

LEITANTRAG

ZUR FORTSETZUNG UND AUSWEITUNG DES DAM-PROJEKTES “UNTERWEGS”- FORSCHUNGSDATEN

Laufzeit: 01.01.2023 - 31.12.2025

Verbundkoordinator:

Deutsche Allianz Meeresforschung e.V.

Wissenschaftsforum
Markgrafenstraße 37
10117 Berlin

Verbundpartner:

Bundesamt für Seeschifffahrt und Hydrographie

Bernhard-Nocht-Str. 78
20359 Hamburg

Institut für Chemie und Biologie des Meeres, Carl
von Ossietzky Universität Oldenburg

Carl-von-Ossietzky-Str.9
26129 Oldenburg

Christian-Albrechts-Universität zu Kiel, Kiel Marine
Science

Ludewig-Meyn-Str. 10
24118 Kiel

Leibniz-Institut für Ostseeforschung Warnemünde

Seestraße 15
18119 Rostock

MARUM - Zentrum für Marine
Umweltwissenschaften, Universität Bremen

Leobener Str. 8
28359 Bremen

Max-Planck-Institut für Chemie

Hahn-Meitner-Weg 1
55128 Mainz

Version: 1

Letzte Änderung: 16.08.2022

Autor:innen: Vertreter:innen des DAM-Projektes “Unterwegs”-Forschungsdaten

Inhaltsverzeichnis

1. Zusammenfassung	2
2. Wissenschafts-politischer Hintergrund	3
3. Ziele	5
4. Stand der Wissenschaft und Technik	7
5. Arbeitsplan und Vernetzung des Verbundes und der beteiligten HGF-Einrichtungen	9
5.1. Fortsetzung und Ausbau der Aktivitäten an großen Forschungsschiffen	10
5.1.1. Arbeitspaket 1: Sensor Monitoring sowie Datenqualitätssicherung und Bereitstellung	10
5.1.2. Arbeitspaket 2: Datenkuration	17
5.1.3. Arbeitspaket 3: Portalentwicklung und Datenvisualisierung	19
5.1.4. Arbeitspaket 4: Automatisierte Anbindung an internationale Portale	20
5.1.5. Arbeitspaket 5: Technische Integration sowie Betreuung und Ausbau der Dateninfrastruktur	21
5.2. Erweiterung der Aktivitäten auf mittelgroße Forschungsschiffe	25
5.2.1. Arbeitspaket 6: Vorbereitende Maßnahmen	26
5.2.2. Arbeitspaket 7: Umsetzung der Maßnahmen	28
5.2.3. Arbeitspaket 8: Projektkoordination	31
5.3. Sachmittel	33
5.4. Strukturplan	35
5.5. Meilensteinplan	37
5.6. Abgrenzung des Projektes zur NFDI und zum Datenmanagement innerhalb der DAM-Forschungsmissionen	45
6. Verwertungsplan	46
7. Verbundweiter Reiseplan	46
8. Finanzplan des Verbundes	48

1. Zusammenfassung

Die Deutsche Allianz Meeresforschung (DAM) will Forschungsdatenmanagement im Bereich Meeresforschung, unter Berücksichtigung bestehender sowie entstehender Forschungsdaten-Infrastrukturen (z.B. die *Nationale Forschungsdateninfrastruktur* (NFDI)) koordinieren, um Innovation und Datenwissenschaften zu fördern.

Forschungsdatenmanagement wird als Basis für die Anwendung moderner Analysemethoden (KI, Maschinelles Lernen) verstanden. Zu diesem Zweck setzt sich die DAM im Kernbereich *Datenmanagement und Digitalisierung* für FAIRe (von Findable, Accessible, Interoperable, Re-usable) und offen (Engl. Open) zugängliche Expeditionsdaten der Deutschen Forschungsflotte ein. Die Initiative wird gestützt durch eine Empfehlung des Wissenschaftsrates aus dem Jahr 2010 (*Empfehlung zur zukünftigen Entwicklung der deutschen marinen Forschungsflotte*), in der ein beträchtliches Defizit an verfügbaren Expeditionsdaten festgestellt wird.

Das bisherige Projekt “Unterwegs”-Forschungsdaten hat den Status eines konkreten Vorzeigebereichs für die instituts- und flottenübergreifende Bereitstellung der “Unterwegs”-Forschungsdaten großer deutscher Forschungsschiffe nach den FAIR- and Open Data-Prinzipien erreicht. Das Kernziel der hier beantragten Fortsetzungsphase ist, die Empfehlung des Wissenschaftsrates aus dem Jahr 2010, bezogen auf die Verfügbarkeit von Expeditionsdaten der deutschen Forschungsflotte inkl. der mittelgroßen Forschungsschiffe, umfassend umzusetzen. Dies bedeutet die Bereitstellung möglichst vieler Expeditionsdaten (“Unterwegs”-Forschungsdaten sowie Primärforschungsdaten) nach den FAIR- and Open Data-Prinzipien. Das beantragte Arbeitsprogramm umfasst zwei Schwerpunkte:

1. Die Fortsetzung und den Ausbau der laufenden Aktivitäten an großen Forschungsschiffen und
2. die Erweiterung der Aktivitäten auf mittelgroße Forschungsschiffe.

Darüber hinaus wird ein Modul Citizen Science (CS) vorgeschlagen. Ein einfacher und durchaus effizienter Ansatz zur Verbesserung der Informationen zum Zustand der Meere, ist die Steigerung der Datenerhebung auf See durch die Einbeziehung der kommerziellen und privaten Schifffahrt. Zentraler Gegenstand des CS-Moduls ist die Umsetzung einer Machbarkeitsstudie zur Etablierung dieses CS-Moduls als zusätzliche Erweiterung des DAM-Projektes “Unterwegs”-Forschungsdaten.

Als Ergebnis dieser Projektphase ist eine vereinheitlichte Dateninfrastruktur für die Bereitstellung FAIRer und offenen Expeditionsmetadaten und Daten der Deutschen marinen Forschungsflotte zu erwarten. Der nutzer:innenfreundliche Zugang zu Expeditionsdaten wird über das *Portal Deutsche Meeresforschung* (marine-data.de)

gewährleistet. Eine abschließende Evaluierung des Projektes wird das Finanzvolumen beziffern, das für einen stetigen Betrieb der zum Ende der beantragten Phase aufgebauten Infra-, Organisations- und Kommunikationsstrukturen notwendig wäre.

2. Wissenschafts-politischer Hintergrund

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

Hinter dem Begriff Forschungsdatenmanagement steckt eine explizite Anforderung an die Wissenschaft, die unter anderem in einer Leitlinie der Deutschen Forschungsgemeinschaft (DFG) oder den Förderrichtlinien der Programme Horizon 2020 und Horizon Europe beschrieben ist. Die verstärkte Aufmerksamkeit für Forschungsdatenmanagement ist dem gesteigerten Interesse an der Professionalisierung und Standardisierung der administrativen und wissenschaftlichen Aufgaben im Umgang mit Daten zu verdanken. Forschungsdatenmanagement bringt der wissenschaftlichen Arbeit einen Mehrwert und fördert den Wissenschaftsbetrieb als Ganzes. Gründe sind die Sichtbarkeit und Nachnutzung von Forschungsdaten, Effizienz und Zeitersparnis im Alltag der Wissenschaft, Datensicherheit und Nachvollziehbarkeit sowie bessere Chancen auf Forschungsförderungen.

Die DAM hat im Kernbereich Datenmanagement und Digitalisierung im September 2019 mit der Umsetzung des DAM-Projektes "Unterwegs"-Forschungsdaten begonnen. Im Kern geht es um die systematische Erfassung, Übertragung an Land, Qualitätssicherung und Veröffentlichung von "Unterwegs"-Forschungsdaten nach den *FAIR- and Open Data*-Prinzipien. Der Fokus lag bisher auf den Forschungsschiffen (FS) MARIA S. MERIAN, METEOR, POLARSTERN und SONNE sowie der Segelyacht EUGEN SEIBOLD. Als „Unterwegs“-Forschungsdaten werden jene Daten bezeichnet, die von bordeigenen, meereswissenschaftlich relevanten Sensoren, zumeist beschränkt auf Transitstrecken in internationalen Gewässern, gemessen werden. Dies dient u.a. der Effizienzsteigerung in der Nutzung der Forschungsschiffe als Messplattformen. Bei der Qualitätssicherung und

systematischen Publikation werden im DAM-Projekt bisher ausschließlich “Unterwegs“-Forschungsdaten berücksichtigt. Dies sind nur Daten, die nicht Teil des wissenschaftlichen Programms der jeweiligen Expedition sind. Forschungsdaten, die als Teil des wissenschaftlichen Programms gesammelt werden, werden hier als “Primärforschungsdaten” bezeichnet. Expeditionsdaten umfassen Primär- sowie “Unterwegs“-Forschungsdaten.

Dass die DAM sich mit ihrer Gründung im Kernbereich Datenmanagement und Digitalisierung für FAIRe und offene Expeditionsdaten einsetzt, geht u.a. auf die *Empfehlung zur zukünftigen Entwicklung der deutschen marinen Forschungsflotte* des Wissenschaftsrates (2010) zurück. Die Notwendigkeit Expeditionsdaten der deutschen Forschungsflotte systematisch bereitzustellen, wurde vom Wissenschaftsrat ausdrücklich angemerkt:

„Die Nutzerinnen und Nutzer von Forschungsschiffen sind gehalten, mit Hilfe von Forschungsschiffen gewonnene Daten zu sichern und in einem angemessenen Zeitraum der wissenschaftlichen Öffentlichkeit zur Nutzung zugänglich zu machen. Schon jetzt müssen Antragsteller in ihrem Antrag auf Schiffszeit ein Konzept zur Nutzung, Archivierung und Freigabe der Forschungsschiffsdaten vorlegen. Mit PANGAEA, der vom MARUM und AWI betriebenen Datenbibliothek, ist ein Instrument geschaffen worden, das die Archivierung und Publikation von Daten aus der Erdsystemforschung ermöglicht. Gegenwärtig werden in PANGAEA, jedoch nur ca. 70 % der Daten eingespeist, die mit Hilfe der POLARSTERN gewonnen werden, die Daten der übrigen Forschungsschiffe der deutschen Flotte gehen sogar nur zu ca. 10 % und unregelmäßig in die Datenbibliothek ein.

Der Wissenschaftsrat empfiehlt dem KDM [Konsortium Deutsche Meeresforschung], allen meeresforschenden Einrichtungen und den Zuwendungsgebern (DFG und BMBF), die Nutzerinnen und Nutzer von Forschungsschiffen noch nachdrücklicher anzuhalten, im Sinne einer Selbstverpflichtung den Richtlinien der DFG-Senatskommission für Ozeanographie zur Archivierung und Publikation von Forschungsschiffsdaten zu folgen. Mit PANGAEA steht dafür ein sehr gutes Werkzeug zur Verfügung, das jedoch besser genutzt werden könnte.“

(Wissenschaftsrat: *Empfehlung zur zukünftigen Entwicklung der deutschen Forschungsflotte*. 2010, S. 101)

Das DAM-Projekt “Unterwegs“-Forschungsdaten führt zu einer systematischen Nutzung von PANGAEA als Repositorium für “Unterwegs“-Forschungsdaten, zur Vereinheitlichung im Umgang mit Expeditionsdaten an Bord der Forschungsschiffe und an Land, zur nutzer:innenfreundlichen Bereitstellung von Expeditionsdaten über das *Portal Deutsche Meeresforschung* und fördert den dafür nötigen kulturellen Wandel innerhalb der wissenschaftlichen Community. Die Einwilligung und Bereitschaft der Wissenschaft,

“Unterwegs“-Daten sowie weitere Expeditionsdaten *FAIR* und offen bereitzustellen, sind Voraussetzungen für alle Serviceleistungen dieses Projektes. Vor jeder Expedition mit einem der großen Forschungsschiffe werden die Fahrtleitungen um Unterstützung für das Projekt “Unterwegs“-Forschungsdaten gebeten. Sie werden über technische und organisatorische Neuerungen im Zusammenhang mit der Übertragung von Expeditionsdaten informiert und angehalten, Expeditionsdaten aus Forschungsgebieten in PANGAEA zu publizieren. Die Publikation in PANGAEA führt automatisch zur nutzer:innenfreundlichen Auffindbarkeit der Daten über das *Portal Deutsche Meeresforschung*. Die aufgebaute Daten-Infrastruktur kann allerdings nicht nur im Sinne der Projektziele mit Bezug zu “Unterwegs“-Forschungsdaten, sondern auch im Sinne der Empfehlungen des Wissenschaftsrates und bezogen auf alle Expeditionsdaten, durch die Wissenschaft genutzt werden. Durch das DAM-Projekt “Unterwegs“-Forschungsdaten werden der Wissenschaft die notwendige Infrastruktur und personelle Kapazitäten bereitgestellt, um Expeditionsdaten sicher und systematisch an Land zu bringen sowie *FAIR* und offen bereitzustellen. Die Inanspruchnahme der Serviceleistung geschieht auf freiwilliger Basis und steht aktuell nur für große Forschungsschiffe zur Verfügung.

Das BMBF hat kürzlich den Wissenschaftsrat beauftragt eine neue Auflage der *Empfehlung zur zukünftigen Entwicklung der deutschen marinen Forschungsflotte* mit besonderem Augenmerk auf mittelgroße Forschungsschiffe gebeten. Der Wissenschaftsrat wird voraussichtlich bis Q3 2022 dafür eine Arbeitsgruppe einsetzen. Diese Arbeitsgruppe wird vermutlich eine vergleichbare Bestandsaufnahme vollziehen wie 2010 und aller Wahrscheinlichkeit nach zu der Erkenntnis kommen, dass weder Expeditionsdaten mittelgroßer Forschungsschiffe in gleichem Maße *FAIR* und offen zugänglich sind, wie es Expeditionsdaten großer Forschungsschiffe durch das DAM-Projekt “Unterwegs“-Daten nun sind, noch wird es einheitliche Strukturen geben, die dies ermöglichen. Die DAM weist somit darauf hin, dass 1. Defizite in der Bereitstellung der Expeditionsdaten mittelgroßer Schiffe der Community bekannt sind, 2. der Wissenschaftsrat parallel zu der hier beantragten Projektphase dieses Defizit ebenfalls erkennen und benennen wird und 3. dieser Antrag u.a. der Behebung dieses Defizits dient.

