

Antrag

für die Anschubphase des Aufbaus einer vernetzten, dezentral operierten Dateninfrastruktur für Meeresforschungsdaten unter dem Dach der Deutschen Allianz Meeresforschung e.V. (in Gründung) (DAM)

Projektlaufzeit: 01.07.2019 bis 31.12.2020

1. Hintergrund

Um den Herausforderungen in der Erforschung der Meere und Ozeane zu begegnen, sollen die wissenschaftlichen und technischen Kapazitäten der deutschen Meeresforschung in einem international herausragenden Verbund von Expertisen und Institutionen gebündelt werden (Deutsche Allianz Meeresforschung e.V. – DAM). Die Allianz will den großen Zukunftsfragen der Meeresforschung auf höchstem Niveau begegnen und den wissenschaftlichen Beitrag zu Handlungswissen für Politik, Wirtschaft und Zivilgesellschaft verstärken. Die DAM wird dabei Aktivitäten in vier Kernbereichen - Forschung, Infrastruktur, Transfer sowie Datenmanagement und Digitalisierung - umsetzen. Die Gründung des gemeinnützigen Vereins (Deutsche Allianz Meeresforschung e.V.) ist für 2019 geplant, initiiert vom Bund und den fünf norddeutschen Bundesländern (Freie Hansestadt Bremen, Freie und Hansestadt Hamburg, Mecklenburg-Vorpommern, Niedersachsen und Schleswig-Holstein).

Angesichts der rapide fortschreitenden Digitalisierung in der internationalen Meeresforschung ist der DAM Kernbereich „Datenmanagement und Digitalisierung“ von herausgehobener Bedeutung und Dringlichkeit für die deutsche Meeresforschung. Der offene und nachhaltige Zugang zu Daten der Meeresforschung ist ein Kernelement der geplanten DAM Aktivitäten. Langfristig soll der Kernbereich „Datenmanagement und Digitalisierung“ einen wichtigen Baustein der Nationalen Forschungsdaten Infrastruktur für Erdsystemforschung (NFDI4Earth) bilden. Im Fokus der DAM - Dateninfrastruktur stehen dabei die Erstellung von Datenprodukten, mit den entsprechenden Metadaten, um die Qualität einzuschätzen, ihre Offenheit sowie ihre (Rück-)Verfolgbarkeit und Nutzbarkeit in mehrdimensionalen Kontexten (FAIR Prinzipien).

Antragsteller sind das Alfred-Wegener-Institut Helmholtz-Zentrum für Polar- und Meeresforschung (AWI) und das GEOMAR Helmholtz-Zentrum für Ozeanforschung Kiel (GEOMAR). Projektpartner sind das Helmholtz- Zentrum Geesthacht Zentrum für Material- und Küstenforschung (HZG), das Centrum für Erdsystemforschung und Nachhaltigkeit der Universität Hamburg (CEN), die Christian-Albrecht Universität zu Kiel (CAU), das Institut für Chemie und Biologie des Meeres der Universität Oldenburg (ICBM), das Leibniz-Institut für Ostseeforschung Warnemünde (IOW) und das MARUM - Zentrum für Umweltwissenschaften der Universität Bremen. Im Falle der Bewilligung dieses Antrags verpflichten sich die antragstellenden Einrichtungen (AWI und GEOMAR) i) zu einer anteiligen Weiterleitung der Mittel an die DAM-Partner (CAU, CEN, ICBM, IOW und MARUM) für deren konkrete Beiträge zur zu entwickelnden gemeinsamen Infrastruktur und ii) zur Durchführung des Controlling. Die anteilige Weiterleitung der Mittel ist gemäß Abschnitt 4.4 zu leisten. Mittel für den Beitrag des HZG müssen direkt vom Bundesministerium für Bildung und Forschung (BMBF) an das Helmholtz-Zentrum übermittelt werden. Die Projektpartner im Kontext des hier beantragten Pilotprojektes 'Unterwegs'-Forschungsdaten des DAM Kernbereichs „Datenmanagement und Digitalisierung“

lisierung“ und im Rahmen der Anschubphase, erfüllen die Voraussetzung, einen nennenswerten Beitrag zum Erreichen der für die Anschubphase gesetzten Ziele, leisten zu können. Eine detaillierte und eine Kurzfassung der Liste mit Partner und Aufgaben sowie befinden sich im Anhang [i] und unter Punkt 4.4.

2. DAM und NFDI

Die Nationale Forschungsdateninfrastruktur (NFDI) ist dazu konzipiert, disziplinübergreifend Datenprodukte und Datendienste anzubieten [1]. Im Gegensatz dazu zielt der Kernbereich „Datenmanagement und Digitalisierung“ der DAM in einem Pilotprojekt darauf ab, in Bezug auf die ‘Unterwegs’-Forschungsdaten der großen Forschungsschiffe den ersten Schritt zu leisten, nämlich Daten vom Ort und Zeitpunkt der Erhebung, koordiniert und kuratiert in das Archiv PANGAEA (und ggf. weitere) zu transferieren. So wird eine Voraussetzung für das Erzeugen nachnutzbarer Datenprodukten in der NFDI geschaffen.

Als ‘Unterwegs’-Forschungsdaten werden hier jene Daten bezeichnet, die von bordeigenen, meereswissenschaftlich relevanten Sensoren gemessen werden. Der DAM Kernbereich „Datenmanagement und Digitalisierung“ sieht seine langfristige Aufgabe in der Entwicklung und Implementierung verbindlicher, nachhaltiger und harmonisierter Arbeitsabläufe innerhalb der Meeresforschungsgemeinschaft ggf. auch über ‘Unterwegs’-Forschungsdaten und hier beantragten Vorhaben hinaus. Basierend auf einheitlichen Standards sollen möglichst vollständig beschriebene, kuratierte, archivierte, publizierte und zu Datenprodukten verarbeitbare Daten für die NFDI und insbesondere für die Erdsystem [2]- und Biodiversitätsforschung [3] angeboten werden. Ziel ist es, ein starker und kompetenter Datenprovider hochwertiger und harmonisierter Daten für die NFDI zu sein.

