

# HENRY

Hydraulic Engineering Repository

Ein Service der Bundesanstalt für Wasserbau

---

Article, Published Version

**Schäfer, Angela; Koppe, Roland**

## **Das Marine Netzwerk für Integrierten Datenzugriff und das Datenportal Deutsche Meeresforschung**

Die Küste

Zur Verfügung gestellt in Kooperation mit/Provided in Cooperation with:  
**Kuratorium für Forschung im Küsteningenieurwesen (KFKI)**

---

Verfügbar unter/Available at: <https://hdl.handle.net/20.500.11970/106354>

Vorgeschlagene Zitierweise/Suggested citation:

Schäfer, Angela; Koppe, Roland (2018): Das Marine Netzwerk für Integrierten Datenzugriff und das Datenportal Deutsche Meeresforschung. In: Die Küste 86. Karlsruhe: Bundesanstalt für Wasserbau. S. 21-32.

### **Standardnutzungsbedingungen/Terms of Use:**

Die Dokumente in HENRY stehen unter der Creative Commons Lizenz CC BY 4.0, sofern keine abweichenden Nutzungsbedingungen getroffen wurden. Damit ist sowohl die kommerzielle Nutzung als auch das Teilen, die Weiterbearbeitung und Speicherung erlaubt. Das Verwenden und das Bearbeiten stehen unter der Bedingung der Namensnennung. Im Einzelfall kann eine restriktivere Lizenz gelten; dann gelten abweichend von den obigen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Documents in HENRY are made available under the Creative Commons License CC BY 4.0, if no other license is applicable. Under CC BY 4.0 commercial use and sharing, remixing, transforming, and building upon the material of the work is permitted. In some cases a different, more restrictive license may apply; if applicable the terms of the restrictive license will be binding.



# Das Marine Netzwerk für Integrierten Datenzugriff und das Datenportal Deutsche Meeresforschung

*Angela Schäfer und Roland Koppe*

## Zusammenfassung

Für die marine Forschung ist die Verknüpfung von interdisziplinären Daten essentiell. Um dahingehend einen zentralen und leichten Zugriff auf die existierenden und vielfältigen, jedoch verteilten Daten der marinen Forschung zu ermöglichen, sind gut funktionierende nationale und internationale Infrastrukturen fundamental wichtig.

In diesem Sinne entwickelt das „Marine Network for Integrated Data Access“ (MaNIDA) einen nationalen und vernetzten Ansatz für die Auffindbarkeit und den Zugriff auf verteilte marine Forschungsdateninfrastrukturen mit entsprechenden Datenmanagementstrategien und Arbeitsabläufen. Um sowohl eine kohärente Datenauffindbarkeit, Visualisierung und einen einfachen Datenzugriff als auch die Veröffentlichung von wissenschaftlichen Daten und Publikationen zu ermöglichen, entwickelt das MaNIDA-Konsortium das zentrale „Datenportal Deutsche Meeresforschung“. Durch ein zentrales, automatisiertes Harvesting-Verfahren und standardisierte Schnittstellen verknüpft dieses Datenportal unterschiedliche und verteilte Datenquellen. In diesem Artikel werden sowohl die speziellen Aspekte des Portals zum Thema Inhalte, Funktionalität, Dienste, Architektur, Schnittstellen und Standards als auch die beitragenden Datenanbieter vorgestellt.

## Schlagwörter

Dateninfrastruktur, Datenportal, marine Forschungsdaten, Datenzugriff, Datenabfrage, Portalarchitektur, Standards, Schnittstellen, Webdienste, Datenanbieter, Datenprovider, Datenmanagement, Datenarbeitsabläufe, Datenpublikation, Datenzierbarkeit

## Summary

*The linkage of data beyond disciplinary boundaries is essential for marine research. Sufficient national and international data infrastructures are fundamental for central and easy access to the variety of existing, but distributed datasets in marine science. The Marine Network for Integrated Data Access (MaNIDA) provides a national networked approach in accessing and mining of federated marine research data infrastructures together with data management strategies and data workflows. In that course the MaNIDA consortium conceptualized and developed the “Data Portal German Marine Research” for coherent discovery, view, download and dissemination of scientific data and publications. The data portal is based on a central harvesting and interfacing approach by connecting distributed data sources. Here we inform about the specific details of the portal in terms of content, functionality, services, architecture, interfaces, standards and the contributing data providers.*

## Keywords

*data infrastructures, data portal, marine research data, data access, data retrieval, portal architecture, standards, interfaces, web services, data providers, data management, data workflows, data publication, data citation*

