



Handbuch

für



Geowissenschaftlicher Datenverlag für Österreich

beinhaltet

Datenpublikationsrichtlinien

Terms & Conditions

Anleitung

Version: 2.1

Geosphere Austria
Hohe Warte 38
1190 Wien

Zweigstelle: Neulinggasse 38, 1030 Wien

Kontakte

GeoSphere Austria: office@geosphere.at
Tethys RDR: repository@geosphere.at

Mittwoch, 16. August 2023

Beteiligte Autoren

Viktoria Haider

Thomas Brus
Christian Cermak
Werner Gesselbauer
Thomas Hofmann
Christoph Janda
Arno Kaimbacher
Christian Linsberger
Johannes Reischer

i. Inhaltsverzeichnis

1. DATENPUBLIKATIONSRICHTLINIEN (DATA POLICY).....	7
1.1 MISSION STATEMENT	7
1.2 LEITLINIE VON TETHYS RDR	7
1.2.1 OPEN SCIENCE PRINCIPLES / DIE OFFENE WISSENSCHAFT	8
1.2.2 FAIR-PRINZIPIEN	9
1.2.3 METADATENBESCHREIBUNG.....	10
1.2.3.1 Dublin Core	11
1.2.3.2 DataCite	11
1.2.3.3 ISO 19139	11
1.2.3.4 Deskriptives Metadatenset: „Dataset Type“ (Datensatzart)	11
1.3 WORKFLOW	14
1.4 DIE RICHTLINIEN FÜR DAS PUBLIZIEREN VON FORSCHUNGSDATEN	15
1.4.1 WORIN LIEGT DER MEHRWERT, DIE FORSCHUNGSDATEN MIT HILFE DES TETHYS RDRs ZU PUBLIZIEREN? ...	16
1.4.2 WELCHE DATEN WERDEN IM FORSCHUNGSDATENREPOSITORY AUFGENOMMEN?.....	17
1.4.3 WELCHE INFORMATIONEN MÜSSEN IN DER DATENPUBLIKATION VORHANDEN SEIN?	18
1.4.4 KÖNNEN BEREITS IM TETHYS RDR PUBLIZIERTE FORSCHUNGSDATEN WIEDER GELÖSCHT WERDEN?	19
1.4.5 KÖNNEN NACHTRÄGLICH ENTDECKTE FEHLER IN DER DATENPUBLIKATION KORRIGIERT WERDEN?.....	19
1.4.6 KÖNNEN EXTERN ERSTELLTE FORSCHUNGSDATEN PUBLIZIERT WERDEN?.....	19
1.4.7 KANN ICH DIE FORSCHUNGSDATEN PUBLIZIEREN, WENN MEHRERE PERSONEN DIESER ERSTELLT HABEN?.....	19
1.4.8 WELCHE LIZENzen STEHEN FÜR DIE DATENPUBLIKATION ZUR VERFÜGUNG?	20
1.4.9 WAS MUSS ICH MACHEN, WENN DIE DATEN AUF FREMDDATEN BASIEREN?	21
1.4.10 IN WELCHEN FORMATEN DARF ICH DIE DATEN PUBLIZIEREN?	21
1.4.11 IN WELCHER SPRACHE KANN ICH DIE DATEN PUBLIZIEREN?	23
1.4.12 KANN ICH MEINE DATENPUBLIKATION EINEM PROJEKT ZUORDNEN?	24
1.4.13 KANN ICH EIN EMBARGO SETZEN UND WAS MUSS ICH DAZU ALLES BERÜCKSICHTIGEN?	24
1.4.14 MUSS ICH MEINE DATENPUBLIKATION MIT KEYWORDS PUBLIZIEREN?	24
1.4.15 GIBT ES EIN UPLOAD-LIMIT?	25
1.4.16 KANN ICH DIE DATENPUBLIKATION AUF BEREITS EXISTIERENDE PUBLIKATIONEN VERWEISEN?	25
1.4.17 WIRD DIE DATENPUBLIKATION ÜBERPRÜFT (REVIEWED)?.....	27
1.4.18 WIE KÖNNEN DIE DATEN VOR DEM FINALEN UPLOAD AN DEN EDITOR ÜBERGEBEN WERDEN?	27
1.4.19 WERDEN DIE ENTWURFSPUBLIKATIONEN AUCH ABGESPEICHERT?	27
1.4.20 WIE HANDHABEN DIE EDITOREN DEN REVIEWPROZESS?.....	27
1.4.21 ICH MÖCHTE VERSCHIEDENE DATENSATZTYPEN GEMEINSAM PUBLIZIEREN. IST DAS MÖGLICH?	28
1.4.22 WIE KANN ICH DIE DATENPUBLIKATION ZITIEREN BZW. DARAUF VERWEISEN?	28
1.5 CHECKLISTE.....	29
2. TERMS AND CONDITIONS.....	30
2.1 WEBAUFRITT	30
2.1.1 VERWEISE/LINKS.....	30
2.1.2 URHEBER- UND KENNZEICHENRECHT	30
2.1.3 DATENSCHUTZ-BESTIMMUNGEN	31
2.1.4 ERFASSUNG ALLGEMEINER DATEN UND INFORMATIONEN.....	32

2.1.5	ANMELDUNG/REGISTRIERUNG AUF UNSERER WEBSITE.....	32
2.1.6	KONTAKTMÖGLICHKEIT ÜBER DIE WEBSITE	33
2.1.7	ÜBERMITTLUNG WISSENSCHAFTLICHER DATEN ÜBER DIE WEBSITE.....	33
2.1.8	NUTZUNG VON CONTENT DELIVERY NETWORKS (CDN).....	33
2.1.9	ROUTINEMÄRIGE LÖSCHUNG UND SPERRUNG PERSONENBEZOGENER DATEN	34
2.1.10	OPENSTREETMAP	34
2.1.11	RECHTSWIRKSAMKEIT DIESES HAFTUNGSAUSSCHLUSSES	34
3.	TECHNISCHE ANLEITUNG ZUR DATENPUBLIKATION	35
3.1	ZEICHENERKLÄRUNG	35
3.2	EINGABEHILFE	35
3.2.1	TETHYS START.....	35
3.2.2	LOGIN.....	36
3.2.3	PUBLISH NEW DATASET (DATENPUBLIKATION VERÖFFENTLICHEN)	36
3.2.3.1	Dataset Language (Sprache der Datenpublikation)	37
3.2.3.2	Mandatory Elements (Verpflichtende Angaben)	37
3.2.3.3	Recommended Elements (Empfohlene Angaben)	41
3.2.3.4	Other Elements (andere Angaben)	43
3.2.3.5	File Upload (Datei hochladen)	43
3.2.3.6	Uploaded 1 file successfully (Hochladen erfolgreich)	44
3.2.4	RELEASE YOUR DATASET FOR EDITOR (FREIGABE DER DATENPUBLIKATION FÜR DEN EDITOR)	45
3.2.5	EDIT DATASET (DATENPUBLIKATION BEARBEITEN)	45
3.2.6	RELEASE SAVED DATASETS (GESPEICHERTE DATENPUBLIKATIONEN FREIGEBEN)	45
4.	DEFINITIONEN	47

ii. Abbildungsverzeichnis

Abbildung 1: Die sechs Prinzipien von Open Science im Überblick.	9
Abbildung 2: FAIR Daten Prinzipien im Überblick, adaptiert von FORCE11	10
Abbildung 3: Deskriptive Auswahlliste „Dataset Type“.	12
Abbildung 4: Workflow-Diagramm.	15
Abbildung 5: Generell mögliche CC Lizenzvarianten, geordnet nach Grad der Offenheit.	21
Abbildung 6: Beispiel eines Zitervorschlages (https://doi.org/10.24341/tethys.193)	28
Abbildung 7: Startseite des Tethys RDR.	35
Abbildung 8: Einstiegsfenster in den Publikationsbereich.	36
Abbildung 9: Erscheinungsbild nach dem Login	36
Abbildung 10: Auswahl der Hauptsprache	37
Abbildung 11: Verpflichtende Angaben zur Datenpublikation.	38
Abbildung 12: Weitere verpflichtende Angaben zu der Datenpublikation.	39
Abbildung 13: Metadaten zu den Mitwirkenden der Datenpublikation.	39
Abbildung 15: Name des Betreibers und die Checkbox für terms and conditions.	40
Abbildung 16: Zweite Publikationsstufe zu den empfohlenen Angaben.	41
Abbildung 17: Ausdehnung der Datenpublikation.	42
Abbildung 18: Möglichkeit zur Erfassung von Referenzen und die Angabe von Schlüsselworten.	42
Abbildung 19: Andere Angaben – Nutzungsbestimmungen.	43
Abbildung 20: Hochladen der Datenpublikation 1	44
Abbildung 21: Hochladen der Datenpublikation 2	44
Abbildung 22: Edit, Release, Delete.	44
Abbildung 23: Hier besteht die Möglichkeit, einen Reviewer auszuwählen.	45
Abbildung 24: Übersicht über die eigene Datenpublikationen während des Publikationsprozesses.	45
Abbildung 25: Release saved datasets.	46

iii. Tabellenverzeichnis

Tabelle 1: Unterstützte Creative Commons Lizenzen.	20
Tabelle 2: Auflistung von unterstützten und geplant unterstützten Formaten.	22
Tabelle 3: Ausgewählte Beispiele für nicht unterstützte Formate.	23
Tabelle 4: Definition der zur Verfügung gestellten Identifier Types.	25
Tabelle 5: Definitionen und Beispiele von möglichen Beziehungstypen.	26
Tabelle 6: Bedeutung der Farben bei „eigene Datenpublikationen“....	46
Tabelle 7: Kurzdefinitionen zu spezifischen Begriffen.	47

1. Datenpublikationsrichtlinien (Data Policy)

1.1 Mission statement

Geosphere Austria hat den gesetzlichen Auftrag, eine zentrale Dateninfrastruktur als Service für Wissenschaft, Wirtschaft, Verwaltung und Gesellschaft mit automatisiertem Zugang aufzubauen und zu betreiben (GSAG §4(3)¹). Eine Säule dieser Infrastruktur ist Tethys RDR, wo verschiedene Arten von georeferenzierten geowissenschaftlichen Forschungsdaten (z.B. Messungen, GIS-Daten, Experimente, Monitoring, 3D-Modelle) frei zugänglich veröffentlicht werden, die innerhalb der Organisation und in Zusammenarbeit mit anderen Institutionen erzeugt werden.