Als Beitrag zur Agenda der NFDI liegt daher besonderes Augenmerk auf:

1. Der Entwicklung einer gemeinsamen Konzeption und Architektur für den DAM Kernbereich „Datenmanagement und Digitalisierung“.
2. Der technischen Weiterentwicklung und Etablierung nutzbarer, standardisierter Forschungsdatenworkflows und vertrauenswürdiger Datenservices.
3. Der nachhaltigen personellen Ausstattung der Institutionen mit Entwickler*innen und Datenspezialist*innen um die Dienstleistungen nachhaltig anbieten zu können.
4. Der Qualifizierung der Forschenden im Bereich Datenwissenschaft.
5. Die Einbettung der Maßnahmen in die internationale Forschungslandschaft, speziell die European Open Science Cloud (EOSC) und die „FAIR“ Initiativen.

Etablierte Dateninfrastrukturen und Arbeitsabläufe sollen dabei berücksichtigt, vereinheitlicht, integriert und weiterentwickelt werden. Die Entwicklung im Rahmen der DAM soll dabei auf der Kooperation aller relevanten Akteure aufbauen: von Universitäten über außeruniversitäre Forschungseinrichtungen bis hin zu Ressortforschungseinrichtungen. Die hier beantragte Anschubfinanzierung gilt dem Einstieg in nachhaltiges Datenmanagement und der effektiven Digitalisierung der ‘Unterwegs’-Daten der deutschen Forschungsschiffe, mit der Priorität auf die vier großen Schiffe der deutschen Meeresforschung (MARIA S. MERIAN, METEOR, SONNE, POLARSTERN).

3. Status Quo

Der Aufbau von ersten wichtigen technischen und organisatorischen Strukturen für den nachhaltigen Zugang von Meeresforschungsdaten erfolgte über das HGF Projekt „Marine Network for Integrated Data Access“ (MaNIDA, 2011-2014) aus dem das „Datenportal Deutsche Meeresforschung“ hervorging. Über das „Datenportal Deutsche Meeresforschung“ könnten mit Hilfe der beantragten Mittel meeresbezogene Forschungsdaten in Zukunft nachhaltig auffindbar und zugreifbar sein, um letztlich der Politik, Wirtschaft und Zivilgesellschaft zugute zu kommen.

Im Jahr 2015 wurde im Kontext des Projektes MaNIDA eine Recherche zur Auffindbarkeit von ‚Unterwegs‘-Forschungsdaten über das „Datenportal Deutsche Meeresforschung“ durchgeführt. Das Ergebnis zeigte, dass nur ein Bruchteil der ‚Unterwegs‘-Forschungsdaten (~ 25 %) dort zentral auffindbar, zugreifbar und zitierbar war. Die maßgeblichen Datenanbieter des Portals waren zu dem Zeitpunkt die Datenarchive PANGAEA (AWI/MARUM) und DOD (BSH). Während mehr als 70 % der Messereignisse der ‚Unterwegs‘-Forschungsdaten der POLARSTERN zu nachnutzbaren und zugänglichen Daten führten, lag der Anteil bei den anderen deutschen Forschungsschiffen deutlich niedriger (zwischen 8 und 20 %). Unabhängig davon variierte die Datenverfügbarkeit stark zwischen den wissenschaftlichen Disziplinen. Die Recherche hat ergeben, dass die Datenverfügbarkeit, die aus Geräteeinsätzen der Atmosphären-Forschung resultiert, bei ~ 96 % lag. In der Ozeanographie waren es ~ 40 %, in anderen Disziplinen tendieren die Verfügbarkeiten teilweise gegen Null. Zusammenfassend lässt sich sagen, dass im Jahr 2015 validierte Daten der Ozeanographie, der Atmosphärenforschung und selbst die Schiffsrouten größtenteils nur von der POLARSTERN zur Verfügung standen, während nur wenig Daten weiterer Schiffe auffindbar waren. Die Lage hat sich durch die Eigenleistung einzelner Forschungseinrichtungen etwas verbessert, aber eine in der Zwischenzeit substantielle Steigerung der Datenverfügbarkeit war und ist unmöglich, da die nötigen Mittel zum Aufbau nachhaltiger Datenströme von den Sensoren zu den Archiven nicht zur Verfügung standen und stehen.

Betrachtet man den Stand der Qualitätskontrolle und Datenverfügbarkeit von ‚Unterwegs‘-Forschungsdaten im internationalen Vergleich, so wird der Handlungsbedarf für die Deutsche Forschungsschiffflotte klar erkennbar. Die USA haben beispielsweise mit dem „Rolling Deck to Repository“ (R2R;[4]) ein System etabliert, das in der Umsetzung für das hier im Rahmen von DAM vorgeschlagene System als Beispiel genutzt werden soll. Es soll größter Nutzen aus existierenden Europäischen und internationalen Standards bezüglich der Qualitätskontrolle und Rohdatenprozessierung gezogen werden. Neben R2R werden auch die Arbeiten des Ship Observations Team (SOT) der World Meteorological Organization (WMO) bei der Standardisierung von Arbeitsabläufen, Datenströmen und Sensorkalibration [5] Berücksichtigung finden. Die Expertise für das hier vorgeschlagene Vorhaben ist in Deutschland durchaus vorhanden. Das Alfred-Wegener-Institut, Helmholtz-Zentrum für Polar- und Meeresforschung (AWI) ist am fortschrittlichsten im Umgang mit ‚Unterwegs‘-Forschungsdaten der POLARSTERN und bietet über das System „O2A – Observations to Archive“ [6], das konzeptionell ähnlich aufgebaut ist wie R2R, eine hervorragende Grundlage für die hier vorgeschlagenen Aktivitäten.