3. Ziele

Das Projekt “Unterwegs“-Forschungsdaten verfolgt im Rahmen des DAM-Kernbereichs *Datenmanagement und Digitalisierung* fachliche, strukturelle und strategische Ziele. Fachlich erreicht das Projekt “Unterwegs“-Forschungsdaten den Status eines konkreten Vorzeigebereichs für die instituts- und flottenübergreifende Bereitstellung der “Unterwegs“-Forschungsdaten großer deutscher Forschungsschiffe nach den *FAIR- and Open Data*-Prinzipien. Das Kernziel der hier beantragten Phase ist, die Empfehlung des

Wissenschaftsrates aus dem Jahr 2010 umzusetzen. Dies bedeutet die Bereitstellung möglichst vieler Expeditionsdaten ("Unterwegs"-Forschungsdaten sowie Primärforschungsdaten) nach den *FAIR- and Open Data*-Prinzipien der Deutschen Forschungsflotte. Eine Ausweitung der Aktivitäten des DAM-Projektes auf mittelgroße Forschungsschiffe soll im Rahmen der hier beantragten Projektphase zunächst konzipiert und anschließend umgesetzt werden. Es wird zusätzlich dafür Sorge getragen, dass (Expeditions-)Daten und Metadaten, die im Rahmen von DAM-Forschungsmissionen gesammelt werden, von den Projektentwicklungen und -erfahrungen profitieren und in PANGAEA möglichst nach identischen Standards behandelt werden.

Für die vollständige Bereitstellung von Expeditionsmetadaten und Daten der Deutschen Forschungsflotte müssen 1. die laufenden Aktivitäten an "Unterwegs"-Forschungsdaten der großen Schiffe optimiert sowie auf mittelgroße Schiffe ausgeweitet und 2. der kulturelle Wandel im Umgang mit Daten aktiv befördert werden. Wissenschaftler:innen müssen von der Nutzung der Daten-Infrastruktur sowie der *FAIR*en und offenen Bereitstellung von Primärforschungsdaten überzeugt werden. Optimierungsmaßnahmen im Umgang mit "Unterwegs"-Forschungsdaten der großen Forschungsschiffe dienen der Effizienzsteigerung, die zu freien personellen Kapazitäten führt, welche für die Ausweitung auf mittelgroße Forschungsschiffe verwendet wird. Konkret geht es um das FS ALKOR, VWFS ATAIR, FS ELISABETH-MANN-BORGESE und FS HEINCKE. Des Weiteren werden weitere Parameter/ Geräte zum Repertoire der "Unterwegs"-Messung hinzugefügt (pCO₂, Lichtfelddaten, Daten des Dust-Monitors, Parasound-Daten).

Strukturell belegt das DAM-Projekt "Unterwegs"-Forschungsdaten beispielhaft, dass der DAM-Kernbereich *Datenmanagement und Digitalisierung* ein geeignetes und erfolgreiches Instrument zur Koordinierung und Weiterentwicklung des institutsübergreifenden Datenmanagements ist. Der *FAIR*e und offene Umgang mit heterogenen und interdisziplinären Daten der Meeresforschung wird beispielhaft durchgeführt. Bestehende (Infra)-Strukturen werden durch die übergreifende Vernetzung von Hochschulen und (Ressort)-Forschungseinrichtungen inhaltlich und technisch ergänzt und verbessert. Die Bedeutung des Kernbereichs für den kulturellen Wandel muss allerdings weiter gestärkt werden. Ziel ist daher die Sichtbarkeit der Errungenschaften und Vorhaben des Projektes "Unterwegs"-Forschungsdaten durch konkrete Maßnahmen, wie z.B. Informationsveranstaltungen für die Wissenschaft, zu steigern.

Fazit: Strategisch bildet das Projekt "Unterwegs"-Forschungsdaten einen zentralen Baustein im Datenmanagement-Konzept der deutschen Meeresforschung. Das Projekt trägt zum koordinierten Vorgehen im Bereich Forschungsdatenmanagement, zur Sichtbarkeit sowie zur Sprech- und Handlungsfähigkeit des DAM-Kernbereichs *Datenmanagement und Digitalisierung* bei. Des Weiteren ist das Projekt eine Maßnahme

der Deutschen Meeresforschungsgemeinschaft, vereint im Sinne der Ziele der *Nationalen Forschungsdateninfrastruktur* zu agieren und Meeresforschungsdaten, als bedeutenden Bestandteil von Erdsystemforschungsdaten und Biodiversitätsdaten, nach *FAIR- and Open Data*-Prinzipien zu behandeln.

4. Stand der Wissenschaft und Technik

In diesem Projekt wird die systematische Erfassung, Übertragung vom Schiff an Land, Qualitätssicherung und Veröffentlichung von "Unterwegs"-Forschungsdaten nach den *FAIR- and Open Data*-Prinzipien geleistet. Zu den Geräten und Sensoren, die aktuell behandelt werden, gehört das CTD (Conductivity, Temperature and Depth) -System, Bathymetrie-Signale (inklusive Schallgeschwindigkeitsmessungen), der akustische Strömungsmesser (ADCP), der Thermosalinograph (TSG), die *Ferrybox* und Bio-optische Sensoren.

Für die Daten der genannten Geräte und Sensoren wurden Arbeitsabläufe, Qualitätskontrollen und Datenflüsse vom Moment und Ort der Datenaufnahme bis zur Publikation im Datenarchiv PANGAEA als Standards innerhalb der jeweiligen Fachgemeinschaft entwickelt. Die Umsetzung sowie die Erstellung von sogenannten „Standard Operation Procedures“ begann in der ersten Phase des Projektes „Unterwegs“-Forschungsdaten (2021-2022). Der etablierte Datenfluss ist schematisch in Abbildung 1 dargestellt.

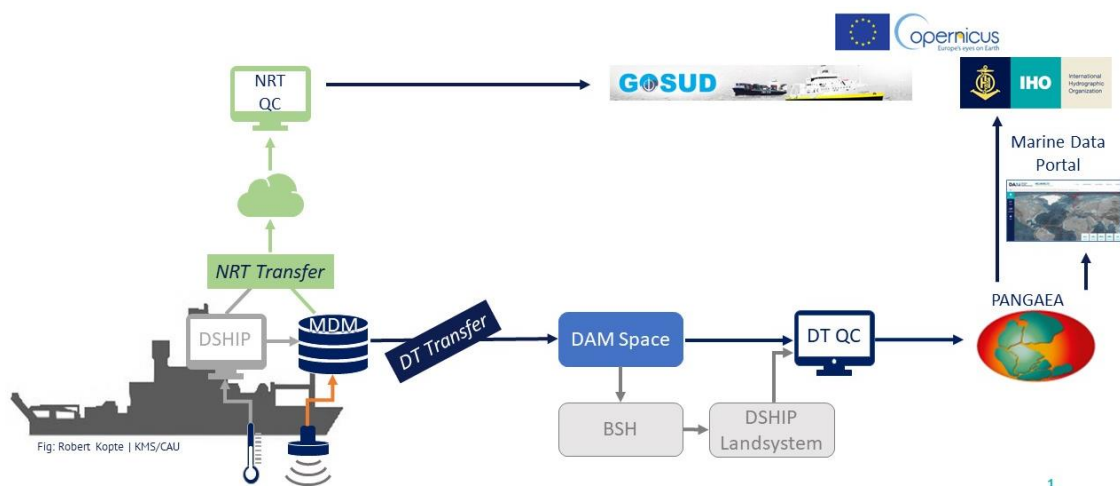


Abbildung 1: Standardisierter Datenfluss vom Sensor bis ins Archiv (PANGAEA) und Portal Deutsche Meeresforschung. Erstellt von Robert Kopte / KMS, CAU

Im Dezember 2020 fand eine Expedition auf der MARIA S. MERIAN statt, die besonders der Einführung und dem Test einer neuen Dateninfrastruktur an Bord diente. Diese neue Dateninfrastruktur ist seit Dezember 2021 auf allen großen deutschen

Forschungsschiffen implementiert und im regulären Dauerbetrieb. Die Datenaufnahme und Übertragung an Land wurden auf allen vier großen deutschen Forschungsschiffen (FS) vereinheitlicht. Der Datenfluss ist auf den genannten Forschungsschiffen (Abb. 1) identisch.

Die Wissenschaft bzw. Expeditions-Fahrleiter:innen werden im Rahmen von Fahrtvorbereitungstreffen der Leitstelle Deutsche Forschungsschiffe durch Projektbeteiligte über alle Neuerungen an Bord sowie den Datenfluss informiert. Da die vereinfachte und vereinheitlichte Übertragung der Daten an das Landsystem ebenfalls die Primärforschungsdaten betrifft, profitiert die Wissenschaft, auch unabhängig von ihrem Bezug zu "Unterwegs"-Forschungsdaten, von der neuen Infrastruktur. Von einer Implementierung auf der EUGEN SEIBOLD wurde abgesehen, da das System für das geringere Datenvolumen überdimensioniert wäre.

Mit der Anlandung der Daten gelangen die Daten in den *DAM-Space*, anschließend wird eine Kopie am BSH archiviert (Abb. 1). Geräteverantwortliche haben im *DAM-Space* Zugriff auf Daten, die während der Expedition im *Massendatenmanagement-System* (MDM) auflaufen. "Unterwegs"-Forschungsdaten werden durch Expert:innen von Primärforschungsdaten getrennt, per Algorithmus qualitätskontrolliert und mit Metadaten angereichert, um die *FAIRness* der Daten zu maximieren. Die Qualitätskontrolle und Publikation von "Unterwegs"-Forschungsdaten wird in Zusammenarbeit mit den Kurator:innen bei PANGAEA und in Rücksprache mit den Expeditions-Fahrleiter:innen systematisch unternommen. Vom Zeitpunkt des Zugriffs auf "Unterwegs"-Forschungsdaten im *DAM-Space* durch die Geräteverantwortlichen bis zur vollständigen Publikation in PANGAEA, vergehen etwa sechs Monate.

Um die Auffindbarkeit und den Zugang zu qualitätskontrollierten und publizierten "Unterwegs"-Forschungsdaten und weiteren Expeditionsdaten sowie bereits vorliegenden Daten der Meeresforschung zu optimieren, wurde das *Portal Deutsche Meeresforschung* weiter überarbeitet und verbessert. Weitere Nutzer:innenworkshops wurden durchgeführt, um die Nutzer:innenfreundlichkeit des Portals zu testen und zu verbessern. Das Portal ist unter <https://marine-data.de/> erreichbar und nutzbar.

Die Anbindung an internationale Portale wurde für ozeanographische Daten analog zur Pilotphase weitergeführt. TSG-Daten werden täglich an das Projekt *Global Ocean Surface Underway Data* (GOSUD) übermittelt. Weitere ozeanographische Daten stehen kurz vor der automatisierten Weitergabe an Europäische Plattformen (CMEMS, Copernicus). Bathymetrische Daten, die in PANGAEA publiziert werden, werden regelmäßig an das *Data Centre for Digital Bathymetry* (DCDB) der *International Hydrographic Organization* (IHO) übermittelt und tragen so zum *Nippon Foundation-General Bathymetric Chart of the Ocean* (GEBCO) *Seabed 2030 Project* bei (Abbildung 1)

Durch eine enge Verzahnung der *MareHub*-Aktivitäten mit den Aktivitäten im Pilotprojekt "Unterwegs"-Forschungsdaten leistet die DAM einen maßgeblichen Beitrag zum koordinierten Aufbau einer Forschungsdateninfrastruktur für die Meeresforschung.

5. Arbeitsplan und Vernetzung des Verbundes und der beteiligten HGF-Einrichtungen

Der Arbeitsplan der hier beantragten Phase II bezweckt die weitgehende Bereitstellung von FAIRen und offenen Expeditionsdaten der Deutschen Forschungsflotte und umfasst zwei Schwerpunkte:

1. Die Fortsetzung und den Ausbau der laufenden Aktivitäten an großen Forschungsschiffen.
2. Die Erweiterung der Aktivitäten auf mittelgroße Forschungsschiffe.

Der Fokus liegt somit auf der systematischen Erfassung und Übertragung von Expeditionsdaten der Deutschen Forschungsflotte an Land, sowie der Qualitätssicherung und zeitnahe Veröffentlichung von "Unterwegs"-Forschungsdaten und weiteren Expeditionsdaten nach den FAIR- and Open Data-Prinzipien. Die Prozesse hinter diesem Schwerpunkt 1 lassen sich auf Grund der vielfältigen und z.T. komplexen Daten, wechselnder Sensorik und der damit verbundenen intensiven Kommunikation mit den Wissenschaftler:innen, Schiffsbetreiber:innen und Bereederungen nicht vollumfänglich automatisieren und binden somit signifikante personelle Kapazitäten. Dies betrifft insbesondere die Qualitätskontrolle (AG-Ozeanographie, AG-Bathymetrie) sowie die Kuration (PANGAEA) der Daten. Diese Aufgaben müssen durch Expert:innen durchgeführt und können durch Algorithmen nur unterstützt werden, um die angestrebte wissenschaftlich validierte Datenqualität zu erreichen.

Der notwendige Ausbau der laufenden Aktivitäten ergibt sich aus der Tatsache, dass die Prozesse hinter Schwerpunkt 1 zum Ende der Phase I zwar erfolgreich vollzogen werden können, jedoch nach der Implementierungsphase die Notwendigkeit einer Optimierung der Datenflüsse und der Effizienzsteigerung bei der Qualitätssicherung und Veröffentlichung ergab. Diese Effizienzsteigerung ist besonders mit Blick auf die Erweiterung der Aktivitäten auf mittelgroße Forschungsschiffe (Schwerpunkt 2) notwendig, um bei konstanter personeller Kapazität mehr Expeditionsdaten systematisch erfassen, übertragen und qualitätsgesichert veröffentlichen zu können.

Die Aufgaben des Schwerpunktes 1 (Abschnitt 5.1) werden entlang der Arbeitspakete (AP) beschrieben, die bereits im Leitantrag der Phase I (2021-2022) ähnlich betitelt sind. Die Arbeitspakete des Schwerpunktes 2 werden gesondert in Abschnitt 5.2 dargelegt.

Das Arbeitspaket 8 (Koordination des Gesamtvorhabens) umfasst in der hier beantragten Phase beide Schwerpunkte. Abweichend ist das Arbeitspaket 8 als Teil von Abschnitt 5.2 beschrieben wenngleich es auch die Koordination der Arbeitspakete umfasst, die in Abschnitt 5.1 beschrieben sind.

5.1. Fortsetzung und Ausbau der Aktivitäten an großen Forschungsschiffen

Der Schwerpunkt 1 umfasst die Fortsetzung und Weiterentwicklung von Prozessen im Zusammenhang mit der Bereitstellung *FAIR*er und offener "Unterwegs"-Forschungsdaten der großen deutschen Forschungsschiffe. Das bedeutet:

1. Die systematische Erfassung und Übertragung von Daten von den Forschungsschiffen an Land sowie der Qualitätssicherung und Veröffentlichung von "Unterwegs"-Forschungsdaten nach den *FAIR- and Open Data-* Prinzipien und
2. die Dokumentation, Optimierung, Anpassung, Automatisierung und Weiterentwicklung der Prozesse aus der ersten Projektphase zur Effizienzsteigerung.