Die Aufgabe von Tethys RDR ist es, Forschungsdaten gemäß den FAIR- und Open Data-Prinzipien so zu veröffentlichen und zu archivieren, dass diese auffindbar, zugänglich, interoperabel und wiederverwendbar sind. Wir legen großen Wert auf die wissenschaftliche Nachvollziehbarkeit sowie auf eine nachhaltige Veröffentlichung und Langzeitarchivierung. Daher bieten wir auch redaktionelle Unterstützung für den Publikationsprozess an. Die Datenveröffentlichungen sind mit einem Digital Object Identifier (DOI) versehen und können eindeutig und einfach zitiert werden.

1.2 Leitlinie von Tethys RDR

Tethys RDR, abrufbar unter www.tethys.at, ist ein digitales geowissenschaftliches Forschungsdatenrepositorium der Bundesanstalt für Geologie, Geophysik, Klimatologie und Meteorologie (GeoSphere Austria).

In diesem digitalen Daten-Repositorium sollen vorrangig digitale, georeferenzierte Forschungsdaten aus den Fachdisziplinen Geologie und Geophysik, die an bzw. im Auftrag der GeoSphere Austria generiert wurden, publiziert und mit einem DOI versehen nachhaltig archiviert werden.

Die von Tethys RDR bereitgestellten georeferenzierten Datenpublikationen und die dazugehörigen Metadaten können frei eingesehen und – den Nutzungsrechten entsprechend – genutzt werden. Grundsätzlich sind alle Datenpublikationen sichtbar, mit Ausnahme jener, die einem Embargo unterworfen wurden. Diese Daten sind erst zu einem definierten Zeitpunkt verfügbar. Die dazugehörigen Metadaten sind jedoch öffentlich zugänglich.

Tethys RDR besteht aus einem Digital Content Management System (Open Source Software Opus 4), das an die Bedürfnisse eines nachhaltigen geowissenschaftlichen Forschungsdatenrepositoriums angepasst wurde. Die Datenbanken werden auf einem Linux SQL Server gepflegt. Als Such-Server wird die Open-Source-Plattform Solr verwendet.

Zum aktuellen Zeitpunkt wird das Tethys RDR noch vorwiegend in deutscher Sprache bereitgestellt. An einem bilingualen Webauftritt wird bereits gearbeitet.

Die Datenpublikationsrichtlinien von Tethys RDR enthalten als RDR Policy grundlegende Informationen für den Umgang mit Forschungsdaten und stellen eine einheitliche Qualität der Daten sicher. Mit deren

¹ <https://ris.bka.gv.at/geltendefassung.wxe?abfrage=bundesnormen&gesetzesnummer=20011885> (abgerufen am 26.05.2023).

Einhaltung kann eine solide, transparente sowie nachhaltige Datenpublikation gewährleistet werden. Hier wird geregelt, welche Daten wie, in welcher Form und in welchem Umfang publiziert werden können. Ebenfalls wird ein Workflow bereitgestellt sowie über Rechte und Pflichten informiert. Die Verfasser (Creator) der Veröffentlichungen behalten grundsätzlich die geistigen Eigentumsrechte an ihren Daten. Die Einreicher (Submitter) übertragen dem Repotorium mit dem Hochladen und Publizieren der Daten im Tethys RDR die Kopier-, Umwandlungs-, Speicher- und Verbreitungsrechte.

Das Tethys RDR baut auf zwei bereits in der Wissenschaftswelt etablierte Prinzipien auf, die nachhaltiges Archivieren sowie Publizieren von weiterverwertbaren wissenschaftlich relevanten Daten ermöglichen: Open Science Principles^{2, 3} und FAIR Guiding Principles⁴.

Bereits publizierte Daten in Tethys können nicht mehr gelöscht werden, es sei denn, es gibt einen rechtlichen Konflikt. In diesem Fall greift ein interner Ablauf ein, der die Metadaten sowie die Daten auf der Landingpage vollständig entfernt. Stattdessen wird der Grund der Löschung angeführt.

Tethys RDR ist verantwortlich für eine reibungslose Publikation der Daten ab dem Zeitpunkt des Hochladens. Entstehen auf dem Weg zur Publikation unbemerkt technische Fehler, so werden diese mitdokumentiert und über ein internes Fehler-Management abgehandelt und schnellstmöglich korrigiert. Bei Bekanntwerden von inhaltlichen Fehlern werden diese ebenfalls dokumentiert, jedoch auf eine andere Weise abgehandelt. Details dazu werden im [Abschnitt 1.4.5](#) beschrieben.

Mit Tethys RDR sind wir bestrebt, aktiv an der internationalen Forschungsdateninfrastruktur teilzunehmen und der Community die Auffindbarkeit der Daten wesentlich zu erleichtern. Derzeit ist Tethys unter re3data.org⁵, unter OAI⁶ und bei OpenDOAR⁷ registriert. Weitere Registrierungen sind in Vorbereitung.

1.2.1 Open Science Principles / Die offene Wissenschaft

Die Open Science Principles bestehen aus sechs Prinzipien ([Abb. 1](#)). Besonderer Fokus liegt im speziellen auf Open Data⁸ und auf Open Access^{9, 10}.

Als Open Data werden alle frei nutzbaren Daten bezeichnet, die uneingeschränkt benutzt werden dürfen. Um gute wissenschaftliche Praxis zu gewährleisten, wird die Nennung der Urheber (CC BY) und bei Bedarf die Verwendung von Copyleft¹¹ bzw. CC BY-SA¹² als einzige Einschränkung der Nutzungsrechte erlaubt.

Open Access steht für einen offenen, uneingeschränkten Zugang zu digitaler wissenschaftlicher Literatur. Im Fall des Tethys RDR soll bis auf gut begründete Ausnahmen (temporäre Restriktion durch Embargo, eingeschränkte Creative Commons Lizenzen) ein uneingeschränkter Zugang zu den publizierten wissenschaftlichen Daten und den dazugehörigen Informationen gewährt werden.

² https://de.wikibooks.org/wiki/Handbuch_Open_Science (abgerufen am 25.05.2023).

³ https://en.wikipedia.org/wiki/Open_science (abgerufen am 25.05.2023).

⁴ WILKINSON et al. (2016): The FAIR Guiding Principles for scientific data management and stewardship. – Sci. Data 3: 160018. <http://dx.doi.org/10.1038/sdata.2016.18>

⁵ RE3DATA.ORG Registrierung: <https://www.re3data.org/repository/r3d100013400>

⁶ OAI Registrierung: <http://www.openarchives.org/Register/BrowseSites?viewRecord=https://www.tethys.at/doi>

⁷ OPENDOAR Registrierung: <https://v2.sherpa.ac.uk/id/repository/10614>

⁸ https://de.wikipedia.org/wiki/Open_Data (abgerufen am 26.05.2023).

⁹ <http://open-access.net/AT-DE/informationen-zu-open-access/was-bedeutet-open-access/> (abgerufen am 26.05.2023).

¹⁰ https://openaccess.mpg.de/mission-statement_de (abgerufen am 26.05.2023).

¹¹ <https://de.wikipedia.org/wiki/Copyleft> (abgerufen am 26.05.2023).

¹² <https://creativecommons.org/licenses/by-sa/4.0/deed.de> (abgerufen am 02.08.2023).

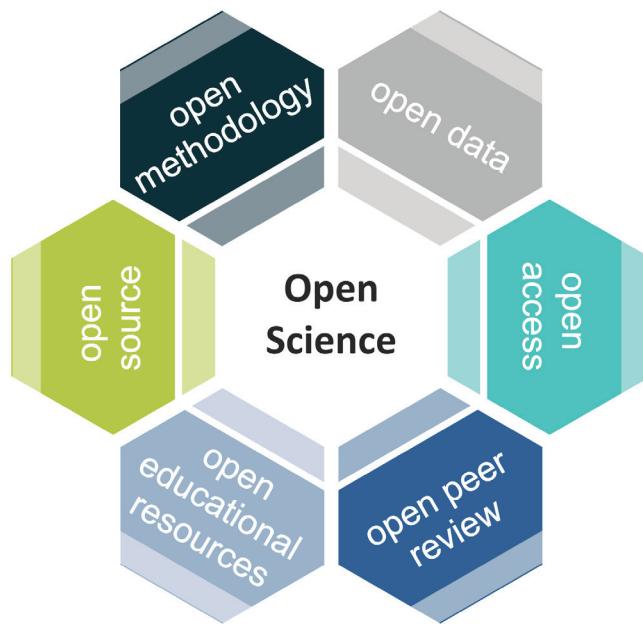


Abbildung 1: Die sechs Prinzipien von Open Science im Überblick.

1.2.2 FAIR-Prinzipien

Die Prinzipien nach FAIR¹³, die ein Akronym für *Findability* (Auffindbarkeit), *Accessibility* (Zugänglichkeit), *Interoperability* (Kompatibilität) und *Re-usability* (Wiederverwendbarkeit) sind (Abb. 2), unterstützen eine nachhaltige wissenschaftliche Datenpflege sowie deren Verwaltung. Erst durch die garantierte Auffindbarkeit und volle Zugänglichkeit zu den Daten, die gewährleistete Kompatibilität mit anderen Datenbeständen sowie die wissenschaftliche Wiederverwendbarkeit wird die Datenpublikation zu einer echten Publikation und erlangt einen großen wie auch nachhaltigen Wert.

Auffindbar

- [1] Die (Meta-) Daten können anhand der eindeutigen IDs mit einem standardisierten Kommunikationsprotokoll abgerufen werden.
- [2] Das verwendete Protokoll ist offen, frei und universal einsetzbar.
- [3] Das Protokoll erlaubt bei Bedarf eine Authentifizierung bzw. Autorisierung.
- [4] Die Metadaten sind vorhanden, auch wenn die Daten selber nicht mehr verfügbar sind.