## Inhalt

|     |                                                                   |    |
|-----|-------------------------------------------------------------------|----|
| 1   | Einleitung .....                                                  | 22 |
| 2   | Das MaNIDA-Konsortium.....                                        | 23 |
| 2.1 | Entwicklung von Datenarbeitsabläufen und Datenpflege .....        | 24 |
| 3   | Das „Datenportal Deutsche Meeresforschung“ .....                  | 25 |
| 3.1 | Inhalt und Funktionen .....                                       | 25 |
| 3.2 | Mehrwertdienste.....                                              | 26 |
| 3.3 | Datenanbieter .....                                               | 27 |
| 3.4 | Architektur .....                                                 | 28 |
| 3.5 | Schnittstellen und Standards .....                                | 29 |
| 3.6 | Datenzugriffsbestimmungen und gute wissenschaftliche Praxis ..... | 30 |
| 4   | Dankeswort.....                                                   | 30 |
| 5   | Schriftenverzeichnis.....                                         | 30 |

## 1 Einleitung

Wissenschaftlicher Erkenntnisgewinn und Fortschritt in der Erdsystemforschung hängen in zunehmendem Maße von der Verfügbarkeit einschlägiger Daten ab. Für die Erforschung des globalen Wandels ist besonders die Verknüpfung von interdisziplinären Daten von größter Bedeutung. Wichtige Beobachtungen und neue Einblicke in globale Umweltveränderungen sind nur mit Hilfe von Daten möglich, die sowohl über einen langen Zeitraum aufgezeichnet wurden als auch kohärent und einfach zugänglich sind. Im Gegensatz zu anderen Forschungsfeldern (Astronomie, Hochenergiephysik, Genetik etc.) existiert in der Erdsystemforschung weder auf nationaler noch auf internationaler Ebene eine Dateninfrastruktur, die einen zentralen und einfachen Zugang zu den vielfältigen, jedoch verteilten Datensätzen der Meeresforschung erlauben würde.

Vor diesem Hintergrund sollen mit dem „Marinen Netzwerk für Integrierten Datenzugriff“ (MANIDA 2014), das von der Helmholtz-Gemeinschaft und mehreren deutschen Meeresforschungsinstituten und Universitäten finanziell unterstützt wird, eine vernetzte Infrastruktur für die Erschließung verteilter mariner Forschungsdaten sowie entsprechende auf Nachhaltigkeit ausgelegte Datenmanagementstrategien bereitgestellt werden. Das Netzwerk strebt die Integration, Harmonisierung und Aggregation verschiedenartiger qualitätskontrollierter Daten und Datenprodukte der deutschen Meeresforschung an und hat als solches Modellcharakter.

Eine der Hauptaufgaben von MaNIDA ist die Implementierung und Wartung einer nachhaltigen digitalen Infrastruktur in Form eines zentralen Datenportals (das DATENPORTAL DEUTSCHE MEERESFORSCHUNG (2014)), das die schlüssige Suche,

Visualisierung, Abfrage und Weitergabe wissenschaftlicher Daten und Publikationen ermöglicht. Damit wird erstmals ein großer Bestand an marinen Forschungsdaten aus national betriebenen Forschungs- und Monitoringseinrichtungen über ein zentrales Portal durchsuchbar und zugänglich gemacht.

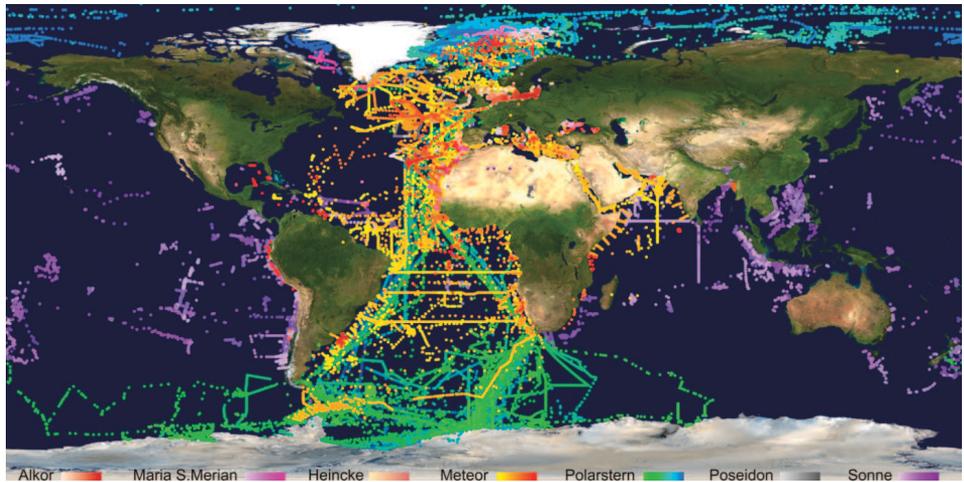


Abbildung 1: Aktueller Status der Probennahmestationen und Fahrten der wichtigsten deutschen Forschungsschiffe (Quelle PANGAEA 2013).

