionsübergreifendes "User Support Network" auf, das sich auf die bestehenden und gut eingebetteten Nutzerunterstützungsstrukturen der beteiligten Institutionen stützt. Ein unter dem Dach der DAM eingerichteter Datenmanagement-Helpdesk stünde der gesamten deutschen Meeresforschung als Serviceeinheit zur Verfügung, wäre spezialisiert auf die Belange der Meeresforschung und schlosse sich dem Netzwerk der NFDI4Earth an. Aufgrund von Erfahrungen in einzelnen DAM-Mitgliedseinrichtungen sollte der Datenmanagement-Helpdesk personell aus vier VZÄ bestehen, um die Unterstützung der Wissenschaft gewährleisten zu können. Aufgaben des Help-Desk sollten in einer Aufgabenbeschreibung ("Terms Of Reference") in Abstimmung mit den Mitgliedseinrichtungen festgehalten werden, um Ambiguität hinsichtlich der Zuständigkeiten zu vermeiden. Folgende Aufgaben könnten beim Datenmanagement Help-Desk liegen:

- Unterstützung der Wissenschaft entlang des Datenlebenszyklus (bzw. der vier Phasen) **durch Nutzung vorhandener Dateninfrastrukturen** (z.B. sensor.awi.de, OSIS (Ocean Science Information System, GEOMAR) (siehe auch folgende Abschnitte).

- Weiterbildung von Datenmanagementbeauftragten in Verbundprojekten
- Koordination der Anbindung an nationale und internationale Datenportale
- Angebot von Workshops zum Thema Datenmanagement mit zusätzlichem Anschluss an die Sektion [6] [“NFDI Cross-cutting Topic: Education And Training”](#)
- Enge Zusammenarbeit mit den zertifizierten Repositorien PANGAEA und WDCC (World Data Center for Climate)
- Die Bereitstellung von Daten zur Erstellung von aggregierten Datenprodukten, die die Wissenschaft und/oder Politik zur weiteren Analyse beantragt.
- Die Bearbeitung spezifischer und zeitlich begrenzter Aufgaben z.B. zur Aufarbeitung und Bereitstellung historischer Daten.

4.3. Gemeinsames Nutzungs- und Betriebsmodell für Dateninfrastrukturen

Im Sinne einer nationalen Infrastruktur sollten Nutzungs- und Betriebsmodelle für vorhandene Dateninfrastrukturen entwickelt werden. Ziel ist es, der Forschung den optimalen Zugang zu den Dateninfrastrukturen zu ermöglichen, dafür einheitliche Rahmenbedingungen zu schaffen und die Dateninfrastruktur im Sinne einer nationalen Infrastruktur weiterzuentwickeln und zu betreiben, wobei die Weiterentwicklung und der Betrieb der Infrastrukturen Aufgabe der betreibenden Einrichtungen bleibt.

Beispielhaft für vorhandene Dienste, die von der gesamten Gemeinschaft genutzt werden sollten, können die Dateninfrastrukturen sensor.awi.de (AWI) und OSIS (Ocean Science Information System, GEOMAR) genannt werden. Sensor.awi.de dient der Erstellung und Verwaltung von plattform-, geräte- und sensorbezogenen Informationen, die für die FAIRness von Datenpublikationen z.B. in PANGAEA essentiell sind. OSIS ermöglicht Wissenschaftler:innen den projektinternen Austausch von Daten, die im Rahmen von Expeditionen, aus numerischen Modellen oder Experimenten generiert werden. OSIS kommt im Rahmen der DAM-Forschungsmissionen zum Einsatz, bislang ausschließlich in Verbünden mit GEOMAR Beteiligung.

Personen im Datenmanagement-Helpdesk sollten z.B. die Kuration von Sensor-Daten in sensor.awi.de projektbezogen, als Service für die Wissenschaft und in enger Absprache mit der Wissenschaft koordinieren. Dies würde für eine Homogenität der Dateneinträge sorgen. Sensor-Information könnte so standardmäßig an Datenpublikationen z.B. in PANGAEA als Metadaten angehängt werden. Auch die Nutzung anderer Infrastrukturen, wie OSIS, braucht unterstützendes Personal, das im Datenmanagement-Helpdesk der DAM ansprechbar, geschult und kompetent zur Unterstützung der Wissenschaft bereitsteht.