4. Anschubförderung

Für die Anschubphase wird hier der Ausbau des webbasierten „Datenportal Deutsche Meeresforschung“ und der dahinterliegenden Strukturen beantragt, um die nachhaltige Nutzung von ‚Unterwegs‘-Forschungsdaten der vier großen Forschungsschiffe (MARIA S. MERIAN, METEOR, SONNE, POLARSTERN) erheblich zu verbessern.

Um dieses Ziel zu erreichen, bedarf es der verbindlichen und nachhaltigen Implementierung harmonisierter Arbeitsabläufe. Die Implementierung bezieht sich auf Arbeitsabläufe sowohl organisatorischer, als auch technischer Art. Ein nahezu vollständiger semiautomatischer Datenfluss zur Archivierung und Publikation in den entsprechenden Datenrepositorien, auf die das „Datenportal Deutsche Meeresforschung“ zugreift, kann auf diese Weise aufgebaut werden. Eine maschinenlesbare Auffindbarkeit der schiffsbezogenen beschreibenden Daten (Metadaten), verknüpft mit den entsprechenden ‚Unterwegs‘- Forschungsdaten, wäre gewährleistet. Etablierte Arbeitsabläufe, Datenflüsse etc – kurz Ergebnisse der Anschubphase sollen auf weitere deutsche Forschungsschiffe und Daten weiterer Geräte und Sensoren einfach und schnell übertragbar sein, um die Nachnutzbarkeit qualitätskontrollierter Daten in die Breite zu bringen. Die Umsetzung der Übertragung hängt jedoch maßgeblich von der zukünftigen Finanzierung des DAM Kernbereichs „Datenmanagement und Digitalisierung“ ab:

Insgesamt wird mit der hier beantragten Ausstattung für den Kernbereich „Datenmanagement und Digitalisierung“ ein Grundstein für den Aufbau einer vernetzten, dezentral operierten Dateninfrastruktur für Meeresforschungsdaten unter dem Dach der DAM geschaffen. Die programmatischen und infrastrukturellen Weiterentwicklungen für die nachhaltige Digitalisierung in der Meeresforschung in Deutschland, die mit der Gründung der DAM angestrebt werden, erfordern jedoch eine langfristige, verlässliche Förderung in einem Umfang der weit über diesen Antrag auf Anschubfinanzierung hinausgeht. Diese weitreichende gemeinsame Zielsetzung vom Bund, den norddeutschen Ländern und den wissenschaftlichen Einrichtungen der Meeresforschung liegt diesem Antrag zugrunde.

4.1. Ziele und Meilensteine

Ziel der Anschubförderung ist eine nachhaltige, gemeinsame und international anschlussfähige Dateninfrastruktur zu initiieren, die eine schnelle, zentrale Verfügbarkeit der qualitätskontrollierten und harmonisierten ‚Unterwegs‘-Forschungsdaten über das „Datenportal Deutsche Meeresforschung“ ermöglicht und damit eine dauerhafte Nachnutzung der Daten gewährleistet. Für den Aufbau der gemeinsamen Dateninfrastruktur und die Entwicklung der Datenrichtlinien ist eine Einstiegsphase von mindestens zwei Jahren notwendig. Datenrichtlinien beinhalten die Regelung von Datenflüssen, rechtlichen Aspekten im Umgang mit Daten, die Einigung auf Standards und Prozeduren zur Beschreibung der Qualität etc.

Zu den Zielen im Umgang mit ‘Unterwegs’-Forschungsdaten gehören folgende Punkte:

- Schaffung eines gemeinsamen organisatorischen Rahmens zur Ausarbeitung, Optimierung und praktischen Umsetzung von Datenrichtlinien und standardisierter Arbeitsabläufe für Forschungsdaten.
- Die Standardisierung von Metadaten, wie Wegstrecken und Stationsdaten der Schiffe aber auch der eingesetzten Geräteplattformen und Sensoren, muss erarbeitet werden und die Metadaten müssen direkt nach Ende einer Expedition in existierenden Repositorien eingespeist werden. Teil der Metadatenverarbeitung ist die Erstellung des „Cruise Summary Reports“, der die europäische und internationale Sichtbarkeit der Expeditionen durch die Nutzung des EXPCODE (www.bco-dmo.org) gewährleistet.
- Metadaten der über das DSHIP-Informationssystem erfassten Schiffssensordaten sollen vereinheitlicht, validiert und in einem Datenrepositorium als nachnutzbare, harmonisierte und mit den Primärdaten verknüpfte Daten, erfasst und zitierfähig veröffentlicht werden.
- Schiffssensordaten sollen dabei webbasiert durch noch zu entwickelnde, zentral gestützte, semiautomatische Prozeduren und Werkzeuge in Bezug auf Vollständigkeit, Einhaltung von Formaten/Standards und, soweit möglich, Plausibilität kontrolliert und mit Metainformationen versehen werden, um anschließend, mit einem Minimum an manueller Qualitätskontrolle in einen zentralen ‚Common Workspace‘ eingespeist zu werden. Von dort hat hier beantragtes wissenschaftliches Personal Zugriff auf die Daten und kann sie weiterverarbeiten. Ein Teil der Daten kann von dort automatisiert in PANGAEA fließen und publiziert werden. Mittelfristig wird die Übermittlung in ein zentrales ‚Common Workspace‘ fast in Echtzeit bereits während Expeditionen möglich sein, so dass eine umgehende Weiterverarbeitung und Publikation der qualitätskontrollierten Daten in PANGAEA möglich wird.
- Die archivierten Forschungsdaten werden anschließend über nationale und internationale Portale und Repositorien abrufbar und - pro Reise oder integriert - zentral im „Datenportal Deutsche Meeresforschung“ auffindbar sein. Der Zugriff für die Wissenschaft soll direkt und einfach zur Auswertung und weiteren Nutzung ermöglicht werden. Suchfunktionen wie Raum, Zeit sowie eine Verschlagwortung sollen die Auffindbarkeit erleichtern.
- Die strukturierte Erfassung und Überführung der großvolumigen Massenspeicherdaten der Hydroakustik wie Fächerecholot, Sedimentecholot, Fischereilot sowie Strömungsmessungen (ADCP) und profilierenden Messverfahren (CTD) sollen in kontrollierter Form mit automatischer Metadatengenerierung in den ‚Common Workspace‘ ermöglicht werden. Damit können die gesicherten Rohdaten vereinheitlicht, mit allen nötigen Informationen zur weiteren Qualitätssicherung und Nachnutzbarkeit, der anschließenden wissenschaftlichen Aufbereitung zur Verfügung gestellt werden.
- Die Forschungsdaten sollen zitierbar mit dem gesamten Berichtswesen (u.a. „Cruise Summary Reports“, Fahrtberichten) und den im Weiteren damit verbundenen, wissenschaftlichen Publikationen nachhaltig im „Datenportal Deutsche Meeresforschung“ über DOI verknüpfbar und auffindbar gemacht werden.
- Es finden Vorbereitungsarbeiten statt, um langfristig den automatischen Austausch von Datensätzen zwischen Schiff- und Land-Servern zu ermöglichen, nahezu Echtzeitverfahren der Datenbereitstellung auszubauen.

Auf Grund des Gesamtumfangs von unterschiedlichen Typen von ‚Unterwegs‘-Forschungsdaten ist es gegenwärtig nicht realistisch alle ‚Unterwegs‘-Forschungsdaten der vier großen Forschungsschiffe innerhalb der Anschubphase in qualitätskontrollierte und standardisierte Daten, die für Datendienste und -produkte nutzbar sind, zu verarbeiten. Im Rahmen der Anschubfinanzierung werden sich die teilnehmenden Einrichtungen daher neben den unverzichtbaren Metadaten zur Navigation (Zeit, Position, Schiffsbewegung, Log, Kompass) auf die Kuration von einer wissenschaftsgetriebenen Prioritätenliste von ‚Unterwegs‘-Forschungsdaten konzentrieren, unter welchen drei Datentypen als ‚show-case‘ behandelt und zu Datenprodukten und Servicezwecken verarbeitet werden sollen. Alle hier nicht berücksichtigten ‚Unterwegs‘-Forschungsdaten werden behandelt und bearbeitet wie bisher – weitestgehend unsystematisch, in der Gemeinschaft unabgestimmt, ohne die systematische Einhaltung, geschweige denn Kontrolle von Standards und einem insgesamt in weiten Teilen mangelhaft durchgesetztem Forschungsdatenmanagementplan.

Zu den Geräten und Sensoren, die die Daten der Prioritätenliste generieren gehören das CTD-System, Bathymetriesignale (inklusive Schallgeschwindigkeitsmessungen), der akustische Strömungsmesser, der Thermosalinograph, die Ferry-Box und Bio-optische Sensoren. Es steht bereits fest, dass Bathymetriedaten sowie Temperatur und Salzgehaltsdaten als ‚show-case‘ behandelt werden. Es sollen sowohl historische als auch ‚Unterwegs‘-Bathymetrie-Daten behandelt werden, da diese im Rahmen zukünftiger Fahrplanungen eine erhebliche Rolle spielen können und zunehmend Anfragen von internationalen Partnerinstitutionen zum Zugang zu diesen Daten gestellt werden. Temperatur und Salzgehalt der Meere und Ozeane sind fundamentale Parameter der Klimaforschung. Es besteht höchste Dringlichkeit, Wege zu etablieren, die die Integration hydrographischer Daten ermöglicht und automatisiert, um der Forschungsgemeinschaft systematisch Daten höchster Qualität anbieten zu können.

Ausgenommen der Bathymetrie-Daten und CTD - Daten, werden dabei nur Daten berücksichtigt die nicht Teil des wissenschaftlichen Programms der jeweiligen Expedition sind. Mit Blick auf die Ziele, dient die Anschubphase anfangs dem Aufbau von Organisationsstrukturen, damit zukunftsfähige technische Strukturen sowie Arbeitsabläufe geordnet und im Sinne der Nutzung innerhalb und außerhalb der Wissenschaft umgesetzt werden können. Abbildung 1 zeigt zehn Meilensteine des Pilotprojektes, die sich in zwei Hauptaufgabenbereich unterteilen lassen:

- a. **Datenfluss Forschungsdatenmanagement:** Entwicklung und Implementierung standardisierter Arbeitsabläufe an Bord sowie an den verantwortlichen Institutionen zur Beschreibung und Validierung der (zunächst priorisierten) ‚Unterwegs‘-Forschungsdaten. Diese geschieht auch im Hinblick auf das langfristige Ziel eines systematischen Datenflusses für alle ‚Unterwegs‘-Forschungsdaten auf allen deutschen Forschungsschiffen.
- b. **Meilensteine technischer Aspekt:** IT-Entwicklungsarbeiten zur technischen Implementierung der Datenflüsse und des „Datenportals Deutsche Meeresforschung“.