## 2 Das MaNIDA-Konsortium

Das marine Netzwerk für integrierten Datenzugriff wird seit Februar 2012 als Projekt des Impuls- und Vernetzungsfonds betrieben und durch Ressourcen der Helmholtz-Gemeinschaft und der Partnerinstitute finanziell unterstützt. Gegenwärtig gehören dem Konsortium fünf Partnerinstitute und zwei angegliederte Universitäten an, die in der Meeresforschung tätig sind und Dateninfrastrukturen für die Forschung aufbauen. Zu den ursprünglichen Initiatoren des Projekts gehören:

- AWI – Alfred-Wegener-Institut, Helmholtz-Zentrum für Polar- und Meeresforschung, Bremerhaven
- MARUM – Zentrum für Marine Umweltwissenschaften, Universität Bremen
- BSH – Bundesamt für Seeschifffahrt und Hydrographie, Hamburg und Rostock-Warnemünde
- GEOMAR – Helmholtz-Zentrum für Ozeanforschung Kiel
- HZG – Helmholtz-Zentrum Geesthacht, Zentrum für Material- und Küstenforschung GmbH
- CAU – Christian-Albrechts-Universität zu Kiel, Institut für Informatik (angegliedert)
- UHH – Universität Hamburg, Leitstelle Deutsche Forschungsschiffe (angegliedert)

MaNIDA und das „Datenportal Deutsche Meeresforschung“ werden vom Alfred Wegener Institut (AWI) koordiniert. Ein Lenkungsausschuss und mehrere partnerübergreifende Arbeitsgruppen (insgesamt mehr als 20 vorwiegend aus Eigenmitteln finanzierte wissenschaftliche Mitarbeiter) arbeiten gemeinsam an Arbeitsabläufen, Harmonisierung,

Standards und technischen Konzepten von Daten sowie an der praktischen Implementierung des zentralen Datenportals, der Schnittstellen und der Infrastrukturanpassungen der verschiedenen beteiligten Datenanbieter.

Seit dem Ende der Anlaufphase im Juli 2014 steht das MaNIDA-Konsortium auch anderen nationalen Partnerinstituten offen, die sich an der Entwicklung und Einrichtung gemeinsamer Arbeitsabläufe, Standards und digitaler Infrastrukturen für marine Forschungsdaten beteiligen können.

## 2.1 Entwicklung von Datenarbeitsabläufen und Datenpflege

Eine wichtige Voraussetzung für akzeptable zentrale Such- und Zugriffsfunktionen sind gut aufbereitete Inhalte, die zuverlässig und nachhaltig bereitgestellt werden. Neben den erforderlichen technischen Lösungen gilt es daher vor allem, datenrelevante Arbeitsabläufe und Verfahren zur Pflege wissenschaftlicher Daten zu entwickeln und umzusetzen. Die folgenden Daten sollen angeboten werden:

- validierte Expeditions- und Kampagnendaten zu deutschen Forschungsschiffen und -plattformen,
- archivierte, qualitätskontrollierte Primärdaten, echtzeitnahe Daten (Near-Real-Time-Daten) und Datenprodukte,
- Daten mit langfristiger Zitierfähigkeit und Wiederverwendbarkeit,
- wissenschaftliche Publikationen und Berichte.

Zu diesem Zweck werden in den partnerübergreifenden Arbeitsgruppen von MaNIDA Grundsätze für die Umsetzung von Best Practices in Bezug auf Validierung, Archivierung und Weitergabe mariner Forschungsdaten sowie Empfehlungen zu geeigneten Forschungsinfrastrukturen für Entscheidungsträger und Gremien erarbeitet.

Auf lange Sicht werden die wissenschaftlichen Nutzer von optimierten Verfahren bei Datenerfassung, Qualitätssicherung und Benutzer-Support profitieren. Insgesamt werden sich erhebliche Verbesserungen in der Abwicklung des gesamten Lebenszyklus mariner Forschungsdaten erzielen lassen. Ausdrücklich geplant ist in der nächsten Projektphase die Einrichtung eines Zentrums für die Datenkuration (Data Curation Center), die Implementierung von Technik- und Organisationsstrukturen sowie Maßnahmen zur Sicherstellung der finanziellen Unterstützung der Infrastruktur. Die Aufgabenschwerpunkte des gemeinsamen Data Curation Centers für marine Forschungsdaten sind:

- Supportleistungen für Benutzer und Arbeitsabläufe
- Kontaktstelle und Redaktionssystem für standardisierte Vokabularien und Ontologien sowie
- Aufbau einer strukturierten Datenmanagementlösung.