Institutionelle Datenarchive werden an das Portal Deutsche Meeresforschung angeschlossen, um die Auffindbarkeit der archivierten Daten zu verbessern. Die dafür

notwendigen technischen Voraussetzungen werden von den jeweiligen DAM-Mitgliedseinrichtungen in Rücksprache mit Entwickler:innen der portalbetreibenden Einrichtung (AWI) implementiert. Die Nutzung standardisierter Schnittstellen und gemeinsamer Metadatenprofile fördern die Auffindbarkeit von Daten innerhalb des Portals Deutsche Meeresforschung. Es sei an dieser Stelle darauf hingewiesen, dass die Zusammenarbeit mit Betreibern der Marinen Dateninfrastruktur Deutschland (MDI.de) und der Meeresumweltdatenbank (MUDAB) ausgebaut werden soll. Die Einbindung spezifischer Dienste in das Datenportal Deutsche Meeresforschung ist bereits über Schnittstellen möglich.

4.4. DAM als Koordinationsplattform

Die Erstellung von institutseigenen Forschungsdatenleitlinien und Datenmanagementkonzepten soll über die DAM-AG "Datenmanagement und Digitalisierung" unterstützt werden. Dafür müssen AG-Mitglieder die Verantwortung und Koordination zur Erstellung und Abstimmung der Dokumente innerhalb ihrer Einrichtungen übernehmen können und hierfür entsprechend von den jeweiligen Hausleitungen mandatiert werden. Die o.g. Dokumente (Forschungsdatenleitlinien, Datenmanagement-Konzepte), die an DAM-Mitgliedseinrichtungen bereits vorliegen, können als Vorlagen dienen.

Die Einrichtung eines Datenmanagement-Helpdesks sollte von den o.g. Forschungsdatenleitlinien und institutionellen Datenmanagementkonzepten unabhängig sein, da die Aufgaben des Helpdesks davon z.T. unabhängig sind. Die Ansiedlung des Datenmanagement-Helpdesks innerhalb des DAM-Kernbereichs würde sich anbieten, um den zentralen Zugang für alle Wissenschaftler:innen aus DAM-Mitgliedseinrichtungen zu unterstreichen. Die Personen könnten entweder von DAM-Mitgliedseinrichtungen entsandt oder neu eingestellt werden. Die Steuerung des entstehenden Teams würde von der Leitung des DAM-Kernbereichs "Datenmanagement und Digitalisierung" gemeinsam mit dem für den Kernbereich verantwortlichen Vorstandsmitglied übernommen. Unter dem Dach der DAM würde für die Vernetzung des Datenmanagement-Helpdesk mit der DAM-AG "Datenmanagement und Digitalisierung", den Datenmanagement-Beauftragten sowie insbesondere dem zentralen Datenmanagement der DAM-Forschungsmissionen gesorgt. Darüber hinaus würde die Anbindung an das "User Support Network" der NFDI4Earth und die Beteiligung an [NFDI-Sektionen](#) eingerichtet.

Die Erstellung eines Nutzungs- und Betriebskonzeptes für die gemeinsame Nutzung bestehender Dateninfrastruktur würde über den DAM-Kernbereich „Datenmanagement und Digitalisierung“ koordiniert. Die Verständigung darüber, welche bestehenden Infrastrukturen neben sensor.awi.de und OSIS für die gesamte Gemeinschaft zugänglich

sein sollen, würde innerhalb der DAM-AG "Datenmanagement und Digitalisierung" geführt. Durch die Benennung weiterer Expert:innen der betroffenen Einrichtungen, könnte ein gut arbeitsfähiger Personenkreis bestimmt werden, der die Ausarbeitung eines Konzeptes vornimmt. Hier würden vermutlich auch die Verwaltungen um Zuarbeit gebeten werden müssen.

5. Datenmanagementkonzept in der Praxis

Unter der Annahme, dass vorhandene und in Entwicklung befindliche Forschungsdateninfrastrukturen gemeinschaftlich und institutsunabhängig genutzt werden können, beschreiben die folgenden Abschnitte wie die Umsetzung des Datenmanagements innerhalb von Projekten geregelt sein könnte.