Die Meilensteine und Teilprojektziele der Bereiche a) und b) sind der Abbildung 1 zu entnehmen. Textlich wird im Folgenden nicht weiter auf die Meilensteine eingegangen. Sie dienen dem Projektmanagement der Pilotphase und werden zu Beginn des Projektes gemeinsam mit dem einzustellenden Personal weiter ausdifferenziert.

DAM Kernbereich „Datenmanagement und Digitalisierung“
Stand: Juni 2019

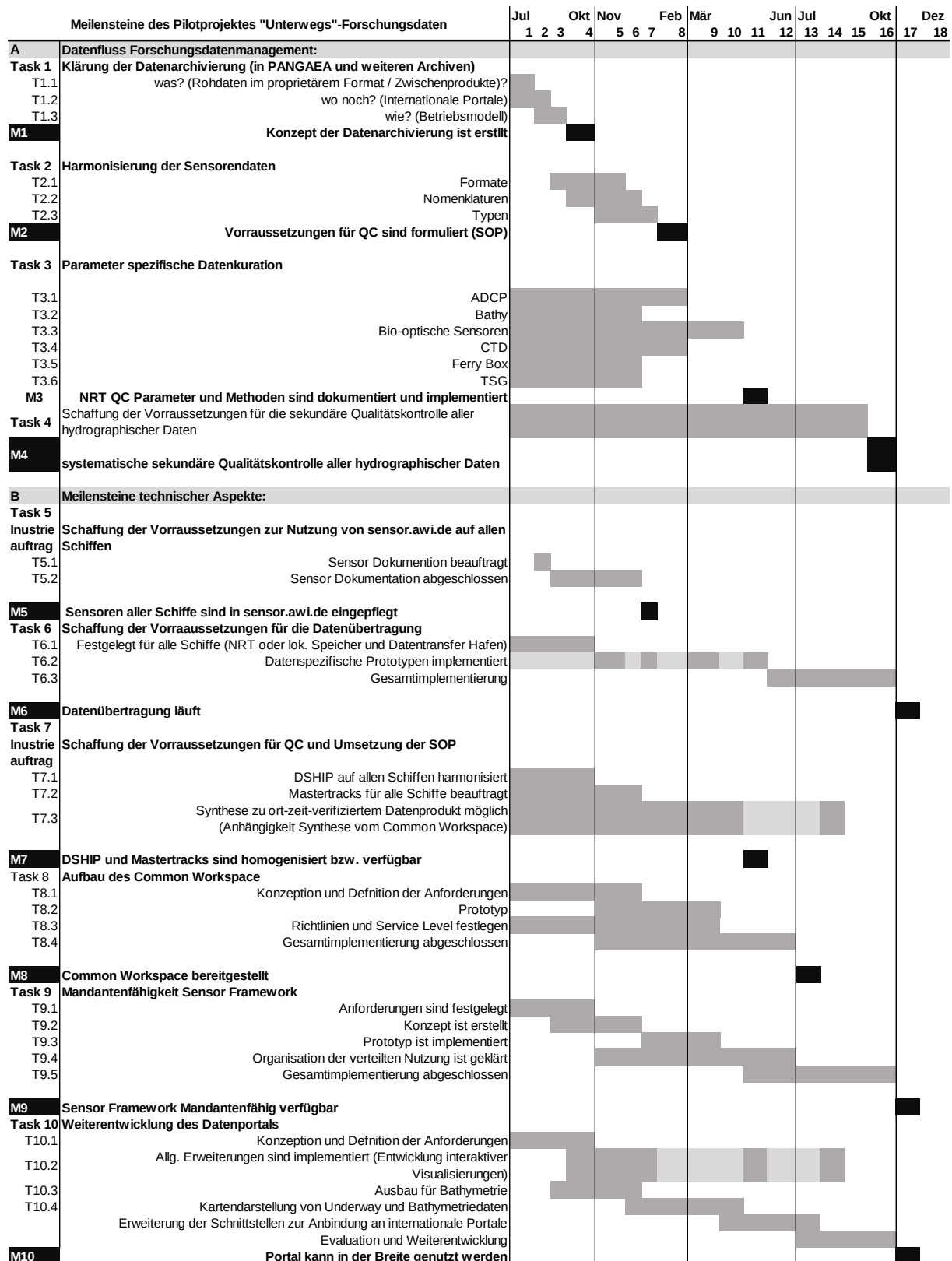


Abbildung 1: Meilensteine und Teilaufgaben

4.2. Partner und Aufgabenpakete

Es wurden fünf Arbeitspakete bestimmt, die zum Erreichen der gesetzten Meilensteine bearbeitet werden sollen. Sowohl die Meilensteine als auch die Arbeitspakete wurden mit Hilfe eines Experten-Teams ausgearbeitet und anschließend vom Leitungsteam des DAM Kernbereich „Datenmanagement und Digitalisierung“ geprüft. Das Experten-Team setzt sich aus wissenschaftlichem und IT-Personal der voraussichtlichen DAM Mitgliedsinstitutionen Alfred-Wegener-Institut, Helmholtz-Zentrum für Polar- und Meeresforschung (AWI), GEOMAR Helmholtz-Zentrum für Ozeanforschung Kiel, Helmholtz- Zentrum Geesthacht Zentrum für Material- und Küstenforschung (HZG), Leibniz-Institut für Ostseeforschung Warnemünde (IOW), MARUM - Zentrum für Umweltwissenschaften der Universität Bremen, und dem ICBM - Institut für Chemie und Biologie des Meeres der Universität Oldenburg zusammen. Das Leitungsteam besteht aus den Direktor*innen Prof. Dr. Antje Boetius (AWI), Prof. Dr. Detlef Stammer (CEN), Prof. Dr. Oliver Zielinski (ICBM), Prof. Dr. Michael Schulz (MARUM), Prof. Dr. Kay-Christian Emeis (HZG) und Dr. Nicolas Dittert (ZMT) sowie Prof. Dr. Frank Oliver Glöckner (Professor am AWI und Mitglied des Rates für Informationsinfrastrukturen) und Dr. Johannes Karstensen (GEOMAR).