Um die hierfür benötigten zuverlässigen Arbeitsabläufe zu schaffen, sucht MaNIDA die Kooperation von Akteuren aus Wissenschaft und Behörden.



Abbildung 2: Vereinfachtes Schema zum Arbeitsablauf bei kuratierten Forschungsdaten.

Dank der Arbeit von MaNIDA und durch das zentrale Datenportal wird die Erschließung interdisziplinärer Daten und die Planung zukünftiger Expeditionen erheblich verbessert, da sich fehlende Daten einfach identifizieren lassen. Das Portal bietet jedoch zusätzlich die Möglichkeit der Zusammenführung von Daten verschiedenen Ursprungs (Fahrtdaten, nach den Fahrten erhobene Daten, per Satellitennetz erfasste Daten, Modelldaten usw.) in erweiterten Datenprodukten.

Von den beteiligten Partnern wurden übergreifende Arbeitsgruppen mit entsprechenden Kompetenzen und Aufgabenstellungen eingerichtet:

- Datenportal: Entwicklung und Implementierung von Architektur und Schnittstellen gemäß internationaler Standards und Anforderungen für beteiligte Datenprovider
- Expeditions katalog: zentrale Verwaltung der Metadaten zu aktuellen und zurückliegenden Expeditionen mit Links zu Forschungsdaten und Publikationen
- Fahrtdaten (Underway-Daten): Validierung und Harmonisierung mit Hilfe eines verbesserten Datenerfassungs- und Informationsmanagementsystems (DSHIP 2014), das die an Bord aller großen deutschen Forschungsschiffe erhobenen technischen, nautischen und wissenschaftlichen Daten aufzeichnet
- Fahrtberichte (Cruise Summary Reports): halbautomatisierte Generierung und Schnittstellenintegration DSHIP für nationale Datenanbieter und internationale Vernetzung der wichtigsten deutschen Forschungsschiffe und -plattformen
- Vokabularien, Ontologien, Qualitätssicherung: Harmonisierung und Festlegung internationaler Standards
- Gemeinsames Data Curation Center und Ticketing-System: Anforderungen, Konzepte und Einrichtung
- Angemessene Handhabung mariner Forschungsdaten: gemeinsame Datenrichtlinien und Datenmanagementpläne

### **3 Das „Datenportal Deutsche Meeresforschung“**

#### **3.1 Inhalt und Funktionen**

Mit dem „Datenportal Deutsche Meeresforschung“ wird sowohl ein integrativer Rahmen für die einheitliche Auffindbarkeit, Visualisierung und den Abruf als auch für die Veröffentlichung von marinen Forschungsdaten aus unterschiedlichen Quellen (Forschungsschiffe, Observatorien, Glider usw.) geschaffen. Die Inhalte des Portals stammen aus den verschiedenen Daten- und Publikationsrepositorien der in der marinen Forschung in Deutschland tätigen Partnerinstitute. Abb. 3 zeigt die große Bandbreite der im Portal integrierten Informationen: Metadaten zu Forschungsfahrten (Expeditionen), Berichte, Publikationen, archivierte und echtzeitnahe Daten sowie Datenprodukte der gesamten Meeres- und Erdsystemforschungsgemeinschaft Deutschlands und ihrer internationalen Projekte. Vorgehalten werden in erster Linie publizierte Datensätze mit den erforderlichen Metadaten und globalen persistenten Identifikatoren (z. B. DOI 2014) für eine korrekte Zitierfähigkeit. Damit soll der Grundsatz des freien Datenaustauschs und der Datenwiederverwendung bei gleichzeitigem Hinweis auf den ursprünglichen Datenproduzenten unterstützt werden.



Abbildung 3: Inhalte des „Datenportals Deutsche Meeresforschung“.

Bislang wurden die folgenden Suchfunktionen im Datenportal implementiert:

- Allgemeine Suche nach Daten oder Publikationen anhand von Schlüsselwörtern, die in den Metadaten der aus den jeweiligen Repositorien gewonnenen Datensätze beschrieben sind. Durch Auswahl von Ortslexika (Gazetteers) oder grafischer Abfrage kann die Suche zeitlich und räumlich eingeschränkt werden.
- Facettensuche und Suche nach gegebenen Kategorien wie z. B. Datenanbieter (Repositorien), Regionen (Gazetteers), Autoren, Plattformen, Expeditionen/Kampagnen, Projekten, Geräten und Parametern.
- Das Portal bietet neben dem Zugriff auf Datensätze, Datenprodukte und Near-Real-Time-Daten auch Recherchemöglichkeiten über verlinkte wissenschaftliche Veröffentlichungen, Berichte und Dokumentationen in den Publikationsbeständen der jeweiligen Institute, die mit den verfügbaren Datensätzen verbunden sind.