Es folgt die Beschreibung der Arbeitspakete. Zu Beginn der einzelnen Arbeitspakete wird die Anzahl der Vollzeitäquivalente (VZÄ) aufgeschlüsselt, die in den jeweils führenden AGs tätig sind. Die Anzahl der Personen, die an HGF-Einrichtungen angestellt sind werden ebenfalls in VZÄ angegeben. Der Schwerpunkt 1 wird in folgenden Arbeitspaketen bearbeitet:

5.1.1. Arbeitspaket 1: Sensor Monitoring sowie Datenqualitätssicherung und Bereitstellung

Dieses Arbeitspaket wird führend durch die AG-Bathymetrie und die AG-Ozeanographie umgesetzt. Beide AGs werden von der AG-Technik sowie der AG-Portal und PANGAEA unterstützt.

Die AG-Bathymetrie umfasst drei VZÄ. Die AG-Ozeanographie umfasst acht VZÄ. Beide AGs umfassen zusammen elf VZÄ. Davon werden acht VZÄ aus diesem Antrag finanziert und drei aus *Informations-Infrastrukturmittel* der Helmholtz-Gemeinschaft. In beiden AGs bindet die Bearbeitung der hier beschriebenen Arbeitspakete etwa 75% der personellen Kapazitäten. Daraus ergeben sich etwa 25% personelle Kapazitäten für die Bearbeitung des Schwerpunktes 2 (siehe Abschnitt 5.2).

AG-Bathymetrie:

Die AG-Bathymetrie wird bis zum Ende der Phase I (2021-2022) folgende Kernpunkte erfüllen und in der hier beantragten Phase fortsetzen, um die aufgebauten Prozesse zu betreuen und zu optimieren:

- Die systematische Publikation von “Unterwegs“-Fächerecholot-Daten der großen Forschungsschiffe.
- Datenkuration eingehender Roh- sowie qualitätskontrollierter Bathymetrie-Daten in PANGAEA.
- Die sukzessive Publikation historischer, in PANGAEA unveröffentlichten Bathymetrie-Daten (Fächerecholot-Rohdaten) der großen Forschungsschiffe aus dem Bestand des BSH.
- Systematische Aufbereitung der Bathymetrie-Daten-Publikationen aus dem PANGAEA-Datenbestand für die (Meta-)Datenvisualisierung im *Portal Deutsche Meeresforschung* und Einbindung in das *Data Center for Digital Bathymetry* (DCDB) der *International Hydrographic Organization* (IHO) und der *General Bathymetric Chart of the Ocean* (GEBCO).
- Die Bereitstellung aktueller und hochauflösender Bathymetrie-Karten und Karten mit GEBCO Fächerecholot-Datenabdeckung in DSHIP. Hiermit wird die Möglichkeit eröffnet, die Kartierung des Meeresbodens spontan oder geplant im Rahmen einer Expedition fortzusetzen und Beprobungsstandorte und Standorte für Meeresbodeninstallationen besser auszuwählen.

Neue aufbauende Aufgabenstellungen für die Themenbereiche Publikation und Visualisierung:

Da der Aufbau der Kommunikationsstrukturen und des Datenflusses bis Ende 2022 weitestgehend abgeschlossen und standardisiert sein wird, kann damit gerechnet werden, dass das bestehende Personal freie Kapazitäten haben wird, um neue Aufgabenstellungen zu bearbeiten:

1. Das BSH archiviert hydrographische Rohdaten deutscher Forschungsschiffe systematisch. Daten werden mit Metadaten angereichert und i.d.R. unter Moratorium gestellt. Auf Anfrage durch die Wissenschaft können die Daten vom BSH manuell herausgegeben werden. Eine systematische Publikation unter Berücksichtigung der *FAIR- and Open Data*-Prinzipien mit einhergehender nutzer:innenfreundlichen Auffindbarkeit durch Visualisierung und direktem Zugang zu den Daten ist aktuell nicht gegeben.

In der beantragten Projektphase II soll die Möglichkeit der Übernahme und Ausweitung, im Sinne von *FAIR* (Open *Data* oder unter Moratorium) der aktuellen Aktivitäten des BSH im Bereich der Langzeit-Archivierung von Bathymetrie-Rohdaten, geprüft werden. Der Fokus liegt zunächst auf den großen deutschen

Forschungsschiffen MARIA S. MERIAN, SONNE und METEOR. Sämtliche Daten, die im Rahmen von Polarstern-Expeditionen erhoben werden, werden bereits durch den AWI *Observations to Archive and Analysis* (O2A) Datenfluss berücksichtigt und werden in PANGAEA publiziert. Ein Konzept für die Übernahme und Ausweitung des O2A-Systems wird erarbeitet und soweit möglich direkt umgesetzt.

2. Weiterhin soll in dieser Projektphase das Angebot der verfügbaren Bathymetrie (Meta-)Daten aus PANGAEA-Datensätzen im *Portal Deutsche Meeresforschung* und internationalen Portalen (IHO DCDB) erweitert werden. Diese Erweiterung baut auf den im Rahmen eines Pilotprojektes der Nationalen Forschungsdateninfrastruktur für Erdsystemforschung (NFDI4Earth) entwickelte verbesserte Auswertungssoftware für Bathymetrie-Daten auf. Damit können grobe Messfehler in Fächerecholot-Rohdaten automatisch detektiert und bereinigt werden. Mehr Information befindet sich in der Beschreibung des Pilotprojektes:
https://nfdi4earth.de/images/nfdi4earth/documents/pilots/proposals/119-Bathy4All_Workflows_for_Multibeam_Processing_and_Visualization.pdf. Sobald diese Entwicklung abgeschlossen ist, sollen in PANGAEA archivierte Rohdaten mittels dieses Algorithmus in den PANGAEA Bathymetrie Map-Service (WMS) integriert werden. Darüber hinaus sollen neben Roh- und hochwertig prozessierten Fächerecholot-Daten, Bathymetrie-Daten anderer methodischer Grundlage aus PANGAEA im *DAM-Space* für die Visualisierung im Portal aufbereitet werden (Verbindung zu AP 3).
3. Neben der standardisierten Daten-Publikation für Sediment-Echolote in PANGAEA soll eine begleitende Visualisierung der Schiffsrouten entlang derer weitere geophysikalische Daten vorhanden sind, im *Portal Deutsche Meeresforschung* entwickelt werden. Hierbei soll die Georeferenzierung der Daten nicht nur für die Visualisierung, sondern auch für die PANGAEA-Datenkuration mittels des *Mastertracks* (siehe auch Abschnitt 5.3, Pos. 1) erfolgen. Da die Arbeitsabläufe und verwendeten Werkzeuge denen der Bathymetrie entsprechen, wird dieses Arbeitspaket der AG Bathymetrie zugeordnet (Verbindung zu AP 2, 3, 5).
4. Damit die Nachnutzbarkeit der Daten im Sinne der *FAIR- and Open Data*-Prinzipien verbessert wird, soll ein Werkzeug für PANGAEA konzipiert werden, welches die Minimalanforderung für die Publikation von Rasterdaten in PANGAEA semi-automatisch prüft. Dies dient zum einen der Qualitätssicherung im Sinne der Nachnutzung, zum anderen soll dadurch der Arbeitsaufwand während der Datenkuration minimiert werden (Verbindung zu AP 2, 3).

5. Eine Open Source basierte Bearbeitung der in PANGAEA archivierten Fächerecholot-Rohdaten im DAM-Space ist derzeit nur für DAM-Mitarbeiter:innen möglich. Es soll eine Cloud-Umgebung auch für externe Nutzer:innen aufgebaut werden, welche eine nutzer:innenfreundliche Cloud-gestützte Bearbeitung der Fächerecholot-Daten ermöglicht (Verbindung zu AP 3, 5).
6. Alle Fächerecholot-Rohdaten aus dem Datenbestand des Alfred-Wegener-Instituts sind bereits in PANGAEA publiziert. Zusätzlich sollen nun auch alle prozessierten Bathymetrie-Daten aus dem Datenbestand des AWI sukzessiv in PANGAEA publiziert werden (Verbindung zu AP 2).
7. Auf den Forschungsschiffen MARIA S. MERIAN, METEOR und SONNE sollen die mit dem Fächerecholot kartierten Flächen unmittelbar nach der Aufnahme der Daten in DSHIP als "living vector", d.h. als Polygonvektor, der regelmäßig um die neu kartierte Fläche ergänzt wird, dargestellt werden. So kann während einer Ausfahrt oder bei aufeinanderfolgenden Ausfahrten an bereits aufgenommene Daten direkt angeschlossen werden und somit eine lückenlose Kartierung gewährleistet werden (Verbindung zu AP 5).
8. Die GEBCO Fächerecholot-Datenabdeckung deckt bereits einen Teil der weltweit existierenden Bathymetrie-Daten ab. Es ist geplant, in Zusammenarbeit mit dem *Seabed2030 Atlantic and Indian Ocean Regional Center*, eine Karte zu erstellen, die neben der GEBCO Datenabdeckung, auch Abdeckungen aus anderen Datenarchiven beinhaltet. Die Datenabdeckung der Forschungsschiffe MARIA S. MERIAN, METEOR und SONNE sollen ebenfalls in diese Karte integriert werden.
9. Prozessierte Fächerecholot-Datensätze sollen zukünftig auch in die *Global Multi-Resolution Topography Data Synthesis* (GMRT) integriert werden. GMRT ist eine kontinuierliche Zusammenstellung von Fächerecholot-Daten in Mehrfachauflösung und umfasst Daten aus allen Ozeanen. Der Workflow der Datenprozessierung soll in der kommenden Phase an GMRT angepasst werden, so dass eine problemlose Integration der Daten in GMRT möglich ist.
10. In einem Workshop soll evaluiert und diskutiert werden, wie man die Bathymetrie-Datenaufnahme nachhaltiger gestalten kann, z.B. durch Reduzierung des Geräuschlevels. Mögliche Ergebnisse könnten direkt in die Workflows auf den großen Forschungsschiffen integriert oder als "How To" veröffentlicht werden.

AG-Ozeanographie:

Die AG-Ozeanographie ist für die systematische Erhebung, Qualitätskontrolle und Publikation ozeanographischer "Unterwegs"-Forschungsdaten der großen Forschungsschiffe verantwortlich. Bis zum Ende der Phase I (2021-2022) sind die dazu notwendigen Prozesse standardisiert und können von den im DAM-Projekt "Unterwegs"-Forschungsdaten beschäftigten Mitarbeiter:innen unterhalten werden. Die damit einhergehenden Aufgaben sollen in der hier beantragten Phase II fortgesetzt werden, um die aufgebaute Infrastruktur zu betreuen. Die betreuten Geräte umfassen das ADCP (akustische Strömungsmessung), die CTD, den Thermosalinograph (TSG), die bio-optischen Sensoren und die *FerryBox*. Sofern nicht im Text anders ausgeführt, betreffen die im Folgenden dargestellten Arbeiten den Betrieb und die Daten aller Geräte gleichermaßen. Eine gerätespezifische, detaillierte Darstellung von Arbeitspaketen ist im Rahmen der Teilanträge vorgesehen.

Zur systematischen Erhebung, Qualitätssicherung und Bereitstellung von ozeanographischen "Unterwegs"-Forschungsdaten ist die Fortführung folgender Aufgaben erforderlich:

- Information der Leitstelle und Fahrtleitungen auf Fahrt-Vorbesprechungen: Diese Tätigkeit umfasst einerseits den Informationsaustausch mit den beteiligten Fahrtleitungen über die Angebote des DAM-Projektes "Unterwegs"-Forschungsdaten, als auch die Abstimmung über die Zuständigkeiten zur Datenerhebung und Prozessierung. Diese Abstimmung erfolgt spezifisch für jede Expedition. Die Präsenz der Mitarbeiter:innen auf den Fahrt-Vorbereitungstreffen ist unerlässlich, um den durch das DAM-Projekt "Unterwegs"-Forschungsdaten geschaffenen Mehrwert zu kommunizieren. Diese Kommunikation stellt gleichzeitig einen essentiellen Beitrag zum kulturellen Wandel in Richtung Digitalisierung und Open Science dar.
- Prozessierung täglich eingehender Daten und near real-time (NRT) Bereitstellung in nationalen und internationalen Portalen (*Cosyna*, *GOSUD*, *CMEMS*, *EMODnet*) durch die Mitarbeiter:innen: Es besteht ein klar von den Datennutzer:innen herausgestellter Bedarf an NRT-Daten, bzw. deren Bereitstellung während der einzelnen Ausfahrten. Dieser Bedarf ergibt sich einerseits aus dem Anspruch zur Qualitätssicherung und andererseits aus den wissenschaftlichen Arbeiten an Bord, bzw. der Notwendigkeit, Entscheidungen zum Expeditionsverlauf anhand von den gemessenen Umweltparametern treffen zu können. Dies wird durch die in Phase I des DAM-Projektes "Unterwegs"-Forschungsdaten implementierten Prozesse ermöglicht, die je nach Sensor zum Monitoring des Betriebs und zur Weiterleitung an operative Portale dienen.
- Weiterführung des Probenkonzepts: Betreuung der Beprobungs- und Transportlogistik, Durchführung der Messung im Labor, Einbeziehung der Laborergebnisse in die Qualitätskontrolle der "Unterwegs"-Forschungsdaten: Die

Erfahrungen aus Phase I des DAM-Projektes "Unterwegs"-Forschungsdaten zur Logistik und Probenmessung werden für eine Weiterentwicklung des Probenkonzepts genutzt, welches den erforderlichen Qualitätsstandard der Durchflussmessungen (TSG, bio-optische Sensoren, *Ferrybox*) sicherstellt.

- Erstellung eines qualitätskontrollierten Datensatzes pro Gerät und Ausfahrt, Erstellung des Processing-Reports, Ergänzung von Metadaten und Einreichung zur Veröffentlichung bei PANGAEA: diese Aufgabe ist essentiell für alle Sensoren und steht in enger Verbindung zum Arbeitspaket 2. Das Ziel ist die in den Fahrt-Vorbereitungsgesprächen vereinbarte verlässliche öffentliche Bereitstellung der Daten nach *FAIR- and Open Data*-Prinzipien.
- Automatisierte Darstellung der publizierten Datensätze im *Portal Deutsche Meeresforschung*: über die etablierten Schnittstellen werden die bei PANGAEA veröffentlichten Datensätze abgerufen und zusammen mit anderen Forschungsdaten im *Portal Deutsche Meeresforschung* (<https://marine-data.de>) visualisiert.