Im Anhang [i] sind die Partnereinrichtungen sowie die ihnen zugeordneten Arbeitspakete und Aufgaben tabellarisch und im Detail aufgeführt. Die Aufgabenpakete wurden Partnereinrichtungen zugeordnet, die aus Sicht des Experten- und Leitungsteams, die höchste Expertise zur Lösung der Aufgabe haben. Einige Aufgaben werden von mehreren Partnern parallel bearbeitet, was die Qualität und die Vernetzung in dezentralen Strukturen sicherstellt. Diese Aufgaben sind daher in der Liste mehrfach an den betroffenen Institutionen verankert. An den schiff-betreibenden Institutionen ist zusätzliche Verantwortung verankert. Die Aufgabe des einzustellenden Personals wird sein, dafür Sorge zu tragen, dass die Aufgaben, die die Schiffsinfrastruktur und den Datenfluss vom Schiff zu Land betreffen, umgesetzt werden.

Abbildung 2 zeigt eine Zusammenfassung aus Arbeitspaketen und beteiligten Einrichtungen, die in Anhang [i] detailliert aufgeschlüsselt sind. Gelistet sind die Aufgabenpakete, beteiligte Partnereinrichtungen sowie die Anzahl von VZÄ, die an den jeweiligen Partnereinrichtungen eingestellt werden müssen, um das Projekt in der vorhandenen Zeit durchzuführen.

Abbildung 2: Zusammenfassung der detaillierten und diesem Antrag angehängten Liste mit Partnern und Arbeitspaketen

Aufgabenpakete	AWI	GEOMAR	HZG	IOW	CAU	CEN	ICBM	MARUM
#1: Erfassung und Beschreibung von Geräten/Sensoren: <i>Sensor.awi.de als zentrales Repository für alle Schiffe und deren Sensoren</i>								
#2: Werkzeuge Monitoring Schiffssensoren: <i>Zusätzliches landbasiertes Monitoring-System integrieren und minimum Quality Control</i>								
#3: Rohdatentransfer: <i>Vereinheitlichen der Schiffsinfrastruktur für den Rohdatentransfer inklusive Datenspeicherung</i>								
#4: Archivierung der Rohdaten, Qualitätskontrolle und Bereitstellung in internationalen Portalen: <i>Archivierung aller Rohdaten mit DOI in PANGAEA UND FAIRe Sensordaten, Metadaten und Daten in internationalen Portalen</i>		Underway Multibeam, TSG	Ferry Box	CTD	ADCP	Secondary QC -T/S	Biopische Sensoren	historisch Multibeam
#5: Datenportal Deutsche Meeresforschung: <i>Auffindbarkeit, Nutzbarkeit und Überführung der 'Unterwegs'-Daten in Produkte</i>								
"In Kind" Beitrag: <i>Institutsübergreifende Aufgaben zum Aufbau einer Organisationsstruktur, von Workflows und gemeinsamen Standards</i>								
Vollzeitäquivalent (Summe: 14)	3	3	1	1	1	1	1	3

4.3. Projektumsetzung

Zur praktischen Umsetzung der Anschubphase sieht der DAM Kernbereich „Datenmanagement und Digitalisierung“ folgende Aufgaben bei der Leitung, den bereits beteiligten Experten und dem Referenten des Kernbereichs sowie dem hier beantragten und einzustellenden Personal:

Aufgaben des wissenschaftlichen Leitungsteams:

- Leitung und Beratung des Expertenteams über dessen Sprecher*in (Referent*in des Kernbereichs).
- Kommunikation mit der Politik und Rückkopplung mit dem Expertenteams über dessen Sprecher*in (Referent*in des Kernbereichs).
- Sorgt für Sichtbarkeit und Akzeptanz des D.A.M. Kernbereichs "Datenmanagement und Digitalisierung" besonders im Rahmen der NFDI4Earth, NFDI4BioDiversity und in der wissenschaftlichen Community.

- Einbindung des Kernbereichs in Forschungsmissionen und D.A.M. Strukturen unter Berücksichtigung der fachlichen Breite.

Aufgaben des Expertenteams und des Referenten des Kernbereichs „Datenmanagement und Digitalisierung“:

Mitglieder des Expertenteams übernehmen folgende „In Kind“-Beiträge

- Die regelmäßige Rückkopplung mit dem Leitungsteam und dem Referenten des Kernbereichs organisiert durch den Referenten.
- Die Vorbereitung, Begleitung und Finalisierung von Anträgen gemeinsam mit und koordiniert durch den/die Referenten*in des Kernbereichs.
- Die Leitung und Steuerung des einzustellenden Personals hinsichtlich der Meilensteine (Abbildung 1) und relevanten projektbezogenen inter-/nationalen Aktivitäten im Bereich Forschungsdatenmanagement.
- Die regelmäßige Evaluierung der erreichten Ziele im Verhältnis zu den festgelegten Meilensteinen und Einleitung korrigierender Maßnahmen.
- Die Vor- und Nachbereitung von Workshops ggf. gemeinsam mit dem/der Referenten*in des Kernbereichs.

