

Pressemitteilung

24/2022

Meeresforschende machen Datenschatz aus der Tiefsee zugänglicher Einheitliche Datenstandards sollen die Nutzung von Fotos und Videos erleichtern

19.07.2022/Kiel. Weltweit lagern auf den Servern von Meeresforschungsinstituten tausende von Aufnahmen aus der Tiefsee. Bislang sind diese Videos und Fotos über das Internet kaum auffindbar, weil sie nicht einheitlich katalogisiert und beschrieben werden. Ein Team der Helmholtz-Gemeinschaft deutscher Forschungszentren hat unter der Leitung des GEOMAR einen einheitlichen Datenstandard entwickelt, der die globale Nutzung von Bilddaten erleichtern soll. Das neue Metadaten-Format wird jetzt im Fachmagazin *Nature Scientific Data* vorgestellt.

Das Leben in der Tiefsee wird heute mehr und mehr mit hochauflösenden Kameras dokumentiert, die an ferngesteuerten Tauchrobotern oder autonomen Unterwasserfahrzeugen angebracht sind. Fachleute werten diese Aufnahmen wissenschaftlich aus, um Informationen über Lebewesen im Freiwasser und am Meeresboden sowie geologische Strukturen zu erhalten. Weltweit lagern ungeheure Mengen solcher Foto- und Videodaten auf den Servern von Meeresforschungsinstituten – jedoch sehr unterschiedlich katalogisiert. Um diesen Datenschatz international nutzbar zu machen, müssen wichtige Suchbegriffe und Informationen etwa über die Position des Tauchroboters während der Aufnahme, die verwendete Kameratechnik sowie die Namen der Expedition und der beteiligten Wissenschaftler:innen in einem universell lesbaren Format in der Bilddatei hinterlegt sein.

Um dies zu gewährleisten, hat eine Arbeitsgruppe der Helmholtz-Gemeinschaft deutscher Forschungszentren unter Mitwirkung des GEOMAR Helmholtz-Zentrum für Ozeanforschung in Kiel, des Alfred-Wegener-Instituts, Helmholtz-Zentrum für Polar- und Meeresforschung (AWI) und des Helmholtz-Zentrums Hereon ein international einheitliches Metadatenformat für Unterwasseraufnahmen entwickelt. Beteiligt waren dabei auch Fachleute des DataHub, einer Dateninitiative des Helmholtz-Forschungsbereichs Erde und Umwelt, und der Helmholtz Metadata Collaboration (HMC). Ihre Vorschläge präsentieren sie in der aktuellen Ausgabe des Fachmagazins *Nature Scientific Data*.

Dr. Timm Schoening, Hauptautor des Artikels und Datenwissenschaftler am GEOMAR, sagt dazu: „Es gibt weltweit bereits seit einigen Jahren Bestrebungen, Daten allgemein zugänglich zu machen. Wir schaffen mit unserem einheitlichen Metadatenstandard die Voraussetzungen dafür, dass auch wissenschaftliche Fotos und Videos aus der Tiefsee entsprechend dieser Initiative international zugänglich werden. Und wir machen Software verfügbar, mit der dieser Standard nutzbar wird.“

Das neue Format setzt die international anerkannten „FAIR“-Prinzipien für ein nachhaltiges Forschungsdaten-Management um. Die Abkürzung steht für „findable, accessible, interoperable and reusable“ – auffindbar, zugänglich, mit verschiedenen Softwaresystemen nutzbar und wiederverwendbar. Dateien, die anhand ihrer Metadaten für andere Forschende auffindbar sind, werden FAIR Digital Objects (FDO) genannt. Das jetzt vorgestellte FAIRe Datenformat für Unterwasseraufnahmen wurde entsprechend „image FAIR Digital Objects“ (iFDO) getauft. Es ist gewissermaßen eine informatische Karteikarte, die übersichtlich all jene Aspekte zusammenfasst, die für eine Aufnahme wichtig sind. Sie beinhaltet nicht nur beschreibende Informationen über die Bilddaten selbst, sondern ebenfalls unveränderliche Weblinks zu den Bilddaten.