Neue Aufgabenstellungen innerhalb der AG-Ozeanographie:

Die AG-Ozeanographie wird in der beantragten Projektphase II eine Reihe weiterer Aufgaben gemeinschaftlich umsetzen, die den folgenden Zielen dienen:

Homogenisierung der Produkte aus "Unterwegs"-Forschungsdaten: Die Behandlung der ozeanographischen Daten wird weiter vereinheitlicht, besonders in Hinblick auf die zusätzlich zu leistenden Dienste im Rahmen der DAM-Forschungsmissionen. Diese Aufgabe beinhaltet die Vereinheitlichung der Darstellung aller durch die AG-Ozeanographie generierten Datenprodukte im *Portal Deutsche Meeresforschung*. Darüber hinaus wird mindestens für die verschiedenen Durchflusssysteme die Qualitätskontrolle vereinheitlicht. Im Rahmen einer Machbarkeitsstudie wird geprüft, inwieweit die Qualitätskontrolle durch Machine Learning Methoden weiter automatisiert werden kann.

Wissenschaft: Ebenfalls mit Blick auf die DAM-Forschungsmissionen wird die Abstimmung mit der ozeanographischen Wissenschaftscommunity verstärkt. Das Ziel besteht darin, jenseits der bereits etablierten Methoden zur Datenprozessierung und -veröffentlichung, die Bedarfe der Wissenschaft im Datenfluss so zu berücksichtigen, dass "Unterwegs"-Forschungsdaten unmittelbar zu wissenschaftlichen Studien genutzt werden können. Die Umsetzung dieser Aufgabe ist gerätespezifisch und dient der verstärkten Nutzung von "Unterwegs"-Forschungsdaten in wissenschaftlichen Projekten. In Kopplung an das Arbeitspaket 3 sollen Auffindbarkeit und Zugriff auf die DAM-Datenprodukte im *Portal Deutsche Meeresforschung* verstärkt auf die Bedarfe der

wissenschaftlichen Zielgruppe zugeschnitten werden. Die Abstimmung geschieht mittels Usability Studien im Rahmen der Nutzer:innenworkshops.

Erweiterung der Sensorik: Die etablierten Methoden zur Datenbehandlung werden genutzt, um weitere Sensoren und entsprechende neue Parameter in den Datenfluss zu integrieren. Dies betrifft einerseits die Integration bereits bekannter Sensoren auf neuen Plattformen wie S/Y EUGEN SEIBOLD und andererseits die Prüfung der Integration neuer Sensoren, wie Lichtfelddaten als Teil der Bio-Sensorik, den Dust-Monitor oder die Messung von CO₂.

Beispielhaft folgen Teilaufgaben, die von den einzelnen Geräteverantwortlichen umgesetzt werden. Alle Teilaufgaben werden später im Rahmen der Teilanträge detailliert beschrieben.

- Machbarkeitsstudie zur Erhöhung des Automatisierungsgrades der Qualitätskontrolle durch Machine Learning Methoden für Daten der Durchflusssysteme, der bio-optischen Sensoren und der *FerryBox*.
- Erweiterte sekundäre Qualitätskontrolle auf Basis der vereinheitlichten DAM-„Unterwegs“-Datenprodukte der Durchflusssysteme und CTD. Auch hier wird das Potential von Machine Learning Methoden evaluiert.
- Machbarkeitsstudie zur Integration von Lichtfelddaten. Prüfung von „Unterwegs“-Lichtfelddaten als neuer Parameter in den Zeitreihen bio-optischer Sensoren. Entwicklung einer Standardprozessierung und eines Verfahrens zur Qualitätskontrolle.
- Weiterentwicklung der Qualitätssicherung von ADCP Daten, basierend auf den Ergebnissen aus Projektphase I. Implementierung neuer Prozessierungsschritte zur Eliminierung der durch akustische Interferenzen hervorgerufenen Artefakte, die im Parallelbetrieb mit anderen hydroakustischen Geräten auftreten.
- Ausweitung der Veröffentlichung von „Unterwegs“-Forschungsdaten in internationalen Plattformen (siehe auch AP 4). Implementierung der erforderlichen Schnittstellen im Fall unterschiedlicher Anforderungen an die Datenformate, wie im Fall von ADCP Daten bei der Veröffentlichung im *Joint Archive of Shipboard Acoustic Doppler Current Profilers* (JASADCP). Integration der Weiterleitung der Daten in den DAM-Datenfluss.
- Berücksichtigung profilierender Daten aus der Wassersäule und Erzeugung entsprechender Datenprodukte. Integration der akustischen Rückstreuintensität von ADCPs in den DAM-Datenfluss. Implementierung neuer Prozessierungsschritte zur Umrechnung in Backscatter und Eliminierung akustischer Interferenzen im Rückstreusignal.
- Prüfung des Betriebs des auf dem FS METEOR installierten Dust-Monitors auf den anderen großen Forschungsschiffen. Evaluation bekannter

Kontaminationsquellen durch Schiffsemissionen und Entwicklung eines Verfahrens zur Qualitätssicherung. Entwicklung von Indikatoren zur Bestimmung der Kontamination unter Verwendung von Aerosol-Messungen von S/Y EUGEN SEIBOLD.

- Machbarkeitsstudie zur Messung von CO₂ an Bord mittelgroßer Forschungsschiffe und auf Ships Of Opportunity am Beispiel der S/Y EUGEN SEIBOLD. Bestimmung der Anforderungen an die Qualität von CO₂-Sensoren, der erforderlichen Metadaten und möglicher Kontaminationsquellen, basierend auf den Daten und Erfahrungen auf der S/Y EUGEN SEIBOLD. Durchführung von Vergleichs- und Testmessungen mit verschiedenen CO₂-Sensoren unter kontaminationsfreien Bedingungen an Bord der S/Y EUGEN SEIBOLD.
- Integration der S/Y EUGEN SEIBOLD *FerryBox* in den DAM-Datenfluss, entsprechend der bereits etablierten Behandlung von *FerryBox* Daten der POLARSTERN.

5.1.2. Arbeitspaket 2: Datenkuration

Dieses Arbeitspaket wird federführend durch PANGAEA-Mitarbeiter:innen umgesetzt. PANGAEA wird von der AG-Bathymetrie und AG-Ozeanographie unterstützt.

PANGAEA benötigt insgesamt drei VZÄ für die Bearbeitung des Arbeitspaketes. Zwei VZÄ werden über diesen Antrag finanziert. Ein VZÄ wird aus *Informations-Infrastrukturmittel* der *Helmholtz-Gemeinschaft* finanziert. Die Bearbeitung des hier beschriebenen Arbeitspaketes bindet etwa 75% der personellen Kapazitäten. Daraus ergeben sich etwa 25% personelle Kapazitäten für die Bearbeitung des Schwerpunktes 2 (Siehe Abschnitt 5.2).

Die Datenkuration bei PANGAEA wird bis zum Ende der Phase I (2021-2022) folgende Kernpunkte umfassen und in der hier beantragten Projektphase II fortgesetzt werden, um die aufgebauten Prozesse kontinuierlich zu betreuen:

- Archivierung von Stationslisten aus *DSHIP ActionLogs* und Daten der Schiffsrouten (*Mastertracks*).
- Die systematische Kuration aller (Expeditions-)daten, die im Rahmen des Projektes und der DAM-Forschungsmissionen in PANGAEA publiziert werden sollen.
- Fortsetzung der Beratung und Unterstützung im Bereich Datenmanagement und Archivierung im Rahmen der laufenden und kommenden DAM-Forschungsmissionen.

Neue Aufgabenstellungen im Zusammenhang mit Daten der deutschen Forschungsflotte

Die Datenkuration bei PANGAEA ist eng mit den anderen Arbeitspaketen verbunden, weshalb die hier gelisteten Aufgaben im Zusammenhang mit Aufgaben aus anderen Arbeitspaketen stehen.

1. Übernahme und Ausweitung der aktuellen Aktivitäten des BSH im Bereich der Langzeit-Archivierung von Bathymetrie-Rohdaten und somit systematische Archivierung unter Berücksichtigung der *FAIR- and Open Data*-Prinzipien (siehe AP1 AG Bathymetrie 1).
2. Kuration und Georeferenzierung von Datensätzen der Sediment Echolote (siehe AP1 AG Bathymetrie 3).
3. Semi-automatische Prüfung von Rasterdaten für die Datenpublikation in PANGAEA (siehe AP1 AG Bathymetrie 4).
4. Archivierung aller prozessierten Bathymetrie-Daten aus dem Datenbestand des Alfred-Wegener-Instituts (siehe AP1 AG Bathymetrie 6).
5. Für eine zeitnahe Archivierung von ADCP und CTD-Rohdaten sollen Konzepte für einen semi-automatischen Ablauf erstellt und etabliert werden, analog zum *Observation to Archive and Analysis* (O2A) Arbeitsflusses des AWI. Diese Konzepte werden so entwickelt, dass sie auf andere Datentypen übertragen werden können.
6. Konzeption, Erstellung und Publikation von Standard Operating Procedures (SOP) für die Datenkuration neuer Sensoren und Archivierung der Daten, zum Beispiel von Lichtfelddaten, CO₂-Daten der S/Y EUGEN SEIBOLD (siehe AP1 AG Ozeanographie) oder des Dust Monitor in Absprache mit Wissenschaftler:innen des TROPOS, oder der Radiosonden des Deutschen Wetterdienstes (DWD) (siehe AP5 AG Technik 4).
7. Verknüpfung der Sensor-Daten mit PANGAEA-Datensätzen sobald eine zuverlässige Dokumentation gewährleistet ist (siehe AP5 AG Technik 6, Abschnitt 5.1.5). Diese Aufgabe soll weitestgehend automatisiert werden.
8. Recherche und Inventarisierung des PANGAEA-Archivs in Hinblick auf verfügbare Veröffentlichungen die für die Aufbereitung nach den Vorgaben der internationalen Datenportale geeignet sind (siehe AP4 AG BSH 1).
9. Auffindbarkeit dezentraler Daten der DAM-Forschungsmissionen im *Portal Deutsche Meeresforschung* gewährleisten in Kopplung an das Arbeitspaket 3.
10. Anbieten von Infoveranstaltungen/Trainings zum Auffinden von Daten der DAM-Forschungsmissionen und weiteren in PANGAEA bzw. im *Portal Deutsche*

Meeresforschung, sowie zur Archivierung von DAM-Daten in PANGAEA (z.B. Daten-Workflows, Vorbereitung der Daten/Metadaten, Prozess der Datenbereitstellung für PANGAEA).

5.1.3. Arbeitspaket 3: Portalentwicklung und Datenvisualisierung

Dieses Arbeitspaket wird führend durch die AG-Portal umgesetzt. Die AG-Portal wird von der AG-Bathymetrie, der AG-Ozeanographie und PANGAEA inhaltlich unterstützt.

Die AG-Portal umfasst vier VZÄ, die aus *Informations-Infrastrukturmittel* und weiteren Mitteln der Helmholtz-Gemeinschaft finanziert werden. Die Umsetzung von Maßnahmen innerhalb des *MareHub/DataHub* werden prioritär behandelt.

Folgende Aufgabe wird durch die AG-Portal fortgesetzt:

- Betrieb des Portals mit den erforderlichen Systemen sowie der Geodateninfrastruktur als Grundlage der Visualisierung von Datenprodukten.

Neue Aufgabenstellungen:

Die folgenden Aufgaben sind in Abhängigkeit von den Entwicklungen im *DataHub* der Helmholtz-Gemeinschaft zu betrachten.

1. Weiterer Ausbau des Portals auf Basis von Nutzer:innen-Feedback und Abstimmungen in der marinen Community sowie der Community des *Fachbereichs Erde und Umwelt der Helmholtz-Gemeinschaft (MareHub, DataHub)*
2. Review und Optimierung der Portal-Inhalte und Portal-Struktur für die Auffindbarkeit in Suchmaschinen
3. Einbindung weiterer “near real-time”-Daten in das Portal zur Darstellung aktueller Messwerte von Forschungsplattformen
4. Konzeption und Implementierung von Funktionen zum zusammenfassenden Herunterladen von Datensätzen
5. Review und Optimierung des Portals für mobile Endgeräte
6. Planung und Durchführung von Nutzer:innen-Workshops und Usability-Studien zum Abgleich und Verbesserung von existierenden Funktionen und Wünschen der Nutzer:innen
7. Ausbau der Dokumentation mit Hilfestellungen und Beispielen für Nutzer:innen

8. Review, Konzeptionierung und Implementierung von verbesserter Skalierbarkeit der Geodateninfrastruktur

5.1.4. Arbeitspaket 4: Automatisierte Anbindung an internationale Portale

Dieses Arbeitspaket wird führend durch das BSH umgesetzt. Das BSH wird von PANGAEA und der AG-Ozeanographie unterstützt.

Das BSH benötigt ein VZÄ, das im Rahmen des Projektes finanziert wird. Die Bearbeitung der hier beschriebenen Aufgaben bindet etwa 100% der personellen Kapazitäten.

Folgende Aufgabe wird fortgesetzt:

- Betrieb des in der Phase I entwickelten *Datenbrokers*. Dazu gehören die regelmäßige Überprüfung, Wartung und notwendigen Anpassungen und Weiterentwicklungen der Funktion, auch im Hinblick auf die in Phase II beantragten und zu erwartenden Neuerungen.

Neue Aufgabenstellungen:

Der Fokus in der Projektphase II liegt auf der *Datenbroker*-Erweiterung, die sich in drei Teilaspekte untergliedert und die Sichtbarkeit von DAM sowie die Nachnutzbarkeit der Beobachtungsdaten durch die wissenschaftliche Community erheblich verbessern.