Zugänglich

- [1] Die (Meta-)Daten sind einem weltweit einzigartigen sowie unvergänglichen Identifier zugeordnet.
- [2] Die Daten sind mit umfangreichen Metadaten beschrieben.
- [3] Die (Meta-)Daten sind in einer durchsuchbaren Ressource registriert bzw. indiziert.
- [4] Die (Meta-)Daten sind fest verknüpft mit dem Identifier.

¹³ <https://www.force11.org/group/fairgroup/fairprinciples> (abgerufen am 07.03.2018).

Interoperabel

- [1] Die (Meta-)Daten verwenden eine formale, zugängliche, gemeinsame und allgemein verwendete Sprache für die Wissensrepräsentation.
- [2] Die (Meta-)Daten verwenden Vokabeln, die dem FAIR Prinzip folgen.
- [3] Die (Meta-)Daten inkludieren korrekte Verweise auf andere (Meta-)Daten.

Wiederverwendbar

- [1] Die (Meta-)Daten haben eine Vielzahl genauer und relevanter Attribute aufzuweisen.
- [2] Die (Meta-)Daten werden mit einer klaren und zugänglichen Nutzungslizenz veröffentlicht.
- [3] Die (Meta-)Daten sind mit der Datenherkunft verbunden.
- [4] Die (Meta-)Daten erfüllen die in der Forschungsgemeinschaft üblichen Standards.



Abbildung 2: FAIR Daten Prinzipien im Überblick, adaptiert von FORCE11¹⁴

1.2.3 Metadatenbeschreibung

Metadatenbeschreibungen sind essenziell, um die Datenpublikationen sowohl innerhalb als auch außerhalb des Research Data Repository such- und auffindbar zu machen. Um einen „Wildwuchs“ zu vermeiden und das Durchsuchen der Daten auf eine systematische Ebene zu bringen, die international verstanden wird, wurden Metadatenstandards entwickelt.

In Tethys RDR werden weitverbreitete Metadatenstandards (Dublin Core¹⁵, DataCite¹⁶, ISO 19139) sowie ein deskriptives Metadatenelement (dataset type) verwendet¹⁷.

¹⁴ <https://www.force11.org/group/fairgroup/fairprinciples> (abgerufen am 07.03.2018).

¹⁵ <https://www.dublincore.org/> (abgerufen am 26.05.2023).

¹⁶ <https://datacite.org/> (abgerufen am 26.05.2023).

¹⁷ <https://www.tethys.at/oai?verb=ListMetadataFormats> (abgerufen am 26.05.2023).

1.2.3.1 Dublin Core

Dublin Core ist ein weit verbreitetes bibliografisches Datenformat zur Beschreibung von Dokumenten bzw. Objekten. Das Dublin Core Schema wurde von „Dublin Core Metadaten Initiative“ (DCMI¹⁵) entwickelt und umfasst 15 Kernfelder, die eine grundsätzliche Beschreibung der Objekte zulassen. Die genaue Definition der einzelnen Elemente können online nachgelesen werden^{18, 19}.

1.2.3.2 DataCite

DataCite, ebenfalls ein internationales Konsortium, stellt ein sehr umfangreiches Metadaten-Schema²⁰ speziell für Datenpublikationen zur Verfügung, um den globalen Zugang für wissenschaftlich relevante Forschungsdaten wesentlich zu erleichtern. Ein großes Ziel des Konsortiums ist auch die nachhaltige, wissenschaftlich nachvollziehbare sowie zitierfähige Publikation von Forschungsdaten. Als einmalige und nachhaltige ID wird von DataCite ein Digital Object Identifier (DOI) bereitgestellt.

Dazu ist eine Mitgliedschaft beim DataCite Konsortium notwendig, die strenge Überprüfungen nach sich zieht.

Die genaue Definition der einzelnen Elemente kann online abgerufen werden¹⁶.

Im Tethys RDR wurde der DataCite Metadatenstandard in vollem Umfang implementiert.

1.2.3.3 ISO 19139

Dieses Metadatenschema beschreibt standardisierte geografische Information²¹.

1.2.3.4 Deskriptives Metadatenset: „Dataset Type“ (Datensatzart)

Zusätzlich zu den bereits beschriebenen Metadatenstandards wurde ein weiteres, rein deskriptives Metadatenset („Dataset Type“) eingeführt. Darunter fallen die sieben Metadatenelemente: Analysis, Measurements, Monitoring, Remote sensing, GIS (Geographic Information System), Models und Mixed Type (Abb. 3).

Diese Metadatenelemente dienen zur Klassifizierung des Datentypus und sollen im Speziellen bei einem angewachsenen Datenbestand das Filtern nach Datentypen erleichtern. Nachfolgend werden die Eigenschaften der Metadatenelemente detailliert beschrieben.

¹⁸ <https://www.dublincore.org/specifications/dublin-core/dc Terms/> (abgerufen am 26.05.2023).

¹⁹ https://de.wikipedia.org/wiki/Dublin_Core (abgerufen am 26.05.2023).

²⁰ <https://schema.datacite.org/> (abgerufen am 26.05.2023).

²¹ <https://www.iso.org/standard/67253.html> (abgerufen am 26.05.2023).

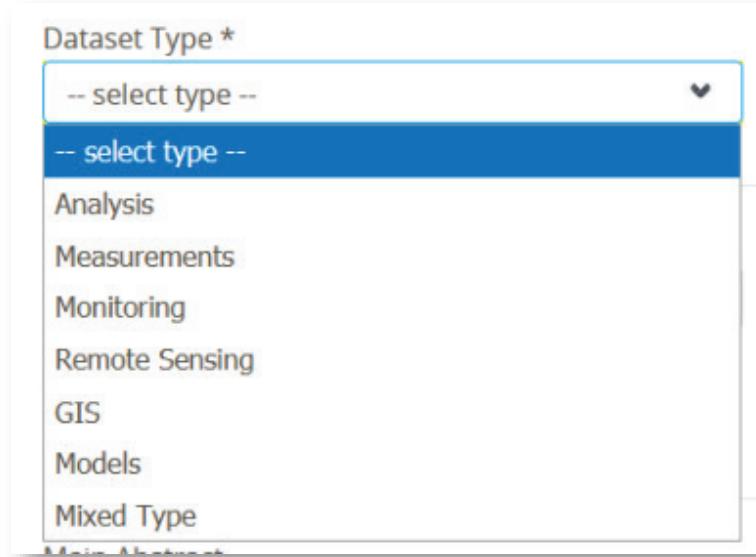


Abbildung 3: Deskriptive Auswahlliste „Dataset Type“.

Analysis

Analysendaten werden bei der Durchführung von Analysen erzeugt und weisen sich durch ihre grundsätzliche Reproduzierbarkeit aus. Analysen geben Rückschlüsse auf Bestandteile und ihre Eigenschaften. Vorherrschende Datenformate für derartige Publikationen sind ASCII Files (z.B. CSV, TXT, DBF).

Beispiele:

- Gebhardt, H., Schenk, B., Enge, A., Coric, S., Ranftl, E., Heinz, P. (2023): Occurrence of planktic foraminifera in drill core NÖ-07 – Diendorf. Tethys RDR, GeoSphere Austria, Wien (<https://doi.org/10.24341/tethys.201>).
- Altenberger, F., Raith, J. (2023): Trace element analyses (LA-ICP-MS) of scheelite from the Felbertal tungsten deposit. Tethys RDR, GeoSphere Austria, Wien (<https://doi.org/10.24341/tethys.210>).
- Geochemische Zusammensetzung von Gesteinen, Altersdatierungen, Zusammensetzung von thermalen Wässern, etc.

Measurements

„Eine Messung ist das Ausführen von geplanten Tätigkeiten zu einer quantitativen Aussage über eine Messgröße durch Vergleich mit einer Einheit.“²² Vorherrschende Datenformate für derartige Publikationen sind ASCII Files (z.B. CSV, TXT, DBF).

Beispiele:

- Messung von Temperatur, Druck, Permeabilität, Porosität, Dichte, elektrische Leitfähigkeit, geophysikalische Messungen etc.

Monitoring

Das sind fachrelevante Daten, generiert durch Analysen oder Messungen, die in Abhängigkeit mit der Zeit gemessen wurden und somit einmalig erfasst werden und dadurch auch nicht mehr als solche

²² Definition entnommen aus DIN 1319.

reproduzierbar sind. Vorherrschende Datenformate für derartige Publikationen sind ASCII Files (z.B. CSV, TXT, DBF).

Beispiele:

- Ottowitz, D., Jochum, B., Pfeiler, S., Römer, A., Supper, R. (2021): Geoelectrical Monitoring – landslide Ampflwang (Upper Austria). Tethys RDR, Geologische Bundesanstalt (GBA), Wien (<https://doi.org/10.24341/tethys.5>).
- Jochum, B., Ottowitz, D., Pfeiler, S., Römer, A., Supper, R. (2021): Geoelectrical Monitoring – landslide Laakirchen (Upper Austria). Tethys RDR, Geologische Bundesanstalt (GBA), Wien (<https://doi.org/10.24341/tethys.178>).
- Messung von Versickerung von Wasser in der Zeit von A bis B; physikalische Messung in der Zeit von A bis B von Massenbewegungen; Rasteranalysen von einem Gebiet XY in Abhängigkeit von Zeit, etc.

Remote sensing (RS)

Im Rahmen der Fernerkundung werden Daten durch berührungslose Erkundung generiert. Als Grundlage und für weitere Messungen und Interpretationen dienen elektromagnetische Wellen oder Schallwellen. Vorherrschende Datenformate für derartige Publikationen sind Rasterdaten in Geopackage (GPKG) und ASCII Files (z.B. CSV, TXT, DBF).

Beispiele:

- Klassifizierte und prozessierte Befliegungsdaten.
- Berechnungen von Wassereinzugsgebieten auf der Grundlage von Rasterdaten.
- Interpretierte Laserscandaten.

Geographic Information System (GIS)

GIS-Daten sind neben den Remote Sensing-Daten ein weiteres Spezifikum von interpretierten Daten und bestehen aus georeferenzierten sowie attribuierten Geometrien. Zusätzlich können andere Datentypen (Analysendaten, Messdaten etc.) enthalten sein. Vorherrschendes Datenformat für derartige Publikationen ist Geopackage (GPKG).