5.1. Datenmanagement in Projekten von Personen

Die Datenmanagementanforderungen an Wissenschaftler:innen, die sich z.B. in einem DFG geförderten Projekt auf "Eigene Stelle", "Emmy-Noether-Programm" oder "Heisenberg-Programm" befinden, sollten in o.g. einrichtungsbezogenen Forschungsdatenleitlinien festgeschrieben sein. Das Datenmanagement-Helpdesk kann bei der Erstellung von Datenmanagementplänen in der Antragsphase, sowie der Aktualisierung der Pläne auch nach der Bewilligung, beansprucht werden.

5.2. Datenmanagement in Verbundprojekten von Organisationen

Bei Antragstellung von Verbundprojekten sollten zwei Dokumente, im Verbund abgestimmt, erstellt werden:

1. Eine Datenleitlinie. Als Vorlage können die bestehende DAM-Forschungsdatenleitlinie oder die Leitlinien der DAM-Forschungsmissionen genutzt werden.
2. Ein oder mehrere Datenmanagementpläne. Auch hier liegen Vorlagen vor, die im Rahmen der DAM-Forschungsmissionen erarbeitet wurden.

Auf Anfrage der Verbundkoordination würde das Datenmanagement-Helpdesk die Koordination dieses Abstimmungsprozesses in Absprache mit Datenmanagementbeauftragten eines Verbundes begleiten. Dokumentvorlagen, Beratungsangebote sowie Koordinationsunterstützung würden über das Datenmanagement-Helpdesk bereitgestellt. Die Voraussetzung dafür ist, dass ein:e Datenmanager:in für ein Verbundprojekt als Ansprechpartner:in zur Verfügung steht. Die Daten-Leitlinie gleicht einer Vereinbarung zum Umgang mit Forschungsdaten, die im Verbundprojekt erhoben werden und regelt gleichzeitig die Zuständigkeiten. Die Erstellung dieser Dokumente muss koordiniert werden. Die Pflege des

Datenmanagementplans muss fortlaufend durchgeführt werden und könnte pro Verbundprojekt zentral vom Datenmanagement-Helpdesk unterstützt werden.

6. "Unterwegs"-Forschungsdaten

Daten, die auf Expeditionen gesammelt werden, sind in der Meeresforschung i.d.R. Teil von „Projekten von Personen“ oder „Verbundprojekten“. Da die Schiffe aber mit vielen Sensoren unterschiedlicher wissenschaftlicher Disziplinen ausgestattet sind, die z.T. permanent laufen und nicht zwangsläufig und ausschließlich im Sinne des Projektes/ der Primärforschung Daten generieren, entstehen zusätzlich wertvolle Daten, die durch die o.g. Überlegungen nicht abgedeckt sind bzw. gehandhabt werden. Hier schafft das DAM-Projekt "Unterwegs"-Forschungsdaten Abhilfe und baut *en passant* eine Dateninfrastruktur an Bord der großen Forschungsschiffe auf, die für jegliche Expeditionsdaten nutzbar ist.

Expeditionsdaten, die im Rahmen eines Projektes erhoben werden, wären von den o.g. Überlegungen abgedeckt. Ein Datenmanagementplan wird im Rahmen von Schiffsanträgen gefordert.

Expeditionsdaten werden hier dennoch gesondert betrachtet, weil Expeditionsdaten systematisch, umfänglich (Primärforschungsdaten und "Unterwegs"-Daten), z.T. in nahezu Echtzeit und/oder möglichst kurz nach Expeditionsende, FAIR und offen bereitgestellt werden sollten. Dafür gibt es mehrere Gründe:

- Expeditionsdaten gehören zu den - monetär und wissenschaftlich betrachtet- wertvollsten und einzigartigen Daten in der Meeresforschung. Anders als in typischen Laborexperimenten oder Modellrechnungen, gibt es keine Chance auf Wiederholung.
- Die Kosten für die sichere „Anlandung“ und langfristige, FAIRe und offenen Bereitstellung der Daten sind im Verhältnis zu den Kosten für die Erhebung gering (etwa 3% bezogen auf die Schiffshaushalte der großen Deutschen Forschungsschiffe, ausgenommen sind die Kosten für die Geräte mit denen die Daten erhoben werden). Das Risiko, Expeditionsdaten „nur“ partiell (z.B. nur Teile der Primärforschungsdaten) und als „long tail output“ (im Zusammenhang mit der Publikation eines wissenschaftlichen Artikels) FAIR bereitzustellen oder gar zu verlieren, sollte nicht eingegangen werden. Ferner sollte Datenmanagement als integraler Bestandteil von Schiffsexpeditionen betrachtet und die Kosten eingerechnet werden, um Expeditionsdaten umfänglich und langfristig bereitzustellen.
- Expeditionsdaten werden z.T. für operationelle Zwecke wie Wettervorhersagen genutzt und müssen daher in nahezu Echtzeit zur Verfügung stehen. Diese

Anwendung würde wegfallen, wenn lediglich einige „long-tail“-Daten bis zu Jahren nach einer Expedition über o.g. Prozess FAIR und offen bereitgestellt würden.

- Der „long-tail“ Prozess umfasst die Expeditionsdaten nicht umfänglich.

Das DAM-Projekt “Unterwegs“-Forschungsdaten hat den Status eines konkreten Vorzeigebispiels für die institutionen- und flottenübergreifende Bereitstellung der “Unterwegs“-Forschungsdaten großer deutscher Forschungsschiffe nach den *FAIR- and Open Data*-Prinzipien erreicht. In einer nächsten Phase werden die Aktivitäten auf möglichst viele Expeditionsdaten von großen und mittelgroßen deutschen Forschungsschiffen ausgeweitet. Es soll zudem ermittelt werden inwieweit das System OSIS gemeinschaftlich und für das Datenmanagement der Expeditionsdaten genutzt werden kann.

7. Referenzen

- [1] Deutsche Forschungsgemeinschaft. (2019). Guidelines for Safeguarding Good Research Practice. Code of Conduct. <https://doi.org/10.5281/zenodo.3923602>
- [2] Wilkinson, M.D., Dumontier, M., Jan Aalbersberg, I. et al. Addendum: The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 6, 6 (2019). <https://doi.org/10.1038/s41597-019-0009-6>
- [3] AWI Directorate (2020): Forschungsdaten-Leitlinie des Alfred-Wegener-Instituts, Helmholtz-Zentrum für Polar- und Meeresforschung [Miscellaneous]
- [4] GEOMAR Direktorium, (2022) Forschungsdaten-Leitlinie. GEOMAR Helmholtz-Zentrum für Ozeanforschung Kiel. doi: 10.3289/data_policy_V1_2022.
- [5] Glöckner, Frank Oliver, Kucera, Michal, Pälike, Heiko, Zabel, Matthias, & Schulz, Michael. (2021). Research Data Policy at MARUM – Center for Marine Environmental Sciences of the University of Bremen (1.0). Zenodo. <https://doi.org/10.5281/zenodo.5643724>
- [6] Bernard, Lars, Braesicke, Peter, Bertelmann, Roland, Frickenhaus, Stephan, Gödde, Hildegard, Keßler, Carsten, Lorenz, Sören, Mahecha, Miguel, Marschall, Horst, Hezel, Dominik, Nagel, Wolfgang E., Reichstein, Markus, Sester, Monika, Thiemann, Hannes, Weiland, Claus, Wytzisk-Arens, Andreas, & NFDI Consortium Earth System Sciences.

(2021). NFDI Consortium Earth System Sciences - Proposal 2020 revised (Revision 1 from 2021-10-01). Zenodo. <https://doi.org/10.5281/zenodo.5718944>

[7] Ebert, Barbara, Fluck, Juliane, Glöckner, Frank Oliver, Koepler, Oliver, Miller, Bernhard, Schmitt, Robert, Schrade, Torsten, Stegle, Oliver, Steinbeck, Christoph, von Suchodoletz, Dirk, Wagemann, Kurt, Knebes, Jennifer, Kraft, Sophie, Seitz-Moskaliuk, Hendrik, Sure-Vetter, York, & Wössner, Elena. (2021). NFDI Cross-cutting Topics Workshop Report. Zenodo. <https://doi.org/10.5281/zenodo.4593770>