### 3.2 Mehrwertdienste

Im Rahmen von MaNIDA werden prototypische Fallstudien zur Optimierung und Erweiterung des Datenabrufs erarbeitet und durch Einbindung der Rückmeldungen von Benutzern fortlaufend weiterentwickelt. Derzeit werden drei Mehrwertdienste angeboten:

- Ein Modul für den direkten Zugriff auf die Datenquellen PANGAEA (2014, AWI/MARUM), COSYNA (2014, HZG) und DOD (2014, BSH). Verschiedene Messwerte können durch Definition einer zeitlichen und räumlichen Ausdehnung für harmonisierte Parameter wie Salzgehalt oder Temperatur abgefragt werden. Die abgefragten und integrierten Daten stehen dann zum Download bereit. Diese direkte Zugriffsmöglichkeit wird durch Harmonisierung auf weitere Parameter ausgedehnt und den Datenanbietern über internationale Vokabularien zugänglich gemacht.
- Zum ersten Mal steht ein Pool validierter Bathymetriedaten der Partnerinstitute zur Verfügung, die über eine gut definierte, erweiterte und einheitliche Metadatenbeschreibung zentral abrufbar sind.

- Messdaten aus der Deutschen Bucht sind im Datenportal durch einen Vergleich der Strömungsvektorfelder (Strömungsgeschwindigkeit an der Wasseroberfläche) nahezu in Echtzeit visualisiert und abrufbar.

### 3.3 Datenanbieter

Die im „Datenportal Deutsche Meeresforschung“ zur Verfügung gestellten Datensätze werden von unterschiedlichen lokalen Quellen abgerufen, die an verschiedenen Instituten vorgehalten und aktualisiert werden. In Folge dessen werden die Datensätze zwar in anbieterspezifischen Formaten bereitgestellt, sind jedoch immer direkt verfügbar und auf dem neuesten Stand.

Die Harmonisierung u. a. der Dateninhalte, Parameter und Vokabularien ist eine langfristige Aufgabe bei der technischen Zusammenführung der Datenbestände der Partnerinstitute. Gemeinsam mit unserer Forschungsgemeinschaft und unseren Datenanbietern arbeiten wir kontinuierlich an der Standardisierung, Klassifizierung und Verlinkung der Daten mit internationalen Initiativen wie ICSU World Data System (ICSU WDS 2014), Global Earth Observation System of Systems (GEOSS 2014), Global Biodiversity Information Facility (GBIF 2014), World Register of Marine Species (WORMS 2014) sowie EU-Initiativen, z. B. European Marine Observation and Data Network (EMODNET 2014), SEADATANET (2014) mit BODC NERC (2014), CF-Standardisierungen (CFC 2014) sowie EUROFLEETS (2014).

Die folgenden Datenanbieter sind am „Datenportal Deutsche Meeresforschung“ beteiligt:

- PANGAEA Data Publisher for Earth & Environmental Science (2014): Dieses Informationssystem ermöglicht die Langzeitarchivierung und Publikation von Daten aus den Erd- und Umweltwissenschaften und wird seit 1993 von AWI und MARUM gemeinschaftlich als Bibliothek mit offenem Zugriff auf Forschungsdaten betrieben. PANGAEA vom International Council for Science (ICSU) als World Data Center akkreditiert worden und von der World Meteorological Organization (WMO) anerkannt. Jeder Datensatz kann über einen Digital Object Identifier (DOI) dauerhaft identifiziert, verbreitet, publiziert und zitiert werden. Die Daten werden in Form von eigenständigen und zitierfähigen Datensammlungen oder als Ergänzungen zu Publikationen archiviert. Als Datenbibliothek verlinkt PANGAEA Primärdaten im Zusammenhang mit Beiträgen in Wissenschaftsmagazinen der Erd- und Umweltwissenschaften aus dem Portfolio der Verlage ELSEVIER (2014) und COPERNICUS PUBLICATIONS (2014), die über SCIENCE DIRECT (2014) kostenlos zur Verfügung stehen. Gegenwärtig bietet PANGAEA mehr als 350.000 Datensätze und 8 Milliarden Messungen.
- COSYNA (2014): Das „Coastal Observing System for Northern and Arctic Seas“ ist ein integriertes, operationelles Beobachtungssystem, das Beobachtungsdaten mit numerischer Modellierung kombiniert. Zu den Messgrößen gehören physikalische, geochemische und biologische Parameter sowie Sedimentschlüsselgrößen mit hoher zeitlicher Auflösung in der Wassersäule und den oberen und unteren Grenzschichten. Die Datenprodukte umfassen in Echtzeit oder nahezu Echtzeit erhobene Datensätze und Webdienste, die aus Vor-Ort-Beobachtung und Fernerkundung abgeleitet wurden, aber auch Vorhersagen auf Basis von