Beispiele:

- Geologische Bundesanstalt Österreich (2022): Geodaten – Bundesland Niederösterreich (1:200.000). Tethys RDR, Geologische Bundesanstalt (GBA), Wien (<https://doi.org/10.24341/tethys.186>).
- Geologische Bundesanstalt Österreich (2022): Geodaten – Blatt 114 Holzgau (1:50.000). Tethys RDR, Geologische Bundesanstalt (GBA), Wien (<https://doi.org/10.24341/tethys.182>).
- Objektarten im Sinne von GeoSciML (Geologische Einheiten, Strukturlinien, Hydrogeologische Objekte, Profile, Bohrungen, Abbaue etc.); Geometrie- und Sachdaten zu geologischen Karten etc.

Models

Ein Modell ist ein vereinfachtes Abbild der Realität, in dem unter anderem auch Veränderungen simuliert werden können. Durch das Quantifizieren von Modelldaten können für die Wissenschaft relevante Daten entstehen. Vorherrschende Datenformate für derartige Publikationen sind Geopackage (GPKG) und ASCII Files (z.B. CSV, TXT, DBF).

Beispiele:

- Porpaczy, C., Bottig, M. (2022): Geological subsurface model of the Austrian Molasse Basin. Tethys RDR, Geologische Bundesanstalt (GBA), Wien (<https://doi.org/10.24341/tethys.191>).
- Porpaczy, C. (2022): Geological subsurface model of Austria and surrounding areas. Tethys RDR, Geologische Bundesanstalt (GBA), Wien (<https://doi.org/10.24341/tethys.184>).
- 3D-Modelle, Modellierte geologische Profile, Tiefenmodellierungen, Versickerungsmodelle etc.

Mixed Type

Unter diesem Metadatenelement werden Datensätze publiziert, die aus zwei oder mehreren verschiedenen Datentypen bestehen (z.B. Analysendaten/interpretierte Daten oder Messdaten/Monitoring-Daten).

Beispiele:

- Rabeder, J., Reitner, H., Wimmer-Frey, I., Benold, C., Heinrich, M., Hobiger, G., Reitner, J., Filzmoser, P., Mert, M. (2021): Geodaten der Rohstoffbezirke und geochemische Analysen österreichischer Löss- und Lösslehmvorkommen. Tethys RDR, Geologische Bundesanstalt (GBA), Wien (<https://doi.org/10.24341/tethys.95>).
- Datensatz zu einer geologischen Karte und eigene Tabelle mit geochronologischen Analysen; Geophysikalische Messung von Grundwasserkörpern und Wasseranalysen; Monitoring einer Hangrutschung und Modellierung einer weiteren Bewegung etc.

1.3 Workflow

Der komplette Arbeitsablauf (Abb. 4) umfasst den gesamten Veröffentlichungsprozess, vom ersten Kontakt über die Aufnahme von Daten und Metadaten bis zur Veröffentlichung des Datensatzes mit einem DOI.

Vor dem eigentlichen Hochladen der kompletten Datenpublikation findet ein Erstgespräch mit einem Tethys-Editor statt. Hier werden die ersten Fragen geklärt, die Daten auf generelle Publikationstauglichkeit geprüft und gegebenenfalls über andere Publikationsmöglichkeiten informiert. Können die Daten im Tethys publiziert werden, bekommt der Einreicher Zugang zum Handbuch und weitere Hilfsmittel, wie das Metadatenformular, zur Verfügung gestellt. Im Workflow-Diagramm (Abb. 4) können die wichtigsten Schritte im Veröffentlichungsprozess nachvollzogen werden.

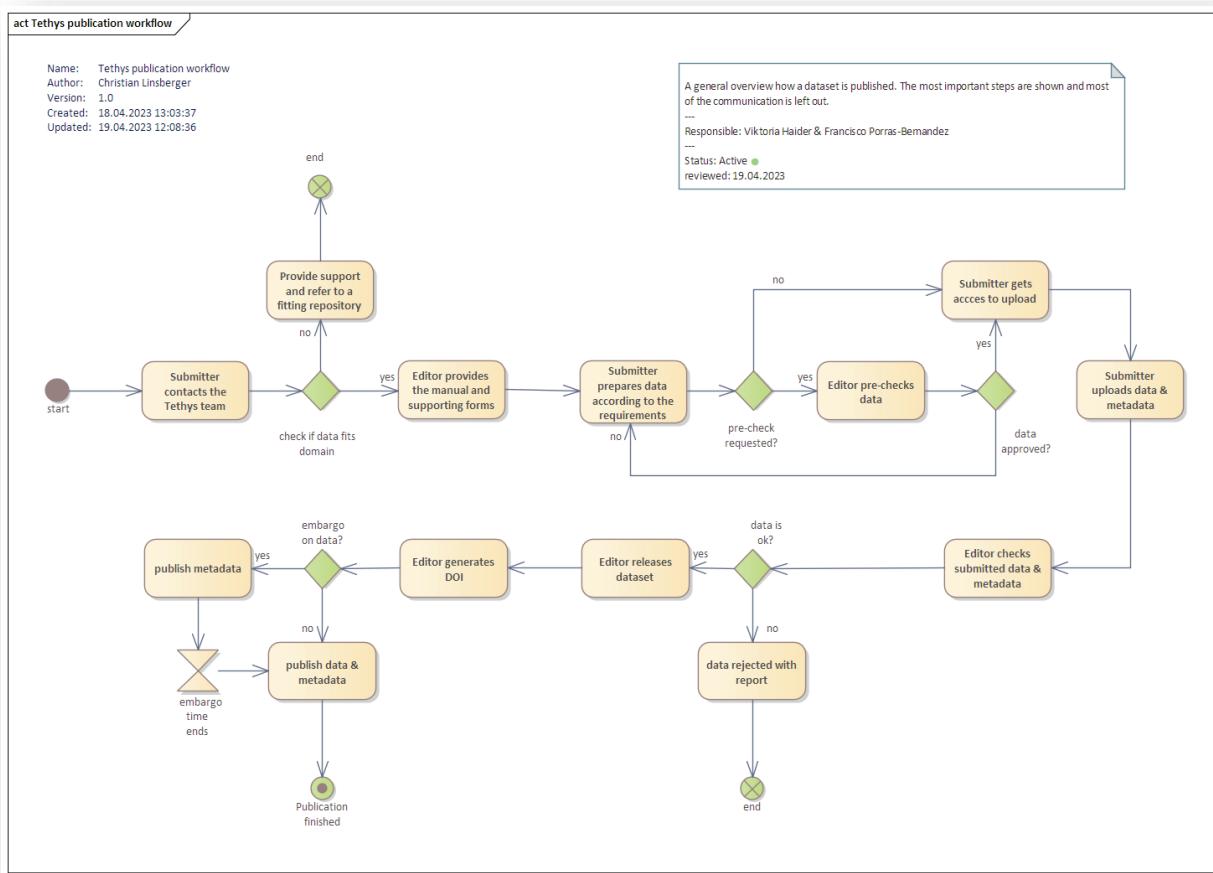


Abbildung 4: Workflow-Diagramm.

1.4 Die Richtlinien für das Publizieren von Forschungsdaten

Das Tethys RDR strebt eine sichere, nachhaltige und auch langfristige Archivierung an. Damit dies auch umgesetzt werden kann, müssen verschiedene Rahmenbedingungen erfüllt werden.

Um einen reibungslosen Ablauf der Datenpublikation zu ermöglichen, wurden die Richtlinien so aufgebaut, dass ein Überblick sehr schnell geschaffen werden kann. Diese wurden in Form eines Fragenkatalogs aufgebaut, um das Abarbeiten der einzelnen Punkte deutlich zu erleichtern. In diesem Katalog befinden sich alle Fragen, die im Zuge der Datenpublikation geklärt werden müssen. Die Beantwortungen der Fragen sind jeweils zweigeteilt. In der eingefärbten Box, direkt nach der Frage, wird diese kurz und kompakt beantwortet. Anschließend gibt es eine vertiefende Beantwortung.

Ziel dieser Strukturierung der sehr umfangreichen Tethys RDR-Richtlinie ist die Verringerung der Vorbereitungszeit für die Datenpublikation. Gleichzeitig wird ein gewisser Überblick gewährleistet.

Am Ende der Datenpublikationsrichtlinie gibt es noch eine Checkliste. Diese kann die Publikationsvorbereitung erleichtern und unterstützt bei der Überprüfung auf Vollständigkeit. Besteht bei einer Aussage eine Unsicherheit oder Unklarheit, so findet man hier den jeweiligen Verweis zu den betroffenen Fragen der Datenpublikationsrichtlinie.

1.4.1 Worin liegt der Mehrwert, die Forschungsdaten mit Hilfe des Tethys RDRs zu publizieren?

Es gibt viele Gründe, die Daten adäquat in einem Forschungsdatenrepositorium zu publizieren. Zusätzlich gibt es einige sehr gute Gründe, die fachrelevanten Forschungsdaten in Tethys zu publizieren. In der folgenden Darstellung sind diese angeführt.

Allgemeiner Mehrwert

- Die Daten werden auf ein Qualitätsniveau gebracht, das eine nachhaltige und sichere Bereitstellung gewährleistet.
- Idealerweise haben diese Forschungsdaten eine internationale Reichweite und werden als solche anerkannt.
- Die Daten bekommen eine persistente ID und sind sowohl permanent auffindbar, als auch zitierbar. Tethys RDR verwendet DOIs als persistente Identifier.
- Interessierte können weltweit auf die Daten zugreifen und im Rahmen der bereitgestellten Lizenzbedingungen diese nachvollziehbar und nach wissenschaftlicher „Best Practice“ weiterverarbeiten.
- Es fördert auf internationaler Ebene auch interdisziplinäre Synergien.
- Es kann den Forschungsförderern gegenüber der Nachweis erbracht werden, dass die Daten unter Berücksichtigung von Open Science und FAIR-Prinzipien für die Nachnutzung zugänglich gemacht werden.
- Wissenschaftlich wertvolle Daten können von der Wissenschaftsgemeinschaft effektiv verwertet werden, ohne neue Ressourcen in die Generierung stecken zu müssen.
- Große Datenbestände können effizient aufgearbeitet werden und dadurch eine schnellere Wissensentwicklung fördern.
- Das Zitieren der Datenpublikation ist eine Anerkennung der eigenen Leistung.
- Die Daten verschwinden nicht in der lokalen oder institutionellen Computerperipherie und gewinnen dadurch auch an Wert.
- Die Metadaten der einzelnen Datenpublikationen sind maschinenlesbar abgelegt. Dadurch können diese in diverse Services eingebunden und somit leichter gefunden und verarbeitet werden.