1. Recherche und Inventarisierung des Archivs (zusammen mit PANGAEA) mit der Fragestellung, welche weiteren, über das DAM-Projekt "Unterwegs"-Forschungsdaten hinausgehenden Veröffentlichungen, zur Verfügung stehen und für die Aufbereitung nach den Vorgaben der internationalen Datenportale geeignet sind. Die Berücksichtigung dieser Veröffentlichungen steigert das Datenvolumen des *Datenbrokers* und erhöht damit zum einen die Sichtbarkeit von DAM. Zum anderen hat die Erschließung von weiteren, qualitätskontrollierten Beobachtungsdaten aus dem PANGAEA-Archiv einen direkten, positiven Effekt auf die Nachnutzbarkeit dieser Daten für die wissenschaftliche Gemeinschaft.
2. Weitere Steigerung der Sichtbarkeit des DAM-Projektes und Nachnutzbarkeit der mittels PANGAEA veröffentlichten Daten durch die Belieferung weiterer Datenportale. In Phase I des Projektes wurde bisher das europäische Datenportal des *Copernicus Marine Service* (CMEMS) bedient. In dieser Phase werden weitere Datenportale wie das *World Ocean Database*-Projekt (WOD) des *International Oceanographic Data and Information Exchange* (IODE) oder die europäische

Initiative EMODnet bedient. Datenströme zwischen einzelnen Datenportalen existieren bereits. Diese gilt es zu identifizieren, um sowohl mögliche Doppelungen von Beobachtungsdaten zu vermeiden als auch die Lücken zu finden, die der *Datenbroker* befüllen kann. Hierfür ist eine umfangreiche Recherche und Absprache mit den Datenportalen nötig, welche noch in Phase I beginnt und in Phase II durch eine Versuchsreihe komplementiert wird: Einzelne Datensätze, die schon in Phase I an CMEMS geliefert wurden, werden hierfür quasi als "Daten-Bojen" genutzt und ihr Weg über die Datenströme durch die Datenportale verfolgt.

3. Die dritte Erweiterung des *Datenbrokers* bezieht sich auf ein Konzept zur Erschließung weiterer Quellen für den *Datenbroker*. So ist zu vermuten, dass es zusätzlich zu reinen Forschungsdaten noch solche gibt, die aus behördlichen Monitoringaufgaben entstanden sind und die der wissenschaftlichen Community nur schwer zugänglich sind. Ein Konzept zur Identifikation und Bereitstellung solcher unerschlossener Datenpools wird in der zu beantragenden Projektphase unter Einbeziehung der Expertise im Projekt und auf Grundlage einer umfangreichen Recherche erstellt.

5.1.5. Arbeitspaket 5: Technische Integration sowie Betreuung und Ausbau der Dateninfrastruktur

Dieses Arbeitspaket wird führend durch die AG-Technik umgesetzt. Die AG-Technik wird von der AG-Bathymetrie, der AG-Ozeanographie und PANGAEA inhaltlich unterstützt. Die AG-Technik umfasst zwei VZÄ, die aus *Informations-Infrastrukturmittel* der Helmholtz-Gemeinschaft finanziert werden. Die Bearbeitung des hier beschriebenen Arbeitspaketes bindet etwa 50% der personellen Kapazitäten. Daraus ergeben sich etwa 50% personelle Kapazitäten für die Bearbeitung des Schwerpunktes 2 (siehe Abschnitt 5.2).

Auf Grundlage der in den zwei vorangegangenen Phasen des Projekts erfolgten Erfassung der Schiffssensorik und Harmonisierung der Sensornomenklatur sowie erfolgreichen Implementierung eines *Massendatenmanagement-Systems* (MDM) auf allen großen Forschungsschiffen, sind nun neben der Aufrechterhaltung der Infrastruktur und des Datenflusses sowie der Interoperabilität von Sensorik und Geräten an Bord der großen Forschungsschiffe weitere Schritte geplant.

- Fortlaufender Betrieb und Wartung der Dateninfrastruktur sowie Aufrechterhaltung der Datenflüsse von den großen Forschungsschiffen in den DAM-Space und weiter zu den Archiven (BSH, PANGAEA)
- Pflege der Metadaten und der Interoperabilität wissenschaftlicher Sensoren und Geräte in für DAM relevanten Informationsinfrastrukturen der großen Forschungsschiffe

- Bereitstellung von qualitätskontrollierten *Mastertracks* der Schiffspositionen (siehe Abschnitt 5.3, Position 1)

Neue Aufgabenstellungen der AG Technik:

Die aufgebaute Infrastruktur soll auch in Zukunft eine insgesamt robuste, zuverlässig funktionierende Lösung für Datenflüsse und Kommunikation zwischen den Forschungsschiffen und Landsystemen sowie zwischen den einzelnen Landsystemen bereitstellen.

Im Folgenden sind die konkreten Maßnahmen aufgeführt, welche z.T. an die Bereitstellung von Sachmitteln (Abschnitt 5.3) gebunden sind. Es wird jeweils auf die Positionen der Sachmitteltabelle (Tabelle 1) verwiesen. Die folgenden Maßnahmen sind ebenfalls im Meilensteinplan (Abschnitt 5.6) erfasst:

1. Nach der erfolgreichen Etablierung von Arbeitsflüssen zur Erfassung, Archivierung und Publikation der "Unterwegs"-Forschungsdaten auf den großen deutschen Forschungsschiffen muss nun eine detaillierte Dokumentation der Abläufe erfolgen. Es müssen Standard Operation Procedures (SOPs) erstellt werden, um eine einheitliche flotten- und institutsübergreifende Nutzung zu erreichen. Der Zugang zu diesen Dokumentationen muss an die Wissenschaft und die Schiffsbesatzungen kommuniziert und bei Bedarf auch Trainings zu deren Anwendung durchgeführt werden.
2. Das Erstellen von SOPs für etablierte technische Abläufe, wird u.a. als Basis für die Ausweitung der Abläufe auf mittelgroße Forschungsschiffe (siehe Abschnitt 5.2) genutzt. Die Möglichkeit der Ausweitung soll in Schwerpunkt 2 dieses Antrages in Betracht gezogen werden. Sollte eine Übertragung zügig und ohne Komplikationen umsetzbar sein, kann auch die Einbindung von Sensordaten auf privaten Segelbooten in Betracht gezogen werden, da es diesbezüglich bereits Anfragen gibt (siehe auch Abschnitt 5.5 Modul Citizen Science).
3. In diesem Rahmen wird auch das Konzept zum zeitnahen Transfer und der Speicherung dieser Daten erweitert werden, um einfach zu nutzende und nachhaltige Lösungen zu finden, die insbesondere auf mittelgroßen (und ggf. kleinen) Forschungsschiffen anwendbar sind.

In der Pilotphase des Projektes (09/2019-12/2020) wurden nur ausgewählte Pilotdaten im Projekt behandelt und Arbeitsflüsse und technische Lösungen für diese entwickelt. In der aktuellen Projektphase I wird das Spektrum der Daten erweitert und Daten weiterer Sensoren verarbeitet.

Zusätzlich soll gemeinsam mit der Leitstelle Deutsche Forschungsschiffe und dem *Leibniz-Institut für Troposphärenforschung* (TROPOS) geprüft werden, ob auf dem FS SONNE und dem FS MARIA S. MERIAN ein s.g. Dust Monitor dauerhaft installiert werden kann. Die Blaupause für dieses Vorhaben wird noch in der Phase I des Projektes fertiggestellt, da ein Dust Monitor dauerhaft an Bord der METEOR installiert wird. Über den DAM-Datenfluss sollen bis zum Ende der laufenden Phase die Daten des Gerätes den Wissenschaftler:innen des TROPOS im DAM-Space bereitgestellt werden. Dort wird die Qualitätskontrolle übernommen bevor die Daten in PANGAEA publiziert und über das *Portal Deutsche Meeresforschung* zugreifbar werden. Des Weiteren sollen Daten der Radiosonden des *Deutschen Wetterdienstes* (DWD) integriert werden. Es soll ein Arbeitsablauf erarbeitet werden, der die Einspeisung der Daten in DSHIP und das MDM ermöglicht. Dies würde die Publikation der Daten durch PANGAEA ermöglichen.

Expeditionsmetadaten wie z.B. Fahrtteilnehmer:innen, Fahrtleitung, geplante Geräteeinsätze, zu erwartende Daten und Moratorien werden bisher nur dezentral und in sehr heterogenen *Excel*-Listen erfasst. Diese Listen sind weder zuverlässig maschinenlesbar noch leicht erreichbar.

Damit die Datenflüsse inkl. der Metadaten skalieren, ist ein Ziel, eine zentrale Erfassung dieser Metadaten in einem Datenbank-gestützten Informationssystem voranzutreiben. Die gemeinschaftliche Nutzung des *Ocean Science Information System* (OSIS), das am GEOMAR betrieben wird, soll geprüft werden. Aktuell kommt OSIS nur in Projekten unter GEOMAR-Beteiligung zum Einsatz. Die Öffnung dieser Infrastruktur für die gesamte DAM wäre beispielhaft im Prozess hin zu einer gemeinsamen dezentralen Dateninfrastruktur. Durch die Anwendung persistenter digitaler Identifikatoren in OSIS werden Metadaten und Daten in ihren Kontexten nachhaltig referenzierbar und stets aktuell nutzbar gemacht. Die maschinenlesbaren Metadaten sollen unter anderem für eine künftige semi-automatisierte Erstellung der Cruise Summary Reports durch das BSH verwendet werden und so deren zeitnahe Erstellung unterstützen.

4. Nach der erfolgreichen Harmonisierung der Parameternamen aller wissenschaftlichen Bordsensoren auf großen Forschungsschiffen ist der nächste Schritt in der jetzt beantragten Phase die Umsetzung einer dauerhaft nutzbaren, referenzierbaren Dokumentation des laufenden Geräte- und Sensorbetriebs an Bord der Schiffe. Diese soll künftig Veränderungen im laufenden Betrieb wie bei Wartungen, Kalibrierungen und tatsächliche Nutzung (periodischer Sensorwechsel etc.) erfassen und dadurch referenzierbar und somit nachvollziehbar machen.

Namen sowie Beschreibungen und Katalogisierung wissenschaftlicher Geräte für das *DSHIP ActionLog* werden weiterhin harmonisiert und kompatibel zu bestehenden Katalogen gehalten. Die Anbindung des *ActionLogs* an die vom AWI betriebene Sensordatenbank (*Sensors*) muss ausgebaut werden, um schrittweise alle Geräte im *Action Log* einzubinden, die in *Sensor* detailliert erfasst sind. Wie schon bei den Sensoren, wird über entsprechende technisch implementierte Referenzierung der Bezug zu erweiterten Metadaten des Meßbetriebs nutzbar. Mit der beabsichtigten Integration von bestehenden Werkzeugen (z.B. *Marine Facilities Planning Tool* (MFP) am GEOMAR) und lokalen Datenbanken an *Sensor* soll die Verknüpfung bereits beschriebener Sensoren und den damit zusammenhängenden Informationen gesichert und so redundante Arbeiten wie auch divergierende Informationen vermieden werden.

5. Neben der Harmonisierung des *ActionLogs* soll langfristig auch eine gewisse Verantwortung für das Führen des *ActionLogs* zurück in die Hände der Wissenschaft gegeben werden. So könnte eine Option sein, dass die Nautik weiterhin essentielle Einträge im *ActionLog* anlegt und darüber hinaus weitere, detaillierte Aufzeichnungen der wissenschaftlich relevanten Aktivitäten durch die Wissenschaft ergänzt werden. Dies soll künftig über die Mobile *ActionLog* App z.B. im Labor, an Deck oder auf dem Eis in Echtzeit erfolgen. Diese App wurde in der ersten Phase entwickelt und produktive Einsätze zeigen bereits, welches Potential eine solche App bei der Steigerung der Metadatenqualität bietet. Eine Weiterentwicklung und Operationalisierung der App werden daher angestrebt. Ergänzend sollen der Wissenschaft auch hierfür Trainings zur praktischen Anwendung angeboten werden.
6. Speicherkapazität an Land wird für Expeditionsdaten der Deutschen Forschungsflotte benötigt. Die koordinierte Erfassung von großvolumigen Massendaten wie Fächerecholot-Daten einschließlich der Wassersäulen- oder ADCP-Daten, erfordert eine große Speicherkapazität, welche in Zukunft zur Verfügung gestellt werden muss. Die benötigte Rechenleistung zur Verarbeitung der Daten muss angepasst werden, um sicherzustellen, dass Daten weiterhin durch Geräteverantwortliche und weitere Nutzer:innen aus der Wissenschaft direkt im DAM-Space verarbeitet und zur Publikation vorbereitet werden können (siehe Abschnitt 5.3, Position 5).

5.2. Erweiterung der Aktivitäten auf mittelgroße Forschungsschiffe

Für Expeditionsdaten, die an Bord mittelgroßer Forschungsschiffe erhoben werden, besteht im gleichen Maße die Notwendigkeit, sie Nutzer:innen bereitzustellen, wie für große Schiffe. Die Partner des DAM-Projektes "Unterwegs"-Forschungsdaten haben in den letzten Jahren viel Erfahrung in der Bereitstellung der Expeditionsdaten großer Forschungsschiffe gesammelt und konnten zeigen, dass eine systematische Bereitstellung der Daten, wie es der Wissenschaftsrat 2010 empfohlen hat, möglich ist. Diese Erfahrung kann und soll genutzt werden, um auch Expeditionsdaten mittelgroßer Forschungsschiffe auf ähnliche Weise aufzubereiten und entsprechend bereitzustellen. Der Fokus soll auf den Forschungsschiffen FS ALKOR, VWFS ATAIR, FS ELISABETH MANN BORGESE sowie FS HEINCKE liegen. Ziel ist, auch hier die bereits erreichte Harmonisierung und Standardisierung der einzelnen Prozessketten umzusetzen.

Eine direkte Übertragung der Erfahrungen aus dem DAM-Projekt "Unterwegs"-Forschungsdaten (Phase I) ist jedoch nicht ohne weiteres möglich. Das liegt u.a. daran, dass:

- der Betrieb der Forschungsschiffe im Vergleich zu den großen Forschungsschiffen anders strukturiert und organisiert ist;
- die vier, hier adressierten, mittelgroßen Forschungsschiffe von vier verschiedenen Einrichtungen (AWI, BSH, GEOMAR, IOW) betrieben werden und damit weitere Instanzen in die Abstimmungsprozesse eingebunden werden müssen;
- die mittelgroßen Forschungsschiffe nicht wie große Forschungsschiffe über einen wissenschaftlich-technischen Dienst (WTD) verfügen, der unmittelbar bei der Umsetzung technischer Maßnahmen an Bord unterstützt;
- die mittelgroßen Forschungsschiffe über unterschiedliche wissenschaftlich-technische Ausstattungen verfügen;
- i.d.R. während einer Seereise mit weniger wissenschaftlichem Personal agiert wird (~12 Wissenschaftler:innen pro Ausfahrt);
- die mittelgroßen Forschungsschiffe vornehmlich küstennah, in ausschließlichen Wirtschaftszonen operieren und der damit verbundene Aufwand für die Beantragung von "Unterwegs"-Messungen u.U. unverhältnismäßig sein könnte;
- die vorhandenen Datenflüsse von Bord an Land, Daten-Infrastrukturen und Messgeräte an Bord sehr heterogen sind;
- der Betrieb der betroffenen Messgeräte in Küstennähe oder allgemein in flachem Wasser neue Anforderungen zur Sicherstellung der erforderlichen

Datenqualität mit sich bringt. Dies ist gerätespezifisch und insbesondere der Fall bei profilierender Hydroakustik;

- Expeditionen mittelgroßer Forschungsschiffe i.d.R. kürzer und damit über ein Jahr gemittelt häufiger stattfinden als Expeditionen mit großen Forschungsschiffen. Der Aufwand für die Qualitätskontrolle, Kuration und Publikation einer zeitlich kurzen Expedition ist allerdings vergleichbar mit dem Aufwand einer zeitlich längeren Expedition eines großen Forschungsschiffes;
- es keine dem Vorgehen bei den großen Forschungsschiffen vergleichbare Fahrtvorbereitungstreffen und damit keine einfache Schnittstelle zum Informationsaustausch zwischen allen Beteiligten gibt. Dementsprechend erhöht sich die Anzahl der Akteure, mit denen Absprachen getroffen werden müssen.