Modellergebnissen. Aufgabe des COSYNA-Datenmanagements ist die Organisation der Datenströme zwischen den Beobachtungs- und zentralen Speichersystemen am Helmholtz-Zentrum Geesthacht und an den Partnerstandorten sowie die Organisation der Datendokumentation und der Benutzeroberflächen für die Datenabfrage und -darstellung. Das Portal enthält ca. 15.000 Datensätze von COSYNA.

- DOD (2014): Das Deutsche Ozeanographische Datenzentrum wurde im Jahr 1967 als Gemeinschaftsprojekt des Deutschen Hydrographischen Instituts (DHI) und der Deutschen Forschungsgemeinschaft (DFG) eingerichtet. Die nationale ozeanografische Datenbank umfasst derzeit Daten und Informationen von deutschen Forschungsfahrten und nahezu 300.000 Stationen, die insgesamt mehr als 45 Millionen Datenwerte zu ca. 900 Parametern liefern. Dabei handelt es sich hauptsächlich um ozeanografische Parameter wie Temperatur und Salzgehalt, chemische Messwerte zu Nährstoffen, organische, anorganische und radioaktive Komponenten in Meerwasser und Sedimenten sowie Schadstoffe in Organismen (Fische und Muscheln).
- In Zusammenhang mit den verfügbaren Datensätzen der Datenanbieter bietet das Datenportal Zugriff auf wissenschaftliche Veröffentlichungen, Berichte und Dokumentationen. Die teilnehmenden Institute stellen die folgenden internationalen Publikationsrepositorien zur Verfügung: EPIC (2014) von AWI, OCEANREP (2014) von GEOMAR und Publikationen von HZG (2014) und MARUM (2014).

### 3.4 Architektur

In Abb. 4 ist die dem „Datenportal Deutsche Meeresforschung“ zugrunde liegende Architektur dargestellt. Die Integration der verteilt verfügbaren Inhalte erfolgt über die automatische Erfassung (Harvesting) und Indizierung der von unseren Datenanbietern bereitgestellten Metadaten. Dieser Ansatz ermöglicht eine sehr leistungsfähige Suche nach wissenschaftlichen Inhalten auf Basis eines APACHE LUCENE (2014) ELASTICSEARCH (2014)-Clusters. Da die Systeme der Datenanbieter verschiedene Metadatenformate über unterschiedliche Schnittstellen anbieten, wurden die folgenden Metadatenfelder als Mindestvoraussetzung definiert:

- Persistent Identifier
- Titel
- Datum der Veröffentlichung
- Autoren, Untersuchungsleiter
- Anfangs- und Enddatum der Datenerhebung
- Forschungsplattform oder -schiff
- Expedition
- Messgrößen

Die Metadaten enthalten darüber hinaus Links zu Downloaddiensten oder dynamischen Diensten wie Sensor Observation Services (SOS) oder Web Map Services (WMS). Um den Benutzern harmonisierte Inhalte (Metadaten) präsentieren zu können, werden die unterschiedlichen Nomenklaturen der Parameter mittels eines Feature Catalogue den

Standardnamen angeglichen und die ursprünglichen Metadaten auf Basis eines standardisierten Gazetteer mit den Namen der Meeresregionen versehen.

Neben den von den Datenanbietern bereitgestellten Metadaten betreibt das Portal eine Datenbank, in der Metadaten zu Expeditionen (z. B. IDs, Beginn, Ende, Häfen, Fahrtleiter) aus unterschiedlichen Quellen zusammengeführt werden. Dieser „Expedition Catalogue“ dient als Masterkatalog für das Datenportal. Die Inhalte der beiden Kataloge werden im bereits genannten „Data Curation Center“ bearbeitet.

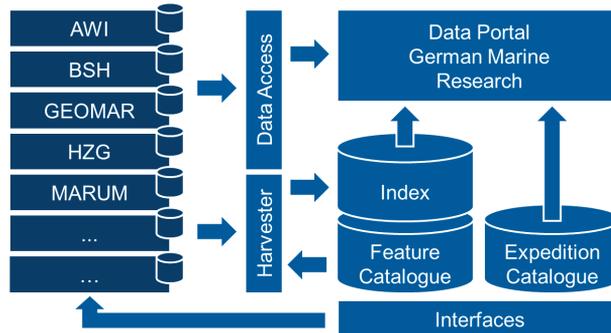


Abbildung 4: Architektur des „Datenportals Deutsche Meeresforschung“.