Spezifischer Tethys-Mehrwert

- Tethys ist ein fachspezifisches Forschungsdatenrepositorium für die Geowissenschaften.
- Die Daten werden alle ausnahmslos räumlich verortet.
- Die zu publizierenden Datensätze und Metadaten werden kuratiert.
- Voller Support bei der Datenpublikation.
- Sowohl Tethys als auch die darin archivierten Daten werden innerhalb der GeoSphere Austria sicher bereitgestellt (gehostet).
- Neben den Datensätzen können ergänzende Datenbeschreibungen publiziert werden.
- Tethys wird laufend zu Gunsten der Benutzenden und der Datenverfasser und im Einklang mit den Forschungsförderern weiterentwickelt.

1.4.2 Welche Daten werden im Forschungsdatenrepositorium aufgenommen?

Im Tethys RDR werden noch nicht als Forschungsdaten veröffentlichte „neue“, fachrelevante, georeferenzierte, aufbereitete sowie strukturierte Daten publiziert. Ergänzend zu den Datensätzen können Grafiken, Bilder oder Texte mitpubliziert werden, vorausgesetzt diese tragen zu einem besseren Verständnis der Daten bei. Die Entscheidung über die Datenaufnahme obliegt grundsätzlich dem Tethys RDR-Redaktionsteam.

„Forschungsdaten bilden unumstritten einen Grundpfeiler wissenschaftlicher Erkenntnis und können die Grundlage für weiterführende Forschung sein. Die Dokumentation von Forschungsdaten nach fachspezifischen Standards und ihre langfristige Archivierung sind daher nicht nur bedeutsam für die Qualitätssicherung wissenschaftlicher Arbeit, sondern auch eine grundlegende Voraussetzung für die Nachnutzbarkeit von Forschungsergebnissen.“²³

Forschungsdaten: Forschungsdaten können nicht eindeutig beschrieben und somit auch nicht scharf von „anderen Daten“ abgegrenzt werden. Damit lassen sie sich nur sehr schwer im allgemeinen Sinne definieren. Jede wissenschaftlich arbeitende Fachdisziplin hat ihre eigenen Anforderungen an die Daten und arbeitet mit individuellen Grundlagen. Damit lassen sich Forschungsdaten nicht grundsätzlich auf eine einzige Definition reduzieren. Eines haben Forschungsdaten jedoch gemeinsam: Forschungsdaten können die reproduzierbare bzw. nachvollziehbare und strukturierte Basis zur Erlangung wissenschaftlicher Kenntnis sein, also sind sie noch unprozessierte Datensätze. Sie können aber auch bereits reproduzierbar verarbeitet sein und auf den ihnen zugrundeliegenden Daten aufbauen. Forschungsdaten können aber auch schon prozessiert sein und gleichzeitig für einen neuen Forschungsansatz unprozessierte Basisdaten sein. Gerade diese Eigenschaft der Forschungsdaten, nämlich gleichzeitig Basis und Ergebnis sein zu können, macht diese Daten erhaltungswürdig und potenziell wertvoll. Die Herausforderung dabei ist die Wahrung der Transparenz und Nachvollziehbarkeit bzw. auch die Reproduzierbarkeit. Das sind die Grundpfeiler von seriösem wissenschaftlichem Arbeiten, welche in der wissenschaftlichen „best practice“ gepflegt werden. Im Verständnis von Tethys RDR sind Forschungsdaten strukturierte Datensätze, die zusammen mit der Datenerläuterung und weiteren Datensupplements publiziert werden.

Neu: Tethys RDR fungiert nicht als Back Up System. Wenn neue Daten zu einer bereits vorhandenen Datenpublikation generiert werden und diese die alten Datensätze ersetzen sollen, so können die neuen Daten in Form einer Versionierung publiziert werden. Die Aktualisierung soll in einem Verhältnis zum Mehrwert für die Nutzenden stehen.

Fachrelevant: Tethys RDR startete als Forschungsdatenrepositorium für die Geologische Bundesanstalt (mittlerweile GeoSphere Austria). Daher werden alle geologischen sowie geophysikalischen Forschungsdaten publiziert, die eine naheliegende fachliche Relevanz für die Institution haben. Forschungsdaten, die im Zuge einer Kooperation mit der GeoSphere Austria entstehen, können ebenfalls über Tethys RDR publiziert werden.

Georeferenziert: Alle im Tethys RDR veröffentlichten Datensätze sind obligatorisch georeferenziert. Der allgemeine Raumbezug des Untersuchungsgebietes wird durch die Metadaten beschrieben. Die einzelnen Verordnungspunkte können in den Datensätzen selber abgebildet werden.

Aufbereitet: Die im Tethys RDR zu publizierenden Forschungsdaten sollen soweit aufbereitet vorliegen, dass die klassischen Nutzenden die publizierten Forschungsdaten mit der dazugehörigen

²³ https://www.dfg.de/foerderung/info_wissenschaft/2015/info_wissenschaft_15_66/index.html (abgerufen am 26.05.2023).

Datenerläuterung und den Supplementen verstehen und verwenden können, ohne zusätzliche Informationen von den Datenerstellern einholen zu müssen.

Strukturiert: Gute Datensätze unterliegen grundsätzlich einer nachvollziehbaren Struktur und bieten einen geeigneten und schnellen Überblick. Das ermöglicht auch eine individuelle Weiterverarbeitung. Je schlechter und inkonsequenter die Datensätze strukturiert sind, umso leichter werden diese als potenzielle Datensätze übersehen und ignoriert. Je einfacher die Struktur, desto eher können diese auch von Fachfremden interdisziplinär weiterverarbeitet werden.

Ergänzende Files zu den Datensätzen: Ergänzend zu den Datensätzen können Fotos, Bilder, Grafiken und Texte mitpubliziert werden, sofern diese zum Verständnis der Datensätze beitragen. Zur Wahrung des guten Überblicks und zur Vermeidung von Missverständnissen soll hier mit Maß und Ziel vorgegangen werden, quasi nach dem Prinzip „so viel wie nötig, so wenig wie möglich.“ Weitere Textinformationen (z.B. Projektberichte) können bei Bedarf ebenfalls zusammen mit den Daten publiziert werden, sofern dies für die Datennachnutzung sowie der Nachhaltigkeit dienlich ist. Textinformationen werden nur im Rahmen der Datenpublikation veröffentlicht, nicht jedoch als eigenständige Publikation, da sie unter anderem auch mehrmals publiziert werden können. Die im Tethys RDR mitpublizierten Texte werden nicht in vollem Umfang redaktionell bearbeitet und haben demnach auch nicht die garantierte Qualität einer klassischen Textpublikation. Werden Texte in klassischen Verlagen publiziert, so können diese mit den Daten im Tethys RDR verlinkt werden.

1.4.3 Welche Informationen müssen in der Datenpublikation vorhanden sein?

Die Datenpublikation besteht schematisch aus drei Teilen: Der erste Teil umfasst die Metadaten. Diese werden im Zuge des Uploads der Datenpublikation abgefragt bzw. generiert. Der zweite Teil befasst sich mit der optionalen Datenerläuterung. Der dritte Teil besteht aus den eigentlichen Datensätzen.

Metadaten: Die Metadaten sind das Herzstück einer jeden digitalen Publikation. Ihre Funktion ist die Daten auffindbar, sichtbar sowie suchbar zu machen. Werden standardisierte Metadatenelemente verwendet, so kann auch automatisiert gesucht werden und die Datenpublikation auch außerhalb des Repositoriums zur Verfügung gestellt werden. Tethys RDR verwendet standardisierte Metadatenelemente und unterstützt die Maschinenlesbarkeit.

Datenerläuterung: Die Datenerläuterung enthält die Beschreibung der publizierten Datensätze, die über die standardisierte Beschreibung durch die Metadatenelemente hinausgeht. Sie ist damit auch ein essenzieller Bestandteil der Publikation, wenn die Daten mit den Metadaten nicht vollumfänglich beschrieben werden können. Ziel der Datenerläuterung ist es, Daten so zu beschreiben, dass diese auch noch in Zukunft – mindestens 10 Jahre, aber idealerweise deutlich darüber – verstanden und nach wissenschaftlicher „best practice“ genutzt werden können. Die Datenerläuterung kann entfallen, wenn es eine vollwertige Textpublikation gibt, welche diese Aufgabe übernimmt und diese für die User zugänglich ist. Der Link zu dieser Textpublikation muss dann in den Metadaten verpflichtend angegeben werden. Fehlt eine selbsterklärende und verständliche Datenbeschreibung, so werden die Daten im Tethys RDR nicht publiziert.

Datensätze: Das Tethys RDR archiviert und publiziert ausschließlich Datensätze. Datensätze zeichnen sich dadurch aus, dass sie aus mehreren logisch zusammengehörenden Daten bestehen²⁴ und diese ohne weitere händisch oder/und technisch durchgeführte Zwischenschritte erst als solche aufbereitet werden müssen.

²⁴ <https://www.techfacts.de/ratgeber/was-sind-daten> (abgerufen am 30.05.2023).

1.4.4 Können bereits im Tethys RDR publizierte Forschungsdaten wieder gelöscht werden?

Sobald die Daten publiziert sind, können diese nicht mehr verändert, ergänzt oder gelöscht werden. In Ausnahmefällen kann der Zugriff auf Datensätze gesperrt werden, die Metadaten bleiben ausnahmslos erhalten, da die Zitierfähigkeit sowie die wissenschaftliche Nachvollziehbarkeit zu jeder Zeit gewährleistet werden muss.