Der Referent übernimmt:

- Die Sprecherfunktion des Expertenteams.
- Die Planung, Steuerung, Organisation, Koordination, Finalisierung, Dokumentation und Kommunikation der Aktivitäten des Expertenteams in Absprache mit dem Expertenteam an das Leitungsteam und vice versa.
- Die Schnittstelle zur NFDI4Earth und NFDI4BioDiversity

Abbildung 3 zeigt die Organisationsstruktur aller beteiligten Personen und Einrichtungen. Mit dieser Struktur wird sichergestellt, dass das Leitungsteam, das Expertenteam und das einzustellende Personal über den Verlauf und Status Quo in den unterschiedlichen Projektphasen und Bereichen informiert sind.

Aufgaben, die die Festlegung von Organisationsstrukturen und Richtlinien betreffen, müssen z.T. in Eigenleistung der Institute finanziert und eingeleitet werden. Bereits während der aktuell laufenden Aufbauphase des Kernbereichs „Datenmanagement und Digitalisierung“ greift die DAM (in Gründung) regelmäßig auf die Expertise von Personal zu, das aus Eigenleistung der Institutionen finanziert wird. Zur Entwicklung organisatorischer und technischer Strukturen würden in Summe etwa vier Vollzeitäquivalente (VZÄ) durch die beteiligten Einrichtungen in Eigenleistung aufgebracht werden müssen. Beginnend mit der Anschubphase, bedarf es zur Umsetzung der praktischen Aufgaben, der Aufwendung substanzieller Personalmittel sowohl für wissenschaftliches als auch IT- Personal.

In Abstimmung mit dem Leitungsteam hat das Expertenteam konkrete Aufgaben und Anforderungsprofile an das beantragte und einzustellende Personal aufgestellt. Diese Profile sind in Anhang [ii] pro Einrichtung aufgeschlüsselt.



4.4. Kostenaufstellung und Mittelverteilung

Die Projektkosten setzten sich maßgeblich aus Kosten für Personal und Industrieaufträge zusammen. Die Aufschlüsselung der Personalkosten sowie die Mittelflüsse zwischen antragstellenden Einrichtungen und weiteren Partnereinrichtungen sind in Tabelle 1 dargestellt. Zentren, die der Helmholtz-gemeinschaft angehören (AWI, GEOMAR und HZG) werden Mittel direkt zugewendet. Ausschließlich die Antragstellenden Einrichtungen, AWI und GEOMAR, leiten einen Teil der Mittel an Partnereinrichtungen weiter. Das HZG empfängt Mittel für ein VZÄ (0.182 Mio, €) direkt und leitet keine Mittel weiter. Das AWI empfängt insgesamt 1.9 Mio. €. Darin enthalten sind die Mittel für die Partnereinrichtungen MARUM und ICBM sowie Mittel für Industrieaufträge und Sachmittel zur Finanzierung von Vernetzungs- und Arbeitstreffen wie z.B. Workshops. Das GEOMAR erhält 0.918 Mio. €. Diese Mittel enthalten die Anteile von den Partnereinrichtungen CAU, CEN und IOW.

Sachkosten für den “Common Workspace” und die Archivierung in PANGAEA, werden mit etwa 600€/TB/Jahr beziffert. Grob geschätzt, rechnen wir mit 8 TB Datenvolumen pro Jahr und Schiff. D.h. es wird derzeit mit etwa 20.000€/Jahr gerechnet. Kosten für Lizenzen und Wartung von Infrastrukturkomponenten werden auf zusätzliche 13.000€ geschätzt. Weitere Kosten für Transportkosten von z.B. Wasserproben müssen vom beauftragten Institut als „in Kind“ Beitrag geleistet werden. Dabei geht es um den Transport innerhalb Deutschlands. Gleiches gilt für Referenzmaterialien, die zur Kalibration von Messinstrumenten benötigt werden. Aufträge und geschätzte Kosten für Industrieaufträge sind in Tabelle 2 gelistet. Die Einholung von Kostenvoranschlägen und ggf. Ausschreibung der Aufträge wird zu Beginn des Projektes von der antragstellenden Einrichtung AWI übernommen werden müssen.

Die Umsetzung der gelisteten Aufgaben bedarf mindestens 1.5 Jahre bei vollem Mittelfluss und voller Personalstärke. Es wird angestrebt, die Rest-Paktmittel über die Laufzeit Mitte 2019 bis Mitte 2021 zu verteilen wenngleich die Darstellungen mit 2020 enden. Für eine nachhaltige Verstetigung der Aktivitäten zu ‚Unterwegs‘-Forschungsdaten werden jenseits der Anschubphase Mittel in Höhe von etwa 1.5 Mio. € jährlich benötigt, um dann etablierte

Prozesse nachhaltig zu betreiben. Für die Erweiterung der Prozesse auf Daten weiterer Geräte werden zusätzliche Mittel benötigt, deren Höhe von der Zahl der Geräte und Sensoren abhängt, die berücksichtigt werden sollen.

Tabelle 1: Liste der Partnereinrichtungen und der direkten und weiterzuleitenden Mittel und Kosten.