Inzwischen hat das Helmholtz-Team das iFDO-Konzept international mehrfach vorgestellt. „Unser Ansatz stößt auf großes Interesse“, berichtet Dr. Schoening. „Insofern sind wir zuversichtlich, dass er als Vorlage für einen neuen internationalen Standard für Unterwasseraufnahmen dienen wird.“

Ergänzend zum iFDO-Metadatenformat hat die Gruppe mehrere Software-Tools entwickelt, die es möglich machen, das iFDO-Format für verschiedene biologische oder geologische Interpretationen zu übernehmen. Eine weitere Idee ist, Kamerasysteme künftig so auszustatten, dass sie bereits während einer Aufnahme automatisch Metadaten im iFDO-Format generieren. Wie gut das funktioniert, testeten Forschende des GEOMAR während der Expedition M182 mit dem Forschungsschiff METEOR im Atlantischen Ozean. Die Kameras an Bord der autonomen Unterwasserfahrzeuge ANTON, LUISE und ABYSS sowie an geschleppten Instrumenten und stationären Verankerungen speicherten schon während ihrer Einsätze die iFDO-Metadaten ab. „Diese Erfahrung war sehr positiv und hat noch zu einigen Ergänzungen an der Dokumentation und den Software-Tools geführt – die iFDOs selbst haben diese Expedition mit Bravour bestanden“, sagt Timm Schoening.

Eine besondere Herausforderung bei Bild- und Videoaufnahmen besteht darin, dass ein Computer sie nicht ohne weiteres auswerten kann. Bei Temperatur- oder Tiefenmessungen ist das anders: Die Zahlenwerte lassen sich einfach speichern und in einem Diagramm darstellen. Ein Video hingegen ist für einen Computer nur ein Datenstrom aus Pixeln. Daher muss zunächst im Bildmaterial jedes Objekt markiert und definiert werden – ein längliches Objekt etwa als Seegurke. Fachleute sprechen dabei von Annotation und nutzen dafür spezialisierte Software wie zum Beispiel BIIGLE (Bio-Image Indexing and Graphical Labelling Environment), die an der Universität Bielefeld entwickelt wurde und auch am GEOMAR eingesetzt wird.

„Wir haben den Annotationsschritt bei der Entwicklung der iFDOs direkt mit in das Datenformat eingebaut. Und wir haben die Funktionalität direkt in die Software BIIGLE eingebaut, so dass dieses weit verbreitete Tool jetzt das iFDO Format bereits unterstützt“, sagt Dr. Schoening. „Das sind zwei große Pluspunkte, die bei unseren Präsentationen bereits viel Aufmerksamkeit bekommen haben: iFDOs können nicht nur als Standard für die Metadaten eingesetzt werden, sondern auch als Standard für die Annotationen und es gibt benutzbare Software, die das Format unterstützt.“

Hier greifen direkt die Vorteile der FAIR Prinzipien, da eine effektive Wiederverwendung der Daten ermöglicht wird: Die Bilddaten und Annotationen können so ebenfalls zum Training von Algorithmen des Maschinellen Lernens (ML) verwendet werden. Da sie in den iFDOs bereits in einem FAIRen Datenformat vorliegen, ist die Entwicklung von ML-Algorithmen deutlich weniger komplex.

Publikation:

Schoening, T., Durden, J.M., Faber, C. et al. (2022): Making marine image data FAIR, *Nature Scientific Data*, doi: <https://doi.org/10.1038/s41597-022-01491-3>

Links:

<https://www.awi.de> Alfred-Wegener-Institut, Helmholtz-Zentrum für Polar- und Meeresforschung (AWI)
<https://www.hereon.de> Helmholtz-Zentrum Hereon
<https://www.geomar.de> GEOMAR Helmholtz-Zentrum für Ozeanforschung Kiel
<https://www.marine-imaging.com/fair/ifdos/iFDO-overview> FAIR marine images
<https://datahub.erde-und-umwelt.de> DataHub des Helmholtz-Forschungsbereichs Erde und Umwelt
<https://helmholtz-metadaten.de> Helmholtz Metadata Collaboration (HMC)
<https://www.go-fair.org/fair-principles> Die FAIR-Prinzipien (englisch)
<https://biigle.de> Bio-Image Indexing and Graphical Labelling Environment (BIIGLE) (englisch)

Bildmaterial:

Unter www.geomar.de/n8544 steht Bildmaterial zum Download bereit.

Kontakt:

Ann Kristin Montano (GEOMAR, Kommunikation & Medien), Tel.: 0431 600-2811, media(at)geomar.de