In Tethys RDR publizierte Daten werden wie Printpublikationen gehandhabt. Sobald die Daten veröffentlicht wurden, können weder inhaltliche Korrekturen, noch Korrekturen an den Metadaten vorgenommen werden. Daher ist es von großer Bedeutung, die Datenpublikation sorgfältig vorzubereiten und gut zu überlegen, wie die Datensätze und unter welcher Lizenzbedingung diese publiziert werden.

Es ist grundsätzlich nicht vorgesehen, dass die an die Datenpublikation gebundene Metadatenbeschreibung von Tethys RDR verändert, gesperrt oder gelöscht wird. In Ausnahmefällen können die publizierten Datenfiles gesperrt werden, sodass diese nicht mehr eingesehen, heruntergeladen bzw. verbreitet werden können.

1.4.5 Können nachträglich entdeckte Fehler in der Datenpublikation korrigiert werden?

Fehler innerhalb der publizierten Datensätze können nicht verändert bzw. korrigiert werden. Allerdings kann der korrigierte Datensatz als eine Version publiziert werden.

Handelt es sich um Rechtschreibfehler innerhalb der Metadaten, können diese korrigiert werden. Sobald es zu einer Veränderung der Information durch das Hinzufügen oder Entfernen dieser kommt, so muss eine neue Version publiziert werden.

1.4.6 Können extern erstellte Forschungsdaten publiziert werden?

Zum derzeitigen Stand werden ausschließlich Forschungsdaten in das Tethys RDR aufgenommen, welche von und in Kooperation mit der GeoSphere Austria generiert wurden.

Zum Zeitpunkt August 2023 ist die Aufnahme extern generierter Forschungsdaten in das Tethys RDR nicht vorgesehen. Die Möglichkeit einer Erweiterung des Tethys RDR zu einem späteren Zeitpunkt für die Publikation externer Forschungsdaten ist gegeben und daher nicht kategorisch ausgeschlossen.

1.4.7 Kann ich die Forschungsdaten publizieren, wenn mehrere Personen diese erstellt haben?

Sofern die beteiligten Personen davon nachweislich in Kenntnis gesetzt wurden und diese in die Datenpublikation eingewilligt haben, ist dies möglich.

Der Einreicher (Submitter) übernimmt die Verantwortung für die Datenpublikation und ist somit auch verpflichtet, von seinen Co-Erstellern (Creators) eine Einverständniserklärung einzuholen. Der Tethys RDR-Editor kann sich diese Einverständniserklärung vorlegen lassen bzw. die mitangeführten Co-Ersteller persönlich über die bevorstehende Datenpublikation in Kenntnis setzen.

1.4.8 Welche Lizenzen stehen für die Datenpublikation zur Verfügung?

Für die Datenpublikation stehen ausschließlich die Open Data Lizenzen nach Creative Commons²⁵ (Tab. 1) zur Verfügung. Das Konzept von Creative Commons umfasst sieben verschiedene Lizenztypen, welche die Daten unterschiedlich weit öffnen (Abb. 5). Die Lizenz CC BY (nur Namensnennung) ist die einzige CC-Lizenz, die zu 100 % mit den Open Science Prinzipien konform ist. Diese sollte auch grundsätzlich verwendet werden. Deckt diese Lizenz die Notwendigkeiten nicht ab, so steht mit CC BY-SA eine eingeschränktere Lizenzierung zur Verfügung. Die Details dazu sind im unten angeführten Text enthalten.

Achtung! Sind in die Daten Fremddaten eingeflossen, so ist zu prüfen, ob die vorhandene Nutzungslizenz mit der zu vergebenen Lizenz kompatibel ist (Abb. 5).

Die Creative Commons Initiative ist eine gemeinnützige Organisation, die im Jahr 2001 Open-Content-Lizenzen entwickelte, um den Schöpfer eines Werkes (z.B. Datenerstellung) ein Mittel in die Hand zu geben, um sicher und verständlich die Rechte über sein Werk zu verteilen. Dadurch muss kein eigenes Nutzungsrecht formuliert werden und von der Anwenderseite nicht jedes Mal die Erlaubnis für die Verarbeitung des Werkes eingeholt werden. Creative Commons Lizenzen sind international bekannt und anerkannt.

Creative Commons stellt verschiedene Nutzungslizenzen unterschiedlicher Einschränkung zur Verfügung (Abb. 5). CC BY 4.0 und CC BY-SA (Tab. 1) sind jene Lizenzen, die mit den Open Data Prinzipien konform sind²⁶, uneingeschränkten Wissenstransfer fördern und grundsätzlich für Forschungsdaten anzuwenden sind. CC BY-SA verhindert bereits die freie Wahl der Lizenz für Derivate des publizierenden Werkes und sollte nur dann eingesetzt werden, wenn ein konkreter Bedarfsfall vorliegt.

Tabelle 1: Unterstützte Creative Commons Lizenzen.

CC Code	Link zur Definition	Kurzbeschreibung
CC-BY	http://creativecommons.org/licenses/by/4.0	Nur Namensnennung
CC-BY-SA	http://creativecommons.org/licenses/by-sa/4.0	Namensnennung; Weitergabe unter gleicher Bedingung

Nach erfolgreicher Veröffentlichung der Forschungsdatenpublikation können einmal vergebene Lizenzen nicht mehr verändert werden. Daher ist im Vorfeld gründlich zu prüfen, welche Lizenz für die zu publizierenden Forschungsdaten eingesetzt werden soll. Die gewählte Lizenz kann auch einen Einfluss auf die darauf aufbauenden Daten- oder Textwerke haben.

²⁵ <https://creativecommons.org> (abgerufen am 30.05.2023).

²⁶ https://de.wikipedia.org/wiki/Open_Data (abgerufen am 30.05.2023).

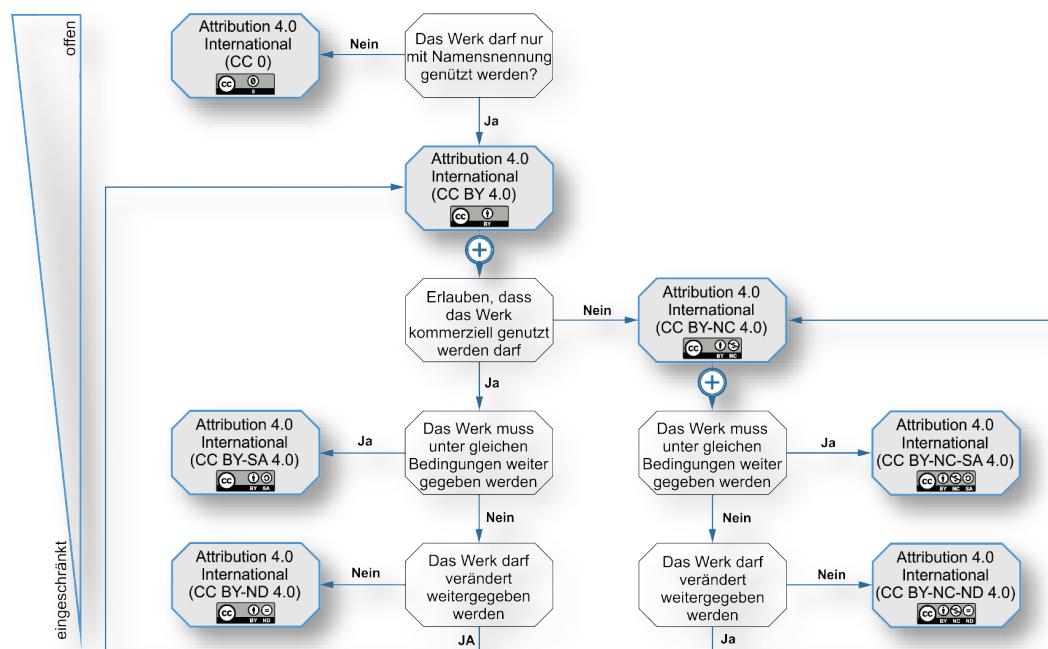


Abbildung 5: Generell mögliche CC Lizenzvarianten, geordnet nach Grad der Offenheit.

1.4.9 Was muss ich machen, wenn die Daten auf Fremddaten basieren?

Es müssen alle Rechte der verwendeten Fremddaten vollständig geklärt sein. Das bestehende Nutzungsrecht muss mindestens eine Weiterverarbeitung erlauben. Zusätzlich ist zu prüfen, ob das Nutzungsrecht mit der zu vergebenden Creative Commons Lizenz kompatibel ist.

Werden bei der Erstellung der Daten unter anderem auch Fremddaten verarbeitet, so ist zu überprüfen, welche Urheber- bzw. Nutzungsrechte an die Fremddaten gebunden sind. Wurden bereits die Fremddaten mit Creative Commons Lizenzen versehen, so kann anhand der Lizenzen abgeleitet werden, ob die Daten verwertet werden dürfen bzw. unter welcher Lizenz die neu entstandenen Daten publiziert werden können.

Wurden die Fremddaten mit individuellen Urheber- bzw. Nutzungsrechten belegt, so ist genauestens zu prüfen, ob die Daten verarbeitet werden dürfen und mit welcher Creative Commons Lizenz diese Rechte kompatibel sind. Im Zweifelsfall ist dies mit Hilfe von Rechtsexperten zu klären.

Wurden mehrere unterschiedlich lizenzierte Fremddaten verwendet, so müssen alle Nutzungsbedingungen allumfassend betrachtet werden, um die neu erstellten Daten korrekt zu lizenziieren.

1.4.10 In welchen Formaten darf ich die Daten publizieren?

Es gibt eine Gruppe von Formaten, welche die Kriterien „Offenheit“, „Sicherheit“ sowie „Nachhaltigkeit“ erfüllen. Derzeit werden folgende Formate von Tethys RDR für die Daten und dessen Supplements unterstützt: `ascii-grid`, `csv`, `gpkg`, `jpeg`, `pdf (PDF/A)`, `png`, `shp`, `txt`, `xlsx`.

„Die Wahl des richtigen digitalen Formats ist essentiell für eine nachhaltige Aufbewahrung wissenschaftlicher Daten (z.B. Forschungsdaten). Die Formatierung sollte den nationalen und

internationalen Vorgaben der jeweiligen Fachdisziplin entsprechen. Um die Nachnutzbarkeit digitaler Daten sicherzustellen haben sich im Laufe der letzten Jahre einige Formate bewährt.“²⁷

Tabelle 2: Auflistung von unterstützten und geplant unterstützten Formaten.