Zweck der Mittel im Pilotprojekt "Unterwegs-" Forschungsdaten	AWI	GEOMAR	HZG	CAU	CEN	ICBM	IOW	MARUM
3 VZÄ am AWI	341,2							
3 VZÄ am GEOMAR		360						
1 VZÄ am HZG			112,35					
Weiterleitung MARUM	421							421,0
- Personalkosten								324,0
- Sachkosten								21,4
- Reisemittel								10,7
- Gemeinkosten								64,9
Weiterleitung ICBM	140					140,0		
- Personalkosten						108,0		
- Sachkosten						6,7		
- Reisemittel						2,0		
- Gemeinkosten						23,3		
Weiterleitung CAU		140		140,0				
- Personalkosten				108,0				
- Sachkosten				6,7				
- Reisemittel				2,0				
- Gemeinkosten				23,3				
Weiterleitung CEN		140			140,0			
- Personalkosten					108,0			
- Sachkosten					6,7			
- Reisemittel					2,0			
- Gemeinkosten					23,3			
Weiterleitung IOW		140					140,0	
- Personalkosten							116,3	
- Sachkosten							9,8	
- Reisemittel							2,3	
- Gemeinkosten							11,6	
Zur Finanzierung Industrieraufträge (746 T€ (Tab. 2)); Material, Sachkos- ten, Reisekosten	758,5	36	6,254					
Gemeinkosten	239,3	102	63,396					
Summe der Zuwendungen	1.900	918	182					

Insgesamt werden hier 3 Mio. €, für eine Projektlaufzeit von 1.5 Jahren und zur Bearbeitung der dargelegten Aufgaben, beantragt. Die hier beantragten Mittel für den DAM Kernbereich „Datenmanagement und Digitalisierung“ sollen verwendet werden, um das essentielle organisatorische Grundgerüst für die operative Vorbereitung der oben dargestellten Vorhaben zu realisieren. Um diese Zielsetzungen des Kernbereichs wirksam und sichtbar zu erfüllen und inhaltlich wie organisatorisch zu entwickeln, sind eine langfristige Perspektive, ein weiterer Ausbau des Kernbereichs und eine stufenweise Erhöhung sowie Verstetigung des Mittelzuflusses erforderlich.

Tabelle 1: Liste geschätzte Kosten für notwendige Aufträge an Industriepartner

Aufträge an Industriepartner	geschätzte Kosten [k€]
Erstellung der Sensorbeschreibung, durch Rücksprachen mit den Reedereien und den Bestandsaufnahmen an Bord der Schiffe.	11
FIELAX Track-QC (für PS vorhanden)	27
DSHIP Homogenisierung + Konzeption für die stetigen Aktualisierung Konzeption und ggf. Implementierung homogener Hard & Software an Bord der Schiffe	75 158
• redundanter Massenspeicher an Bord	
Datenübertragung für alle Schiffe	
• Integration in DSHIP-System, bord-seitig, in Anlehnung an System für die Polarflugzeuge	
• DSHIP-Erweiterung für autom. Extraktionsjob mit NRT Versendung	100
• Erweiterung DSHIP MapViewer zur Darstellung von near real-time Satelliten-Daten	
DSHIP ActionLog offline Version (Sampling)	75
Hardware / Software für Storage, 300 TEUR - Ausbau der Infrastruktur am AWI	300
• Erweiterung Isilon für Common Workspace	
• Erweiterung für Archiv (Tapes, Laufwerke)	
Summ der Kosten für Industrieaufträge [k€]	746

5 Fazit

Die Deutsche Allianz Meeresforschung will den großen Zukunftsfragen der Meeresforschung auf höchstem Niveau begegnen und den wissenschaftlichen Beitrag zu Handlungswissen für Politik, Wirtschaft und Zivilgesellschaft verstärken.

Um dieses Ziel zu erreichen ist der DAM - Kernbereich „Datenmanagement und Digitalisierung“ von herausgehobener Bedeutung und Dringlichkeit. Der offene und nachhaltige Zugang zu Daten der Meeresforschung ist eine wichtige Grundlage für die Bereitstellung von Handlungswissen.

Das „Datenportal Deutsche Meeresforschung“ ging aus dem Projekt „MaNIDA“ hervor, welches dem Aufbau von ersten wichtigen technischen und organisatorischen Strukturen für den nachhaltigen Zugang von Meeresforschungsdaten dient. Jedoch wurden Datenflüsse von Sensoren zu Archiven nicht nachhaltig und schiffsübergreifend etabliert, so dass über das „Datenportal Deutsche Meeresforschung“ in weiten Teilen auf nicht genutzte Archive zugegriffen wird.

Für die Anschubphase wird hier der Ausbau des webbasierten „Datenportal Deutsche Meeresforschung“ und der dahinterliegenden Strukturen beantragt, um die nachhaltige Nutzung von ‚Unterwegs‘-Forschungsdaten der vier großen Forschungsschiffe (POLARSTERN, SONNE, MARIA S. MERIAN, METEOR) erheblich zu steigern. Es soll eine stabile, nachhaltige gemeinsame Dateninfrastruktur und eine schnelle, zentrale Verfügbarkeit über das „Datenportal Deutsche Meeresforschung“ entstehen, so dass eine einfache Nachnutzung dauerhaft gewährleistet und letztlich Handlungswissen bereitgestellt und die DAM ein solider Partner der NFDI werden kann.

6 Links

- [1] <http://www.rfii.de/de/start/>
- [2] <https://www.nfdi4earth.de/>
- [3] <https://www.nfdi4biodiversity.org/>
- [4] www.rvdata.us/
- [5] www.oceanbestpractices.org
- [6] [https://www.awi.de/fileadmin/user_upload/AWI/Forschung/Daten_und IT-Systeme/O2A/Koppe et all 2015 O2A.pdf](https://www.awi.de/fileadmin/user_upload/AWI/Forschung/Daten_und_IT-Systeme/O2A/Koppe_et_all_2015_O2A.pdf)

7 Anhang

- [i] DAM-KB-DM&Digit_Bedarfe_Stand190527
- [ii] DAM-KB-DM&Digit_Personalprofile_Stand190510