Für eine Erweiterung der Aktivitäten ist es daher von Bedeutung, diese Unterschiede zu berücksichtigen, um die wissenschaftliche Verwertbarkeit bzw. die Anschlussfähigkeit der "Unterwegs"-Forschungsdaten, die an Bord von mittelgroßen Forschungsschiffen generiert werden können, zu gewährleisten und im Idealfall eine operationalisierbare Erweiterung zum regulären wissenschaftlichen Betrieb auf alle Expeditionsdaten zu erreichen. Dafür bedarf es der Kooperation mit der wissenschaftlichen Community, die bezüglich der FAIRen und offenen Bereitstellung öffentlich finanzierte Forschungsdaten für die Öffentlichkeit noch sehr heterogen aufgestellt ist. Nach wie vor bedarf es eines kulturellen Wandels, um Expeditionsdaten systematisch FAIR und offen bereitzustellen.

Positiv könnte sich die einfachere Zugänglichkeit des Projektteams zu den mittelgroßen Forschungsschiffen auswirken. Mittelgroße Forschungsschiffe sind häufiger in deutschen Häfen, nicht selten in unmittelbarer Nähe zu den betreuenden Institutionen. Sie können daher einfacher und schneller betreten werden, um z.B. technische Anpassungen durchzuführen oder fehlgeschlagene Datenübertragungen zu korrigieren. Abgesehen vom VWFS ATAIR werden die hier betrachteten Forschungsschiffe durch die Briese Schifffahrts GmbH & Co. KG bereedert, was den Abstimmungsaufwand bezgl. der Einsatzlogistik erleichtert.

5.2.1. Arbeitspaket 6: Vorbereitende Maßnahmen

Dieses Arbeitspaket wird durch alle AGs im Projekt bearbeitet. Das Arbeitspaket wird mit personellen Kapazitäten bearbeitet, die mit Zunahme der Automatisierung der Aktivitäten an großen Forschungsschiffen frei werden. Freie Kapazitäten variieren mit den AGs. Die AG-Technik wird aller Voraussicht nach zu Beginn der Phase II mehr freie Kapazitäten haben als die AG-Bathymetrie, AG-Ozeanographie oder PANGAEA-Mitarbeiter:innen. Dies hängt damit zusammen, dass die AG-Technik einen geringeren Anteil an Modifikations- und Wartungsaufgaben im Zusammenhang mit dem Betrieb und der Instandhaltung der aufgebauten Dateninfrastruktur hat, als die AG-Bathymetrie,

die AG-Ozeanographie sowie PANGAEA mit der Qualitätskontrolle und Kuration von Daten haben. Die AG-Technik wird daher dieses Arbeitspaket führend umsetzen und vom Projektkoordinator (Arbeitspaket 8) unterstützt.

Vorbereitende Maßnahmen für die Vereinheitlichung im Umgang mit Expeditionsdaten mittelgroßer Forschungsschiffe sowie die offene Bereitstellung der Daten nach den FAIR-Prinzipien sind:

1. Bestandsaufnahme

Da das Ziel die Vereinheitlichung im Umgang mit Expeditionsdaten der Deutschen Forschungsflotte ist, muss zunächst der Status Quo im Umgang mit Expeditionsdaten mittelgroßer Forschungsschiffe erfasst werden. Es soll ermittelt werden, welche Messinstrumente und Dateninfrastrukturen an Bord und an Land vorhanden sind. Die genutzten Datenflüsse müssen für jedes Schiff mit den jeweils betreibenden Einrichtungen erfasst und beschrieben werden. Prozesse im Zusammenhang mit der Expeditionsplanung, Durchführung und Abschluss sowie allen am Prozess beteiligten Akteuren muss für die vier Forschungsschiffe dokumentiert und analysiert werden. Zusätzlich muss der Bestand, z.B. an "historischen" Bathymetrie-Daten und deren Nutzbarkeit erfasst werden, um auch diese analog zu den "historischen" Bathymetrie-Daten großer Forschungsschiffe zu publizieren und der internationalen Gemeinschaft über den Map-Service (WMS) der *International Hydrographic Organization* (IHO) bereitzustellen.

Unter der Federführung des IOW wurde in der laufenden Projektphase I ein Fahrtantrag für das mittelgroße FS ELISABETH MANN-BORGESE (EMB) gestellt, welcher mittlerweile bewilligt wurde. Die Fahrt mit dem Titel *DAM Underway Data Acquisition and Provision Sea Trial 2022* (kurz *DAM UDAP 2022*) findet voraussichtlich im November 2022 statt und dient bereits dem oben genannten Ziel der Bestandsaufnahme. Ein zu erwartendes Ergebnis der Ausfahrt wird die Erkennung und Beschreibung konkreter, notwendiger Arbeitsschritte sein, die für die Übertragung der auf großen Forschungsschiffen etablierten Abläufe auf mittelgroße Forschungsschiffe nötig sind.

2. Konzeptentwicklung

Nach der Bestandsaufnahme und Analyse des Status Quo im Umgang mit Expeditionsdaten der mittelgroßen Forschungsschiffe soll die Übertragbarkeit vorhandener Lösungen geprüft werden, die an großen Forschungsschiffen bereits erprobt sind. Wie bereits beschrieben, ist davon auszugehen, dass die prozessualen und technischen Lösungen, die an Bord großer Forschungsschiffe bereits erfolgreich im Einsatz sind, für mittelgroße Forschungsschiffe noch modifiziert werden müssen. Eine konzeptionelle Anpassung wird aller Voraussicht nach erst erfolgen müssen. Das Ziel ist

es, in einem Konzept den betreibenden Institutionen Optionen darzulegen, die unterschiedliche Stufen einer Harmonisierung und Standardisierung in Hinblick auf die Erfassung und der zu erreichenden Weiterverarbeitungsmöglichkeiten der Expeditionsdaten darlegen.

Ein solches Konzept umfasst den gesamten Datenfluss vom Sensor bis zur Publikation in PANGAEA sowie die nutzer:innenfreundliche Bereitstellung über das *Portal Deutsche Meeresforschung*, in dem alle AGs involviert sind.

Ein großer Teil der ozeanographischen “Unterwegs”-Forschungsdaten wird aus Durchflussmesssystemen gewonnen. Ausgehend von der Situation auf den mittelgroßen Forschungsschiffen, setzt die methodisch konsequente, standardisierte Datenerhebung eine Vereinheitlichung der Messinfrastruktur voraus. Dementsprechend hat die Deutsche Allianz Meeresforschung dem BMBF im Jahr 2020 eine Antragskizze mit dem Titel *Infrastrukturprojekt zu Entwicklung und Bau standardisierter Messboxen für en route Messungen von Wasserparametern auf Deutschen Forschungsschiffen im Rahmen der Deutschen Allianz Meeresforschung (DAM)* vorgelegt. Diese Antragskizze ist als Anlage 1 beigelegt. Steht die Förderung der dort vorgeschlagenen, technisch optimalen Lösung nicht in Aussicht, müssen die verschiedenen, aktuell verfügbaren Messsysteme hinsichtlich ihrer Integration in den Datenfluss des DAM-Projektes “Unterwegs”-Forschungsdaten geprüft und entsprechende Übergangslösungen gefunden werden. Die Vereinheitlichung der Infrastruktur, wie im Antrag (Anlage 1) von 2020 beschrieben, wäre diesem Vorgehen natürlich vorzuziehen.

5.2.2. Arbeitspaket 7: Umsetzung der Maßnahmen

Dieses Arbeitspaket wird durch alle AGs im Projekt bearbeitet. Eine erfolgreiche Umsetzung des Arbeitspakets setzt voraus, dass personelle Kapazitäten entlang des gesamten Datenflusses vom Sensor bis zur Archivierung und Bereitstellung im Portal eingesetzt werden. Das Arbeitspaket wird mit personellen Kapazitäten bearbeitet, die mit Zunahme der Automatisierung der Aktivitäten an großen Forschungsschiffen frei werden. Wie bereits erwähnt, variieren freiwerdende Kapazitäten innerhalb der AGs.

Ziel dieses Arbeitspaketes ist die systematische Erfassung und Übertragung von Daten mittelgroßer Forschungsschiffe an Land, sowie die Qualitätssicherung und Veröffentlichung von Expeditionsdaten nach den *FAIR- and Open Data*-Prinzipien analog zu den großen Forschungsschiffen. In Abhängigkeit vom Stand der Entwicklungen werden in der Projektphase für einige Daten bzw. Systeme unterschiedliche Prozessierungs- und Verwertungspfade verfolgt werden müssen. So wird für ADCP- und CTD-Daten die unmittelbare Publikation in PANGAEA angestrebt, sobald die Prozeduren zur Qualitätssicherung auf den Schiffen bzw. landseitig etabliert sind.

Der genaue Umfang von AP7 wird im Zug der o.g. Vorarbeiten, siehe AP6, und basierend auf den Ergebnissen der Ausfahrt *DAM UDAP 2022* auf dem FS ELISABETH MANN BORGESE herausgearbeitet. Dies betrifft insbesondere die Feststellung, welche Daten im Sinne des DAM-Projektes "Unterwegs"-Forschungsdaten auf mittelgroßen Forschungsschiffen unter Berücksichtigung der jeweils zur Verfügung stehenden personellen Kapazitäten systematisch erfasst, bearbeitet und übertragen werden können.

Im beantragten Projektverlauf sollen jedoch mindestens die Expeditionsdaten eines mittelgroßen Forschungsschiffes behandelt werden. Weitere Forschungsschiffe sollen sukzessive hinzukommen. Zunächst soll die Umsetzbarkeit an Bord des FS ELISABETH MANN BORGESE behandelt werden, da die Expedition *DAM UDAP 2022* zu einem Erfahrungsvorsprung im Vergleich zu den anderen Forschungsschiffen führen wird. Ausgehend davon werden auf den übrigen Forschungsschiffen alle Vorkehrungen getroffen, um für die Umsetzung des Projektvorhabens einen gemeinsamen Standard zu erreichen.

Dies umfasst neben der Überprüfung der sensorischen Infrastruktur in Hinblick auf die Ausstattung von Messsystemen an Bord der mittelgroßen Forschungsschiffe auch die Identifikation eines gemeinsamen methodischen Standards unter Nutzung verfügbarer IT-Infrastrukturen. Hier existieren Unterschiede im Umgang mit teils vorhandenen Strukturen, etwa die Verwendung von DSHIP, die Übertragung von Daten an Land oder die Nutzung von Cruise-Assistance-Systemen. Um im Projektverlauf eine effiziente Umsetzung der Maßnahmen zu gewährleisten, sind Sachmittel zur Schaffung einer standardisierten IT-Infrastruktur auf den Forschungsschiffen vorgesehen (siehe Abschnitt 5.3 Sachmittel). Diese Sachmittel werden vom Verbundpartner IOW beantragt und im Sinne des Verbundes verwaltet. Die Zuarbeit und Unterstützung des IOW bei der Verausgabung der Sachmittel muss im besonderen Maße durch die AG-Technik erfolgen. Sachmittel, die vom IOW beantragt werden müssen teilweise für Infrastruktur an Bord der ATAIR, ALKOR und HEINCKE verwendet werden. Dies ist mit einem erheblichen administrativen Aufwand verbunden und kann zu Verzögerungen führen.

Damit werden die Schiffe aus informationstechnischer Sicht dahingehend ertüchtigt, dass sie die notwendigen Voraussetzungen zur Umsetzung der im Projekt entwickelten Maßnahmen erfüllen. Im Vergleich zur Entwicklung individueller Sonderlösungen für die einzelnen Schiffe sind die Aufwendungen für die veranschlagten Sachmittel geringer und erlauben zudem eine spätere Skalierung der Methoden.

Eine wesentliche Grundlage bildet die Verfügbarkeit einer qualitätsgesicherten Schiffsroute (*Mastertrack*) zur Validierung von Daten der Positionsgeber für die geographische Kopplung der an Bord erhobenen wissenschaftlichen Daten, was für die weitere Prozessierung und Sicherstellung der wissenschaftlichen Verwertbarkeit von

zentraler Bedeutung ist. Im Rahmen der Projektbearbeitung sollen die Prozesse und Schnittstellen weiter optimiert werden. Für den Betrieb werden an Bord der mittelgroßen Forschungsschiffe 45 T€ und an Bord der großen Forschungsschiffe 55 T€ für den gesamten Projektzeitraum veranschlagt (siehe Abschnitt 5.3, Sachmittel Pos. 1).

Zur adäquaten Prozessierung ist eine leistungsfähige Infrastruktur zur föderalisierten Datenspeicherung und Bereitstellung der Expeditionsdaten auf den Schiffen erforderlich. Diese Aufgabe übernimmt im Kontext der "Unterwegs"-Forschungsdaten an Bord der großen Forschungsschiffe bereits ein *Massendatenmanagement-System* (siehe auch Abb. 1). Perspektivisch ist die Schaffung einer standardisierten Architektur für die gesamte Forschungsflotte vorgesehen, sodass für die Installation der MDM Server an Bord der mittelgroßen Forschungsschiffe 40 T€ veranschlagt werden (Abschnitt 5.3, Sachmittelt Pos. 2).

Ferner werden Sachmittel in Höhe von 75,5 T€ für die erforderlichen Softwarekomponenten und Lizenzen der MDM Server benötigt, um im Zeitraum der Projektbearbeitung auf den adressierten mittelgroßen Forschungsschiffen eine standardisierte Arbeitsumgebung aufbauen und den Betrieb gewährleisten zu können (Abschnitt 5.3, Sachmittelt Pos. 3). Dies erlaubt, Entwicklungen der einzelnen AGs schnell und effizient für Testzwecke auszurollen und in den operationellen Betrieb der Forschungsschiffe überführen zu können.