Unterstützte Formate	Anmerkung	Referenz
Textfiles		
ASCII	TAB-getrennte TEXT-Dateien (z.B. txt, csv, tab, asc)	https://en.wikipedia.org/wiki/ASCII https://desktop.arcgis.com/de/arcmap/10.3/manage-data/raster-and-images/esri-ascii-raster-format.htm
GML	Unterstützung geplant ISO TC 211	https://de.wikipedia.org/wiki/Geography_Markup_Language https://www.spotzi.com/en/help-center/what-is-a-gml/
DXF	AutoCAD mit ASCII Support	https://en.wikipedia.org/wiki/AutoCAD_DXF http://www.crlf.de/Verlag/DXF-intern/DXF-intern.html
Dokumente		
PDF/A	ISO19005	http://en.wikipedia.org/wiki/PDF/A
ODF	ISO26300	http://en.wikipedia.org/wiki/OpenDocument
MS Office files – standard OOXML	ISO/IEC 29500:2008; seit Office 2013	https://de.wikipedia.org/wiki/Microsoft_Office
Bilder		
JPEG	ISO10918-1:1994	https://en.wikipedia.org/wiki/Jpeg https://www.iso.org/standard/18902.html
PNG	ISO/IEC 15948	https://en.wikipedia.org/wiki/Portable_Network_Graphics
Geodaten		
Shapefile (SHP)		https://www.esri.com/content/dam/esrisites/sitecore-archive/Files/Pdfs/library/whitepapers/pdfs/shapefile.pdf https://de.wikipedia.org/wiki/Shapefile
GeoPackage (GPKG)		https://de.wikipedia.org/wiki/GeoPackage
Models		
NetCDF (CDF, NC)	Unterstützung geplant Panoply netCDF, HDF und GRIB Data Viewer (Panoply netCDF, HDF und GRIB Data Viewer)	http://en.wikipedia.org/wiki/NetCDF

²⁷ <https://www.forschungsdaten.org/index.php/Dateiformat> (abgerufen am 13.02.2020).

Damit Daten auch nach 10 Jahren und deutlich später lesbar und somit auch verwendbar bleiben, ist es nötig, diese in den adäquaten Formaten zu publizieren. Es sind Formate grundsätzlich auszuschließen, die spezifisch nur von einem oder wenigen ausgewählten proprietären Softwaretypen gelesen werden können und spätestens bei den Software-Updates ihre Gültigkeit und somit auch ihre uneingeschränkte Lesbarkeit verlieren. Dies würde auch stark im Widerspruch mit den FAIR Guiding Principles stehen.

Daher sind Formate zu verwenden, die offene Standards umsetzen und bei denen keine Abhängigkeiten zur proprietären Software bestehen. Selbstgestrickte und womöglich auch verschlüsselte Formate werden ebenfalls als Datenpublikation für Tethys RDR kategorisch ausgeschlossen. Hier empfiehlt sich die Konvertierung der Datensätze in zugelassene Formate.

Die derzeit unterstützten Formate sowie die geplanten Formate werden in [Tabelle 2](#) aufgelistet und geben einen Überblick über die Eignung betreffend Langzeitarchivierung. [Tabelle 3](#) gibt einen Überblick über die Nichteignung von Formaten betreffend Langzeitarchivierung. Diese Tabellen haben keinen Anspruch auf Vollständigkeit. Sie werden auch laufend an den aktuellen Stand der Technologie angepasst.

Tabelle 3: Ausgewählte Beispiele für nicht unterstützte Formate.

Document Format ArcGIS (mxd)
FileGeodataBase (gdb)
MS Excel < 2013 (xls)
PersonalGeodataBase (mdb)
Projektdatei ArcGIS Pro (aprxF)
QGIS (qgz)
Skriptsprachen
Proprietäre Formate im Allgemeinen

1.4.11 In welcher Sprache kann ich die Daten publizieren?

Tethys RDR ermöglicht eine Datenpublikation in englischer oder deutscher Hauptsprache. Die Abstracts wie auch der Haupttitel sollen unabhängig von der gewählten Publikationssprache in beiden Sprachen bereitgestellt werden. Die Publikation in Englisch wird empfohlen.

Grundsätzlich wird nahegelegt, die Datensätze in Englisch zu publizieren, sodass eine internationale Reichweite gewährleistet werden kann. Unabhängig von der Publikationssprache müssen die Datenpublikationen sowohl einen englischen, als auch einen deutschen Haupttitel und Abstract zur Verfügung stellen.

Schon am Beginn des Publikationsprozesses ist die Sprache zu wählen. Wurde diese ausgewählt, kann nicht mehr gewechselt werden. Je nach gewählter Sprache ist die Publikation einheitlich in einer Sprache zu beschreiben.

Neben dem klassischen Abstract kann zusätzlich noch ein Kapitel zu den angewandten Methoden und/oder zur technischen Beschreibung hinzugefügt werden. Es wird jedoch ausdrücklich empfohlen, detailliertere Beschreibungen in Form eines pdf zusammen mit den Datensätzen zu publizieren.

1.4.12 Kann ich meine Datenpublikation einem Projekt zuordnen?

Es besteht die Möglichkeit, die Datenpublikation einem Projekt zuzuordnen. Wenn das Projekt bereits angelegt wurde, kann es beim Upload der Datenpublikation ausgewählt werden. Wurde das Projekt noch nicht bereits für eine andere Publikation angelegt, so muss man die Projektdaten (Projektname, Projektbezeichnung, Beschreibung) dem Tethys RDR-Editor mitteilen.

Bevor die Datenpublikation einem bestimmten Projekt zugeordnet werden kann, muss der Editor zunächst das Projekt anlegen. Hierfür muss vor dem Upload der Datenpublikation dem Editor der Projektname, die Projektbezeichnung und eine kurze Beschreibung des Projektes mitgeteilt werden (< 100 Wörter). Nach dem Eintragen des Projektes in die Datenbank, kann direkt beim Upload-Prozess der Datenpublikation das Projekt in der Auswahlliste gefunden und ausgewählt werden.

1.4.13 Kann ich ein Embargo setzen und was muss ich dazu alles berücksichtigen?

Die Datenpublikation kann bis zu maximal zwei Jahre vor Freigabe der Daten ohne Downloadmöglichkeit der Daten publiziert werden.

Es gibt verschiedene Gründe, die eine Sperre der Datenpublikation notwendig machen. So ist es möglich, dass ein Zeitschriftenverlag die Vorabveröffentlichung einer Datenpublikation einfordert, bevor die Textpublikation freigegeben wird oder der wirtschaftlich orientierte Forschungsförderer die generierten Daten zunächst (maximal zwei Jahre) für sich beanspruchen will.

Sollte ein Embargo notwendig sein, können die Daten und die dazugehörigen Metadaten schon im Tethys RDR publiziert werden, allerdings werden die Daten erst mit dem Ende des Embargos automatisch freigegeben. Die Datenpublikation selbst, mit den dazugehörigen Metadaten, kann aber schon aufgefunden und gelesen werden.

Der Vorteil derartiger Vorab-Publikationen ist das Absichern der eigenen Forschungsdaten. Diese können dadurch nicht durch Dritte widerrechtlich veröffentlicht werden.

Vor dem Setzen des Embargos ist zu prüfen, ob es notwendig ist und nicht im Widerspruch mit den Interessen des Fördergeldgebers steht.

1.4.14 Muss ich meine Datenpublikation mit Keywords publizieren?

Die Keywords (Schlagworte) ermöglichen dem Nutzer ein schnelleres sowie leichteres Finden der potenziellen Datenpublikation. Es müssen mindestens drei Keywords abgegeben werden.

Keywords haben den Zweck, die Datensuche innerhalb des Repositoriums sowie in allen Services, in denen die Daten bereitgestellt werden, zu erleichtern. Daher ist es essenziell, repräsentative sowie auch unmissverständliche Keywords zu verwenden.

Da in der Datensuche auch der Titel der jeweiligen Publikation miteinbezogen wird, ergibt es wenig Sinn, Wörter aus dem Titel als Keywords zu verwenden. Es wird daher auch davon abgeraten.

Aktuell werden vom Tethys RDR nur „unkontrollierte Keywords“ unterstützt. An einem kontrollierten Vokabular wird bereits gearbeitet. Die Auswahlliste soll demnächst für die Datenpublikation zur Verfügung stehen. Das heißt, jeder Submitter muss selbstständig entscheiden, welche Keywords die Datenpublikation beschreiben und wie sehr sich die Keywords voneinander unterscheiden, sodass die Daten mit maximaler Wahrscheinlichkeit von Suchenden aufgefunden werden können.

1.4.15 Gibt es ein Upload-Limit?

Derzeit wird von Tethys RDR ein Upload von 512 MB unterstützt. Die Anzahl der Einzelfiles soll so gering wie möglich gehalten werden, um eine Nutzerfreundlichkeit für die Anwender zu gewährleisten.

Der Download der Datenpublikation erfolgt nicht nur innerhalb der GeoSphere Austria, sondern weltweit mit unterschiedlichen technischen Voraussetzungen. Um einen schnellen und unkomplizierten Download zu ermöglichen, wurde der Datenumfang auf eine Gesamtgröße von 512 MB limitiert. Sollten größere Datenmengen publiziert werden, so ist mit dem Tethys RDR Team Kontakt aufzunehmen.

1.4.16 Kann ich die Datenpublikation auf bereits existierende Publikationen verweisen?

Es besteht beim Submission-Prozess die Möglichkeit, sowohl auf bereits bestehende Daten-, als auch auf Textpublikationen zu verweisen. Dazu können mehrere Beziehungen definiert werden (Tab. 4 bzw. Tab. 5).

Dabei ist zu beachten, dass nur auf nachhaltige Referenzen Bezug genommen werden soll. Andernfalls führen die bereitgestellten Referenzen ins Leere.

Nachhaltige Referenz: Damit sind Referenzen gemeint, die einen persistent Identifier (PID) aufweisen. Solche Identifier sind im Tethys RDR unter dem Metadatenelement „Dataset References“ als Identifier Types genau definiert und beschrieben (Tab. 4).