Das Portal bietet zusätzlich Schnittstellen in Form von Webdiensten für den Feature Catalogue und den Expedition Catalogue. Datenanbieter können diese berechtigten und aufbereiteten Informationen und Metadaten für ihre eigenen Kataloge nutzen.

### 3.5 Schnittstellen und Standards

Der Harvesting-Ansatz unterstützt verschiedene Standards für die Metadatenerfassung. Unter anderem werden die folgenden drei Protokolle unterstützt:

- OAI-PMH (2014): Open Archives Initiative Protocol for Metadata Harvesting,
- OGC CSW (2014): Open Geospatial Consortium Catalogue Service for Web und
- OGC WFS (2014): Open Geospatial Consortium Web Feature Service.

Metadaten werden üblicherweise in Form von XML-Serialisierungen bereitgestellt. In diesem Fall wird das mit INSPIRE (2014) konforme ISO 19115/19139-Format empfohlen. Der Datenanbieter definiert das Format, in dem die in den Metadatensätzen beschriebenen Informationen weitergegeben werden. Übliche Datenformate sind Formate TSV (durch Tabulatoren getrennte Werte) oder CSV (durch Kommata getrennte Werte) – möglicherweise in einem ZIP-Archiv komprimiert – sowie das PDF-Format (portables Dokumentformat) für Publikationen und Berichte. Das Angebot an Webdiensten reicht von Dateien auf Basis von OPENDAP (2014) über Sensor Observation Service (SOS 2014) bis hin zu OGC Web Feature Services (WFS) und Web Map Services (WMS 2014).

Beim direkten Datenzugriff werden gleichzeitige Datenanfragen an unterschiedliche Datenanbieter unterstützt. Für die Datenanfrage und -übermittlung bietet das Portal sowohl den direkten SQL-Lesezugriff (z. B. Data Warehouse von PANGAEA (2014) als auch WFS-Anforderungen (Web Feature Service), z. B. aus COSYNA (2014) und DOD (2014) an.

### 3.6 Datenzugriffsbestimmungen und gute wissenschaftliche Praxis

Das Datenportal basiert auf offenen Technologien. Der Zugriff steht Wissenschaftlern, Förderstellen und der Öffentlichkeit frei zur Verfügung. Bislang ist keine Registrierung erforderlich. Um Rückmeldung von Benutzern wird gebeten, damit Funktionalität und Inhalt des gemeinsamen Datenportals fortlaufend verbessert werden können.

Das Zitieren von verwendeten Datensätzen hat sich in der Wissenschaft bewährt. Daher sollten die heruntergeladenen Datensätze immer mit einem Zitat referenziert bzw. mit dem entsprechenden vom Datenanbieter angebotenen Zitat verwendet werden. Um eine lückenlose Referenzierung sicherzustellen, ist zusammen mit dem Datenzitat immer die Zugriffsnummer bzw. die DOI des Datensatzes anzugeben.

## 4 Dankeswort

Zum Kernteam des MaNIDA-Konsortiums gehören mehr als zwanzig wissenschaftliche Mitarbeiter aus allen Partnerinstituten und darüber hinausgehenden Kooperationen, die sich zumeist in Eigeninitiative und mit eigenen Ressourcen, im Dienst der Wissenschaft, an der Gestaltung und Einrichtung gemeinsamer Dateninfrastrukturen und Arbeitsabläufe beteiligen. Diesem unermüdlichen Einsatz gebührt unser Dank und unsere Anerkennung, zumal Datenmanagement und -veröffentlichung in der Regel im wissenschaftlichen Bewertungssystem nicht ausreichend gewürdigt oder ausgezeichnet werden. Dem Einsatz aller an MaNIDA Mitwirkenden ist es zu verdanken, dass das „Datenportal Deutsche Meeresforschung“ verwirklicht werden konnte und die verteilten Datenanbieter der wesentlichen Repositorien für die Erdsystem- und Meeresforschung in Deutschland vernetzt. Das MaNIDA-Projekt wird in erster Linie durch den Impuls- und Vernetzungsfonds der Helmholtz-Gemeinschaft (SO-071) und zu einem nicht unbedeutenden Teil durch eigene Ressourcen aller Partner finanziert und unterstützt.

## 5 Schriftenverzeichnis

APACHE LUCENE: Stand 14.06.2014.