Für die bedarfsgerechte und kontinuierliche Anpassung der Software an die besonderen Anforderungen im Schiffsbetrieb sowie die Pflege und Wartung der Server, hat sich der Abschluss von Wartungsverträgen als effizient und ökonomisch erwiesen, sodass auch hier für die Projektlaufzeit entsprechende Sachmittel vorgesehen sind. Da die Server-Infrastrukturen auf den großen Schiffen bereits implementiert sind und vorgesehene Anpassungen und Wartungen für alle Schiffe verfügbar sein werden, umfassen die vorgesehenen Sachmittel alle im Rahmen des DAM-Projekts betreuten Forschungsschiffe. Für die mittelgroßen Forschungsschiffe sind daher Ausgaben in Höhe von 27 T€ zu erwarten bzw. 48k T€ für die großen Forschungsschiffe (Abschnitt 5.3, Sachmittel Pos. 4). Dadurch wird sichergestellt, dass die Entwicklungen allen Schiffen mit entsprechendem Rahmenvertrag zur Verfügung stehen (analog zu DSHIP).

Neben der Infrastruktur an Bord der Schiffe ist auch die landseitige Anbindung des MDM von entscheidender Bedeutung für die Umsetzbarkeit der durch die DAM formulierten Ziele zur Bereitstellung der Expeditionsdaten. Aus diesem Grund soll eine gemeinsame Cloud-Infrastruktur im Rahmen der Projektbearbeitung bereitgestellt werden, die neben der Speicherung der Daten auf Servern an Land insbesondere die Bereitstellung der Daten über Webservices sicherstellt. Ohne eine solche Service-Infrastruktur wird die Umsetzung einer nutzer:innenfreundlichen Verfügbarkeit der Daten nicht nur erschwert, auch die Akzeptanz zur Nutzung der Daten wird reduziert, etwa wenn Schnittstellen

(APIs) fehlen. Die dafür veranschlagten Kosten belaufen sich auf 30 T€ (Abschnitt 5.3, Sachmittel Pos. 5).

Mit Schaffung der oben genannten Infrastrukturen ist davon auszugehen, dass ein Großteil der an Bord der mittelgroßen Forschungsschiffe gewonnenen Expeditionsdaten entsprechend der Motivation des DAM-Projektes weiterverarbeitet bzw. *FAIR* bereitgestellt und nachgenutzt werden kann. Eine wesentliche Herausforderung des hier beantragten Projektvorhabens wird jedoch darin bestehen, die dafür nötigen Prozesse so zu optimieren bzw. zu entwickeln, dass sie als niederschwellige Dienste (scientific services) Einzug in die wissenschaftlichen Abläufe an Bord der mittelgroßen Forschungsschiffe halten. Daher ist die Bereitstellung einer leistungsfähigen, einheitlichen IT-Infrastruktur insbesondere für die mittelgroßen Forschungsschiffe wichtig, um auch ohne wissenschaftlich-technischen Dienst an Bord die komplexen Vorgänge dauerhaft gewährleisten zu können.

Im Fall des Betriebs von ADCPs auf mittelgroßen Schiffen sind von dem Geräteverantwortlichen zur Erfüllung der Qualitätsstandards und zum Testen der Erfassung schiffsspezifischer Navigations-, Gyro- und weiterer Metadaten, Arbeiten auf den Schiffen zu leisten. Voraussetzung dafür ist die Ausstattung des Teilprojekts mit Sachmitteln zur Beschaffung von Messrechnern und Zubehör zum Betrieb der im flacheren Wasser genutzten, hochfrequenten und nicht fest installierten ADCPs (Abschnitt 5.3, Sachmittel Pos. 6).

Bathymetrische Daten betreffend, lässt sich nach derzeitiger Einschätzung sagen, dass die systematische Publikation in PANGAEA, Auffindbarkeit und Zugriff auf Bathymetrie-Daten mittelgroßer Forschungsschiffe über das *Portal Deutsche Meeresforschung / PANGAEA* durch die Nutzung der in Phase I des DAM-Projektes "Unterwegs"-Daten aufgebauten Strukturen gewährleistet werden könnte. Eine Bestandsaufnahme "historischer" Daten muss dennoch vollzogen werden, um den Umfang der Datenbestände zu quantifizieren und die Anzahl möglicher Datenpublikationen adäquat zu prognostizieren.

5.2.3. Arbeitspaket 8: Projektkoordination

Das Projekt wird weiterhin vom Leiter der Kernbereiche *Datenmanagement und Digitalisierung* sowie *Koordinierung der Infrastrukturen* koordiniert. Es sei an dieser Stelle darauf hingewiesen, dass die adäquate Erfüllung der Leitungsposition in beiden Kernbereichen sowie die Projektkoordination ein VZÄ übersteigt. Die angemessene Koordination des hier beantragten Verbundvorhabens (Phase II), das zehn Partner und ein Finanzvolumen von etwa 1.5 Mio. € pro Jahr (inkl. HGF) umfasst, erfordert die Koordinationsunterstützung des mittleren Managements der Partnereinrichtungen als in kind Beitrag.

Die Koordinationsaufgaben müssen besonders hinsichtlich des Schwerpunktes 2 ausgeweitet werden, um alle Beteiligten am Betrieb der mittelgroßen Forschungsschiffe adäquat einzubinden. Darüber hinaus muss für mehr "Sichtbarkeit" der Errungenschaften im Zusammenhang mit dem Umgang von Expeditionsdaten der Deutschen Forschungsflotte innerhalb der wissenschaftlichen Community gesorgt und stärker für die Nutzung der aufgebauten (Infrastruktur-)Strukturen geworben werden. Die Vernetzung mit internationalen Aktivitäten zu "Unterwegs"-Forschungsdaten muss aufgebaut werden. Der kulturelle Wandel muss aktiv gefördert werden. Diese Aufgaben werden nach Möglichkeit erfüllt.

Das Arbeitspaket 8 umfasst folgende Aufgaben:

- Projektplanung und Begleitung
- Koordination der Einhaltung der Zielvorgaben des Projekts
- Koordination der Passfähigkeit der Arbeitspakete untereinander
- Kommunikation mit allen internen und externen Projektbeteiligten
- Bedienung der organisatorischen Schnittstellenfunktion zwischen den AGs, der Leitstelle Deutsche Meeresforschung, den Reedereien, den schiffsbetreibenden Einrichtungen sowie internen und externen Partnern und der DAM-Geschäftsstelle
- Unterstützung bei der Arbeitsvorbereitung und Planung des Ressourceneinsatzes
- Projektcontrolling
- Qualitätskontrolle
- Berichterstattung
- Dokumentation der Arbeitsergebnisse
- Organisation, Dokumentation und Protokollierung der geplanten Projektveranstaltungen (Kick-Off, Projekttreffen zur Vorbereitung der Evaluation, etc.)
- Koordination der Vernetzung und Anschlussfähigkeit der Aktivitäten an weitere nationale und internationale Datenmanagement-Initiativen (*Marine Dateninfrastruktur Deutschland (MDI.de)*, *MareHub*, *Helmholtz Metadata Collaboration Plattform (HMC)*, *Nationale Forschungsdateninfrastruktur (NFDI)*, *Global Ocean Observing Systems (GOOS)*, *IOC / UNESCO*, etc)
- Vertretung der DAM und des Projektes nach innen (DAM-Gremien) und außen
- Förderung des kulturellen Wandels durch Infoveranstaltungen für die Wissenschaft
- Moderation und Vermittlung bei Uneinigkeiten und Konflikten

5.3. Sachmittel

Für die Umsetzung der oben genannten Maßnahmen und Erweiterung der Dienste und Infrastrukturen zur Erhebung und Bereitstellung von “Unterwegs“-Forschungsdaten auf Deutschen Forschungsschiffen sind zur erfolgreichen Operationalisierung Sachmittel einzuplanen (Tabelle 1) welche in Abschnitt 5.2.2, AP 7 begründet sind. Anhand der fachlichen Diskussionen in den einzelnen Arbeitsgruppen und bisherigen Aktivitäten im Projekt ist deutlich geworden, dass die Einbindung der mittelgroßen Forschungsschiffe nur durch eine vorhergehende Harmonisierung sowohl der an Bord befindlichen Systeme als auch der landseitig vorgehaltenen Infrastruktur erfolgreich umgesetzt werden kann. Die im Projektvorhaben adressierten mittelgroßen Forschungsschiffe unterscheiden sich derzeit sehr stark hinsichtlich Ausstattung an Bord und landseitiger (Daten-) Anbindung, sodass ausgehend von den Erfahrungen mit den großen Forschungsschiffen eine homogene Struktur über die gesamte Flotte erreicht werden soll. Die notwendige Ausrüstung, Hardware-Erweiterung und Anpassungen der Software sowie der Arbeitsabläufe soll durch zentral Koordination unterstützt werden (siehe Arbeitspaket 8). Die Sachmittel werden als Bestandteil des Teilprojektes beantragt, das am IOW bearbeitet wird. Das IOW bietet sich als Verbundpartner und betreibende Einrichtung des mittelgroßen Schiffs ELISABETH-MANN-BORGESE für die Verwaltung der Sachmittel an. Technische und infrastrukturelle Anpassungen, die als Ergebnis der Expedition DAM-UDAP 2022 an Bord der EMB vorgenommen werden sollen, können so unmittelbar und mit vergleichsweise wenig administrativem Aufwand umgesetzt werden. Eine Weiterleitung der Sachmittel an die Einrichtungen AWI, BSH und GEOMAR aller Voraussicht nach nötig, um die Infrastruktur an Bord der ATAIR, ALKOR und HEINCKLE entsprechend aufzubauen.

Geplanten Sachmittel setzen sich aus folgenden Komponenten zusammen, die nicht aus Mitteln der Verbundpartner bestritten werden können:

Position	Sachmittelposten	Kosten (in T€ für 3 Jahre)		Summe (in T€ für 3 Jahre)
		große FS	mittelgroße FS	
1	Mastertracks	45	55	100
2	MDM Server Hardware	-	40	40
3	MDM Server Software und Lizenzen	-	57,5	57,5
4	MDM Wartungsverträge	48	27	75

Position	Sachmittelposten	Kosten (in T€ für 3 Jahre)		Summe (in T€ für 3 Jahre)
		große FS	mittelgroße FS	
5	MDM Landseitig Cloud Speicher und Webservices für 3 Jahre	15	15	30
6	TP ADCP Ausstattung	5	5	10
Summe (T€)		113	199,5	312,5

Tabelle 1: Sachmittelposten für die Vereinheitlichung der Daten-Infrastruktur der Deutschen Forschungsflotte kalkuliert über die Gesamtlaufzeit von 3 Jahren.

5.4. Strukturplan

Der Strukturplan (Abb. 2) stellt den Zusammenhang zwischen den Arbeitsgruppen und den Arbeitspaketen dar. Die starke Vernetzung der Arbeitsgruppen ergibt sich durch die vernetzte Beteiligung der Partnereinrichtungen an jeweils mehreren Arbeitspaketen (Tabelle 2).

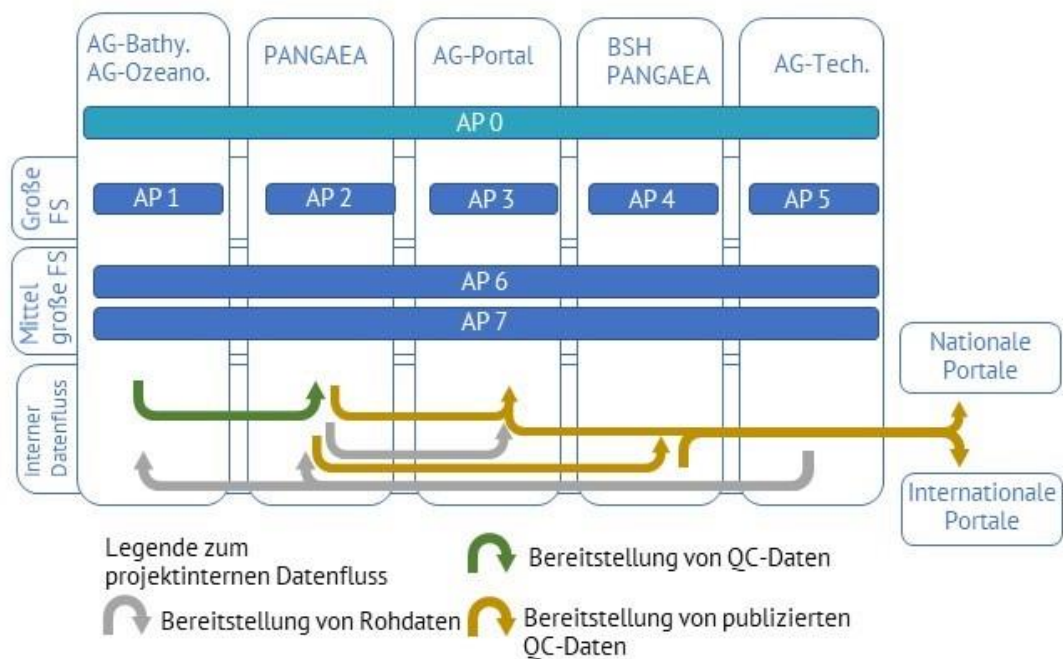


Abbildung 2: Modifiziert nach Abbildung 4 aus dem Leittrag der Phase I zur Darstellung der starken Vernetzung der Arbeitsgruppen.

Partner	Arbeitspaket
AWI	AP 2,3,5,6,7
BSH	AP 4
CAU/KMS	AP 1,6,7
DAM	AP 8
GEOMAR	AP 1,5,6,7

Partner	Arbeitspaket
Hereon	AP 1,6,7
ICBM	AP 1,6,7
IOW	AP 1,6,7
MARUM	AP 1,2,6,7
MPI-Chemie	AP 1,6,7

Tabelle 2: Darstellung der Partnereinrichtungen und der Arbeitspakete an denen sie beteiligt sind. Fett gekennzeichnete Arbeitspakete werden von der jeweiligen Partnereinrichtung geleitet.

5.5. Meilensteinplan

Im Folgenden werden die Meilensteine der einzelnen Arbeitspakete tabellarisch dargestellt. Der Meilenstein 0 jedes Arbeitspaketes umfasst die Aufgaben im Zusammenhang mit der Fortsetzung, Optimierung und Erweiterung der Tätigkeiten, die in Phase I des Projektes aufgebaut werden. Die ausführlichere Darstellung der Inhalte ist jeweils unter den Stichpunkten innerhalb der Beschreibung der Arbeitspakete zu finden. Die folgenden Meilensteine (Nummerierung 1-x) umfassen jeweils neue Aktivitäten. Die Nummerierung der Meilenstein im Meilensteinplan bezieht sich auf die Nummerierung der neuen Aufgabenstellungen, die ebenfalls inhaltlich detaillierter in innerhalb der Beschreibung der Arbeitspakete zu finden sind.

AP	AG	Meilenstein	2023	2024	2025
1	Bathymetrie	0: Fortsetzung der Aufzeichnung und Publikation von "Unterwegs"-Bathymetrie auf den großen Forschungsschiffen und kontinuierliche Aktualisierungen der Karten in DSHIP			
		0: Datenkuration eingehender Roh- sowie qualitätskontrollierter Bathymetrie-Daten (inkl. historischer Daten) in PANGAEA und gleichzeitige Aufbereitung der Daten für den PANGAEA Bathymetrie WMS			
		1: Konzept zur systematischen Publikation bathymetrischer Daten deutscher Forschungsschiffe			
		2: Erweiterung der Datenverfügbarkeit von Bathymetrie Daten im <i>Portal Deutsche Meeresforschung</i> (via PANGAEA Bathymetrie Web Map Service)			
		3: Visualisierung von Daten mit dem SEG-Y Datenformat als Tracklinie im <i>Portal Deutsche Meeresforschung</i> + Georeferenzierung der Daten für PANGAEA			

AP	AG	Meilenstein	2023	2024	2025
		4: Konzeptionierung eines Werkzeuges für PANGAEA zur Qualitätssicherung von Raster Daten und vereinfachten Datenkuration			
		5: Konzeptionierung einer nutzer:innenfreundlichen Cloud Umgebung zur Bearbeitung von Fächerecholot-Daten für die Wissenschaft			
		6: Sukzessive Publikation aller prozessierten Fächerecholot Rohdaten aus dem Datenbestand des Alfred-Wegener-Institutes			
		7: "near real-time" Visualisierung der Fächerecholot Datenabdeckung in DSHIP auf den großen Forschungsschiffen			
		8. Erstellung einer Karte für DSHIP, die die globale Datenabdeckungen aus verschiedenen Bathymetrie Datenarchiven enthält			
		9. Integration von prozessierten Fächerecholot Daten in die Global Multi-Resolution Topography Data Synthesis (GMRT)			
		10. Workshop: nachhaltige Bathymetrie-Datenaufnahme			
	Ozeano-graphie	0: Fortführung Akquise, QC, und Publikation ozeanographischer "Unterwegs"-Forschungsdaten			
		1: Internationale Portale I: Gerätespezifische Anpassung der Datenformate zur Veröffentlichung abgeschlossen			

AP	AG	Meilenstein	2023				2024				2025			
		2: Internationale Portale II: Datenveröffentlichung ist in den DAM-Datenfluss integriert												
		3: Datenbehandlung I: Anwendung von Machine Learning Methoden ist evaluiert												
		4: Datenbehandlung II: Darstellung ozeanographischer Daten im <i>Portal Deutsche Meeresforschung</i> ist vereinheitlicht												
		5: Datenbehandlung III: Sekundäre Qualitätskontrolle für Durchflussmessungen und CTD ist integriert												
		6: ADCP Daten I: Prozessierung zur Eliminierung akustischer Interferenzen sind implementiert												
		7: ADCP Daten II: Das akustische Rückstreuungssignal ist als DAM-Datenprodukt etabliert												
		8: TSG NRT: Überführung der NRT Prozessierung von der GEOMAR-Cloud zum DAM-Space ist abgeschlossen												
		9: Sensorintegration I: Integration von Lichtfelddaten inklusive QC ist evaluiert												
		10: Sensorintegration II: Integration von CO ₂ inklusive QC ist evaluiert												
		11: Sensorintegration III: Integration der „Dust-Monitors“ inkl. QC ist evaluiert												

AP	AG	Meilenstein	2023	2024	2025
		12: Wissenschaft 1: User-Workshops zur Nutzung von "Unterwegs"-Daten			
		13: Wissenschaft 2: Proaktive Rolle zur direkten Anwendung von "Unterwegs"-Forschungsdaten, Potentiale werden evaluiert und entsprechende wissenschaftliche Projekte werden identifiziert			
2	PANGAEA	0: Archivierung von Expeditions-Metadaten, <i>Mastertracks</i> und qualitätskontrollierten Daten und Rohdaten der großen Forschungsschiffe sowie der DAM Forschungsmissionen			
		1: Systematische Publikation bathymetrischer Daten deutscher Forschungsschiffe (siehe AP1 AG Bathymetrie 1)			
		2: Kuration und Georeferenzierung von Datensätzen der Sediment Echolote (siehe AP1 AG Bathymetrie 3)			
		3: Semi-automatische Prüfung von Rasterdaten für die Daten-Publikation in PANGAEA (siehe AP1 AG Bathymetrie 4)			
		4: Archivierung aller prozessierten Bathymetrie-Daten aus dem Datenbestand des Alfred-Wegener-Instituts (siehe AP1 AG Bathymetrie 6)			
		5: Konzeptionierung für semi-automatische Rohdaten-Archivierung von ADCP und CTD Daten			

AP	AG	Meilenstein	2023	2024	2025
		6: Konzeption von SOPs für die Datenkuration neuer Sensoren und Archivierung der Daten neuer Sensoren			
		7: Verknüpfung der Sensor Daten mit PANGAEA Datensätzen			
		8: Recherche und Inventarisierung des PANGAEA Archivs (siehe AP4 AG BSH 1)			
		9: Gewährleistung der Auffindbarkeit dezentraler Daten aus den DAM Forschungsmissionen im <i>Portal Deutsche Meeresforschung</i>			
		10: Weiterbildungsangebote zur Archivierung und Auffindbarkeit von DAM Daten in PANGAEA bzw. dem <i>Portal Deutsche Meeresforschung</i>			
3	Portal	0: Betrieb des Portals			
		1: Ausbau des Portals			
		2: Review und Optimierung der Portal-Inhalte und Portal-Struktur für die Findbarkeit in Suchmaschinen			
		3: Einbindung weiterer near real-time Daten in das Portal			
		4: Konzeption und Implementierung von Funktionen zum zusammenfassenden Herunterladen von Datensätzen			

AP	AG	Meilenstein	2023				2024				2025			
		5: Review und Optimierung des Portals für mobile Endgeräte												
		6: Planung und Durchführung von Nutzer:innenworkshops und Usability-Studien												
		7: Ausbau der Dokumentation mit Hilfestellungen und Beispielen für Nutzer:innen												
		8: Review, Konzeptionierung und Implementierung von verbesserter Skalierbarkeit der Geodateninfrastruktur												
4	BSH	0: Fortlaufender Betrieb des <i>Datenbrokers</i>												
		1: Inventarisierung geeigneter Daten innerhalb von PANGAEA für den <i>Datenbroker</i>												
		2: Anbindung an weitere Datenportale												
		3: Proof of Concept - "Data tracking"												
		4: Identifikation und Erschließung weiterer Datenquellen für den Data-Broker												
5	Technik	0: Fortlaufender Betrieb, Wartung der Dateninfrastruktur, Aufrechterhaltung der Datenflüsse												

AP	AG	Meilenstein	2023	2024	2025									
		0: Pflege der Metadaten, Interoperabilität wissenschaftlicher Sensoren und Geräte												
		0: Bereitstellung qualitätskontrollierter <i>Mastertracks</i>												
		1: Dokumentation und transparente Darstellung der aufgebauten Abläufe; Erstellung von SOPs.												
		2: Übertragbarkeit der SOPs aus 1. auf mittelgroße und kleine Schiffe												
		3: Sicherstellung der technischen Skalierbarkeit (4 mittelgroße Schiffe)												
		4: Technische Voraussetzung für eine Datenübertragung über den DAM-Datenfluss von Dust Monitor-Daten auf SONNE und MSM geschaffen.												
		5: Nutzung von OSIS zur Erfassung von Expeditionsmetadaten und maschinelle Unterstützung der Erstellung von CSRs												
		6: Systematische Kuration und Integration der verfügbaren Sensor- und Geräteinformationen via sensor.awi.de regulär nutzbar												
		7: Anwendung der <i>ActionLog</i> -App durch die Wissenschaft												
		8: Speicherkapazitäten erweitert und verfügbare Rechenleistung gesichert												

AP	AG	Meilenstein	2023	2024	2025
6	Alle	1: Bestandsaufnahme			
		2: Konzept zur Erfassung und Übertragung von Daten der Deutschen Forschungsflotte an Land, sowie der Qualitätssicherung und Veröffentlichung von Expeditionsdaten nach den <i>FAIR- and Open Data</i> -Prinzipien analog zu den großen Forschungsschiffen besteht.			
7	Alle	1: Umsetzung des o.g. Konzeptes.			
8	DAM	Projektkoordination			

Tabelle 3: Meilensteinplan für die Projektphase II

5.6. Abgrenzung des Projektes zur NFDI und zum Datenmanagement innerhalb der DAM-Forschungsmissionen

Die Nationale Forschungsdateninfrastruktur (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. Es soll ein dauerhafter digitaler Wissensspeicher als unverzichtbare Voraussetzung für neue Forschungsfragen, Erkenntnisse und Innovationen geschaffen werden. Relevante Daten sollen nach den FAIR-Prinzipien (Findable, Accessible, Interoperable und Reusable – also auffindbar, zugänglich, interoperabel und wiederverwendbar) zur Verfügung gestellt werden.

Die NFDI ist dabei als übergeordnete, nationale und disziplinübergreifende Koordinationsstruktur zu verstehen. 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 im übergeordneten Rahmen der NFDI. 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. Die Nutzung der Architektur und Einhaltung der Standards bleibt Aufgabe der Wissenschaft.

Das Projekt „Unterwegs“-Forschungsdaten ist ein beispielhaftes und beispielgebendes Projekt, das ein Kernziel der NFDI, nämlich die Bereitstellung FAIRer und offener Forschungsdaten, bezogen und begrenzt auf „Unterwegs“-Forschungsdaten (bzw. Expeditionsdaten) der großen deutschen Forschungsschiffe erfüllt. Der dazu notwendige Aufbau von Dateninfrastruktur und Datenflüssen, inkl. Qualitätskontrollen der Daten und Publikation in einem zertifizierten Langzeitarchiv (PANGAEA) sowie die Bereitstellung im u.a. zu diesem Zweck aufgebauten Portal Deutsche Meeresforschung sind Maßnahmen die als Beitrag zur NFDI zu verstehen sind, aber nicht durch die NFDI finanziert oder umgesetzt werden können.

Das Datenmanagement innerhalb der DAM-Forschungsmission ist primär Aufgabe der Wissenschaft. Zur personellen Unterstützung der Wissenschaft wurden im Rahmen der DAM-Forschungsmissionen geringfügig Mittel akquiriert. Im Projekt „Unterwegs“-Forschungsdaten wird eine Stelle bereitgestellt, die dafür Sorge trägt, dass die Datenkuration von DAM-

Forschungsmission seitens PANGAEA standardisiert wird. Diese Standardisierung schließt aktuell Daten von DAM-Forschungsmissionen aus, die in anderen Repositorien abgegeben werden. Zusätzlich wird durch diese Personalstelle für die Publikation von DAM-Forschungsdaten bei PANGAEA geworben und so der kulturelle Wandel vorangetrieben. Es werden Workshops und Seminare zum Thema Forschungsdatenmanagement und Datenpublikation bei PANGAEA angeboten. Wissenschaftler:innen der DAM-Forschungsmissionen erhalten das Angebot, bei der Erstellung von Datenmanagementplänen und Forschungsdatenrichtlinien unterstützt zu werden. Durch das Projekt „Unterwegs“-Forschungsdaten wird dafür gesorgt, dass Daten von Expedition, die im Rahmen von DAM-Forschungsmission stattfinden, in PANGAEA standardisiert mit den entsprechenden Metadaten beschrieben werden.

6. Verwertungsplan

Alle im Projekt behandelten Daten werden in PANGAEA publiziert. Die in PANGAEA langzeitgesicherten Projektdaten stehen nach Projektende, bzw. nach einer zu bestimmenden Moratoriumszeit der Wissenschaft frei online zur Verfügung. Jeder Datensatz ist mit einer DOI (Digital Object Identifier) versehen und damit eindeutig referenziert und zitierbar. Die im Projektverlauf entwickelte Software basiert weitestgehend auf Open Source Komponenten. Softwarekomponenten werden daher entweder unter einer Open Source Lizenz der Allgemeinheit zur Verfügung gestellt oder innerhalb der Infrastruktur verwendet. Eine kommerzielle Verwertung von Software oder anderen Projektergebnissen ist nicht geplant.

7. Verbundweiter Reiseplan

Projektpartner: **A** BSH **B** CAU/KMS, **C** ICBM, **D** IOW, **E** MARUM, **F** MPI-C,
Teilnehmeranzahl je Projektpartner, beantragt + ggf. selbstfinanziert/kostenfrei

Nr.	Jahr	Zielort	Zweck	DAM	A	B	C	D	E	F
Inland										
1	Feb. 2023	Norddeutshl.	Projekttreffen Kick-off	0+1	1+1	1+1	1+1	1+1	4+2	1+1
2	Jun. 2024	Norddeutshl.	Zwischenstatus	0+1	1+1	1+1	1+1	1+1	4+2	1+1

3	Sept. 2025	Norddeutschl.	Projektabschluss	0+1	1+1	1+1	1+1	1+1	4+2	1+1
4	2023 a	Norddeutschl.	PANGAEA-Seminar					1+0		1+0
5	2023 b	Norddeutschl.	PANGAEA-Seminar							1+0
6	2023 c	Norddeutschl.	PANGAEA-Seminar							1+0
7	2023	Bremerhaven, Kiel oder Geesthacht	Data Science Symposium	0+1			1+1		1+1	
8	2024 a	Norddeutschl.	PANGAEA-Seminar					1+0		1+0
9	2024 b	Norddeutschl.	PANGAEA-Seminar							1+0
10	2024 c	Norddeutschl.	PANGAEA-Seminar							1+0
11	2024	Bremerhaven, Kiel oder Geesthacht	Data Science Symposium	0+1			1+1		1+1	
12	2025 a	Norddeutschl.	PANGAEA-Seminar					1+0		1+0
13	2025 b	Norddeutschl.	PANGAEA-Seminar							1+0
14	2025 c	Norddeutschl.	PANGAEA-Seminar							1+0
15	2025	Bremerhaven, Kiel oder Geesthacht	Data Science Symposium	0+1			1+1		1+1	
16	2025	Norddeutschl.	DHyG Fachtagung						1+1	
Ausland										
17	2023	Ocean Data Conference	Konferenzteilnahme					1+1		
18	2023	EGU	Konferenzteilnahme			1+1			1+1	
19	2023	MAP the GAPS	Konferenzteilnahme						1+1	
20	2023	CTD Workshop Askö Schweden	Konferenzteilnahme					1+3		
21	2023	International	Seereise Forschungsschiff Systemtest					1+0		