BODC NERC - BRITISH OCEANOGRAPHIC DATA CENTRE – NATURAL ENVIRONMENT RESEARCH COUNCIL.

Stand 14.06.2014: [http://www.bodc.ac.uk/products/web\\_services/vocab/](http://www.bodc.ac.uk/products/web_services/vocab/)

CFC - CLIMATE AND FORECAST CONVENTIONS.

Stand 14.06.2014: <http://cfconventions.org/>

COPERNICUS PUBLICATIONS.

Stand 14.06.2014: <http://publications.copernicus.org/>

COSYNA - COSTAL OBSERVING SYSTEM FOR NORTHERN AND ARCTIC SEAS.

Stand 14.06.2014: <http://www.cosyna.de/>

DATA PORTAL GERMAN MARINE RESEARCH.

Stand 14.06.2014: <http://manida.awi.de>

DOD - GERMAN OCEANOGRAPHIC DATA CENTRE.

Stand 14.06.2014: [http://www.bsh.de/en/Marine\\_data/Observations/DOD\\_Data\\_Centre/](http://www.bsh.de/en/Marine_data/Observations/DOD_Data_Centre/)

DOI - DIGITAL OBJECT IDENTIFIER.

Stand 14.06.2014: <http://www.doi.org/>

DSHIP Data Acquisition System.

Stand 14.06.2014: <http://www.werum.de/plattformen/dship.jsp>

ELASTICSEARCH.

Stand 14.06.2014: <http://www.elasticsearch.org/>

ELSEVIER.

Stand 14.06.2014: <http://www.elsevier.com/>

EMODNET - EUROPEAN MARINE OBSERVATION AND DATA NETWORK.

Stand 14.06.2014: <http://www.emodnet.eu/>

EPIC - ELECTRONIC PUBLICATION INFORMATION CENTER.

Stand 14.06.2014: <http://epic.awi.de/>

EUROFLEETS.

Stand 14.06.2014: <http://www.eurofleets.eu/>

GBIF - GLOBAL BIODIVERSITY INFORMATION FACILITY.

Stand 14.06.2014: <http://www.gbif.org/>

GEOSS GLOBAL EARTH OBSERVATION SYSTEM.

Stand 14.06.2014: <https://www.earthobservations.org/geoss.shtml>

HZG – Publication database.

Stand 14.06.2014: <http://141.4.217.215/fmi/xsl/publikat/Search.xml>

ICSU WDS - INTERNATIONAL COUNCIL FOR SCIENCE WORLD DATA SYSTEM.

Stand 14.06.2014: <https://www.icsu-wds.org/>

INSPIRE - INFRASTRUCTURE FOR SPATIAL INFORMATION IN THE EUROPEAN COMMUNITY.

Stand 14.06.2014: <http://inspire.ec.europa.eu/index.cfm/pageid/101>

MANIDA - MARINE NETWORK FOR INTEGRATED DATA ACCESS.

Stand 14.06.2014: <http://manida.org>

MARUM Publications.

Stand 14.06.2014: <http://publications.marum.de/>

OAI-PMH - OPEN ARCHIVES INITIATIVE PROTOCOL FOR METADATA HARVESTING.

Stand 14.06.2014: <http://www.openarchives.org/pmh/>

OCEANREP.

Stand 14.06.2014: <http://oceanrep.geomar.de/>

OGC CSW - OPEN GEOSPATIAL CONSORTIUM CATALOGUE SERVICE FOR WEB.

Stand 14.06.2014: <http://www.opengeospatial.org/standards/cat>

OGC WFS - OPEN GEOSPATIAL CONSORTIUM WEB FEATURE SERVICE.

Stand 14.06.2014: <http://www.opengeospatial.org/standards/wfs>

OPENDAP.

Stand 14.06.2014: <http://www.opendap.org/>

PANGAEA – Data Publisher for Earth & Environmental Science.

Stand 14.06.2014: <http://www.pangaea.de/>

SCIENCE DIRECT.

Stand 14.06.2014: <http://www.sciencedirect.com/>

SEADATANET.

Stand 14.06.2014: <http://www.seadatanet.org/>

SOS - OPEN GEOSPATIAL CONSORTIUM SENSOR OBSERVATION SERVICE.

Stand 14.06.2014: <http://www.opengeospatial.org/standards/sos>

WMS - OPEN GEOSPATIAL CONSORTIUM WEB MAP SERVICE.

Stand 14.06.2014: <http://www.opengeospatial.org/standards/wms>

WORMS - WORLD REGISTER OF MARINE SPECIES.

Stand 14.06.2014: <http://www.marinespecies.org/>