Tabelle 4: Definition der zur Verfügung gestellten Identifier Types.

Abkürzung	Definition
DOI	Ein Digital Object Identifier (DOI; auf Deutsch: Digitaler Objektbezeichner) ist ein eindeutiger und dauerhaft digitaler Identifikator für physische, digitale oder abstrakte Objekte. Er wird vor allem für Online-Artikel von wissenschaftlichen Fachzeitschriften verwendet. Das DOI-System baut auf dem Handle-System auf und ist, vereinfacht ausgedrückt, mit ISBN und ISSN vergleichbar, geht jedoch durch seine integrierte Lokalisierungsfunktion darüber hinaus. ²⁸ Beispiel: https://doi.org/10.24341/tethys.183
HANDLE	Ein Handle bezeichnet einen eindeutigen Referenzwert (Persistent Identifier, PID) einer zu verwaltenden Ressource. Handles werden insbesondere bei der Langzeitarchivierung von Daten oder der Datenarchivierung verwendet. Darüber hinaus sind Handles stets mit einer URL verknüpft und können über den globalen Auflösungsdienst des Handle-Systems aufgelöst werden. ²⁹ Beispiel: https://hdl.handle.net/20.500.11756/dd4424fd
ISBN	Die Internationale Standardbuchnummer (ISBN) ist ein eindeutiges Nummernkennzeichen von Büchern und anderen redaktionell bearbeiteten Veröffentlichungen. Beispiel: 978-3-96201-027-0
URL	Eine URL ist ein durch einen Web-Identifier (wie etwa http) definiertes Schema. Innerhalb jedes Schemas kann es beliebig aufgeteilte Namensräume (also Bezeichnungen) geben. ³⁰ Als URL (Abk.: Uniform Resource Locator) versteht man eine definierte Adresse, welche auf die Position einer Datei auf einem Server zeigt und diese abruft. URLs werden in einem Webbrowser eingegeben, um auf Dokumente im Web zuzugreifen oder werden als Hyperlinks innerhalb eines Dokumentes eingebettet. ³¹ Musteraufbau: http://www.domainname.at/verzeichnis1/datei1.html

²⁸ <https://www.forschungsdaten.org/index.php/DOI> (abgerufen am 05.12.2019).

²⁹ <https://www.forschungsdaten.org/index.php/Handle> (abgerufen am 05.12.2019).

³⁰ <https://fischerlaender.de/lexikon/url-uri-urn> (abgerufen am 06.12.2019).

³¹ <https://de.ryte.com/wiki/URL> (abgerufen am 06.12.2019).

URN	Ein Uniform Resource Name (URN) ist ein dauerhafter, ortsunabhängiger Bezeichner für eine Ressource (z.B. Datenpublikation). ³² Das URN-Schema: <i>urn:<NID>:<NID-spezifischer Teil></i>
-----	--

Im Zuge des Datenpublikationsprozesses besteht die Möglichkeit, auf bereits vorhandene Publikationen zu referenzieren. Hierfür können sowohl unterschiedliche ID-Typen (*Identifier Types*) als auch Beziehungstypen (*Relation Types*) ausgewählt werden (Tab. 5).

Tabelle 5: Definitionen und Beispiele von möglichen Beziehungstypen.

Relation Types	Bedeutung	Beispiele bzw. Anmerkungen
IsSupplementTo	Es werden Datensätze zu einer bestimmten Textpublikation als Supplement publiziert.	Datenpublikation https://doi.org/10.1594/PANGAEA.746803 „Is Supplement To“ https://doi.org/10.1016/S0031-0182(03)00619-9
IsSupplementedBy	Werden eigenständige Datenreports bzw. Datenpapers in einem Data Journal oder einem anderen Verlag publiziert, so kann diese als „Supplement“ an die Datenpublikation angehängt werden.	Beispiel für ein Data Journal: Earth System Science Data (https://www.earth-system-science-data.net/about/aims_and_scope.html)
IsContinuedBy	Werden ältere Daten zu bereits veröffentlichten Datensätzen publiziert, die aufeinander aufbauen, so kann auf diese referenziert werden.	z.B. für Monitoringdaten (Zeitabhängigkeit)
Continues	Es bauen die zu publizierenden Datensätze auf schon veröffentlichte Datensätze auf.	z.B. für Monitoringdaten (Zeitabhängigkeit)
IsNewVersionOf	Die Datenpublikation ist eine neue Version von einer schon bereits publizierten Version.	Hier ist darauf zu achten, dass die Publikation der neuen Version in Verhältnis zu dem Mehrwert für den User steht. Bei der Generierung von zu vielen Versionen kann es nachteilig für die weitere Datenverwertung sein.
IsPartOf	Wurden die Daten bereits als Teil von einer größeren Publikation publiziert, so kann auf diese verwiesen werden.	
HasPart	Kann bei einer Teilkompilation verwendet werden, wo neue Daten mit schon vorhandenen publizierten (Fremd-)daten kombiniert werden.	Es ist darauf zu achten, dass die Rechte von den Erstellern der verwendeten Fremddaten nicht verletzt werden.
Compiles	Ist die Datenpublikation eine Kompilation von verschiedenen Datensätzen, welche auch von Textpublikationen abgeleitet wurde.	Es ist darauf zu achten, dass die Rechte von den Erstellern der verwendeten Fremddaten nicht verletzt werden.
IsVariantFormOf	Werden Datensätze kompiliert und neu strukturiert, so kann dieser Beziehungstyp verwendet werden.	

³² https://de.wikipedia.org/wiki/Uniform_Resource_Name (abgerufen am 05.12.2019).

1.4.17 Wird die Datenpublikation überprüft (reviewed)?

Derzeit werden die Datenpublikationen im Ermessen der Editoren auf technische bzw. inhaltliche Vollständigkeit überprüft. Bei Bedarf kann der Editor einen Fachexperten heranziehen, der die geplante Datenpublikation auf inhaltliche Vollständigkeit überprüft.

Vorerst ist nur ein „Review on demand“ vorgesehen. Bei der Abnahme der Datenpublikation überprüft der Editor die Datenpublikation auf formale Kriterien, wie die vollständig bereitgestellten Metadaten, die Datenerläuterung auf die Nachvollziehbarkeit und Konsistenz mit den Datensätzen sowie auf die Lesbar- und Verständlichkeit der Datensätze selber. Im Zweifelsfall wird zur Überprüfung eine Fachexpertise eingeholt. Es wird bereits an einem Konzept zur Implementierung eines umfangreichen Reviewprozesses gearbeitet und soll bald umgesetzt werden.

1.4.18 Wie können die Daten vor dem finalen Upload an den Editor übergeben werden?

Die Daten können über die gängigsten Datenspeichersysteme und via E-Mail übergeben werden.

Kleinere Datenmengen können via E-Mail versendet werden. Größere Datensätze sollen über ein Filesharing-System bzw. interne Verzeichnisse bereitgestellt werden.

1.4.19 Werden die Entwurfspublikationen auch abgespeichert?

Ja, die Entwürfe und die dazugehörigen Korrespondenzen werden in einem eigenen Verzeichnis abgelegt.

Bevor der Ersteller die zu publizierenden Daten in das Tethys RDR hochlädt, findet schon ein pre-Review-Prozess von den Tethys RDR-Editoren statt. Die dabei entstehenden Publikationsentwürfe werden in einem eigenen geschlossenen Tethys-Verzeichnis mit eingeschränktem Zugang gespeichert bzw. archiviert. Damit soll eine Nachvollziehbarkeit der Publikationsvorbereitung gewährleistet werden und den Publikationsprozess nach dem Upload ins Tethys RDR beschleunigen.

Nachdem die publikationsreifen Daten und Metadaten in das Tethys RDR hochgeladen und als Publikation freigegeben wurden, werden diese in das Digital Asset Management System gespeichert und können dort weder manipuliert noch ausgetauscht werden. Alle vom Editor abgelehnten Daten innerhalb des Upload-Prozesses werden nicht archiviert.

1.4.20 Wie handhaben die Editoren den Reviewprozess?

Die Editoren schreiben Anmerkungen und weisen auf Fehler oder Unklarheiten hin. Sie manipulieren die Datenfiles selber nicht. Die Korrekturen müssen von dem Einreichenden bzw. den Erstellern selber durchgeführt werden.

Die Aufgabe der kontrollierenden Editoren ist das Überprüfen der Daten und Datenbeschreibung auf Nachvollziehbarkeit, Vollständigkeit und Plausibilität. Werden Datenbeschreibungen mitgeliefert, so werden die Korrekturen und Anmerkungen direkt darin vermerkt. Gegebenenfalls wird ein eigenes Korrekturblatt verfasst. Die Daten selber werden niemals von den Editoren in irgendeiner Form manipuliert oder verändert. Damit soll eine unbeabsichtigte Datenveränderung durch Dritte vermieden werden.

1.4.21 Ich möchte verschiedene Datensatztypen gemeinsam publizieren. Ist das möglich?

Sofern die Datensätze zueinander passen und es eine vollwertige Publikation ergibt, ist dies möglich. Als Datensatztyp ist „Mixed Type“ zu wählen. Im Zweifelsfall wird dazu geraten, diese getrennt voneinander zu publizieren und durch das Setzen einer Relation (1.4.16) aufeinander zu verweisen.

Datensatztypen: Dieses Metadatenelement dient zur Klassifizierung des Datensatztypus und soll im Speziellen bei einem angewachsenen Datenbestand das Filtern nach Datensatztypen maßgeblich erleichtern. Ebenfalls kann mit diesem Metadatenelement gezielter nach den einzelnen Datensatztypen gesucht werden.

Um den Überblick einer Datenpublikation zu bewahren und eine leichte Auffindbarkeit zu gewährleisten, ist es sinnvoll, die Datenpublikation nicht zu umfangreich zu gestalten. Hier ist abzuwegen, ob die Datensätze eventuell in kleinere Einheiten gesplittet werden können. Diese können auch untereinander verlinkt bzw. in Relation gesetzt werden.

Ist es sinnvoll, verschiedene Datensatztypen in einer Datenpublikation zu publizieren, so muss als Datensatztyp „Mixed Type“ ausgewählt werden.

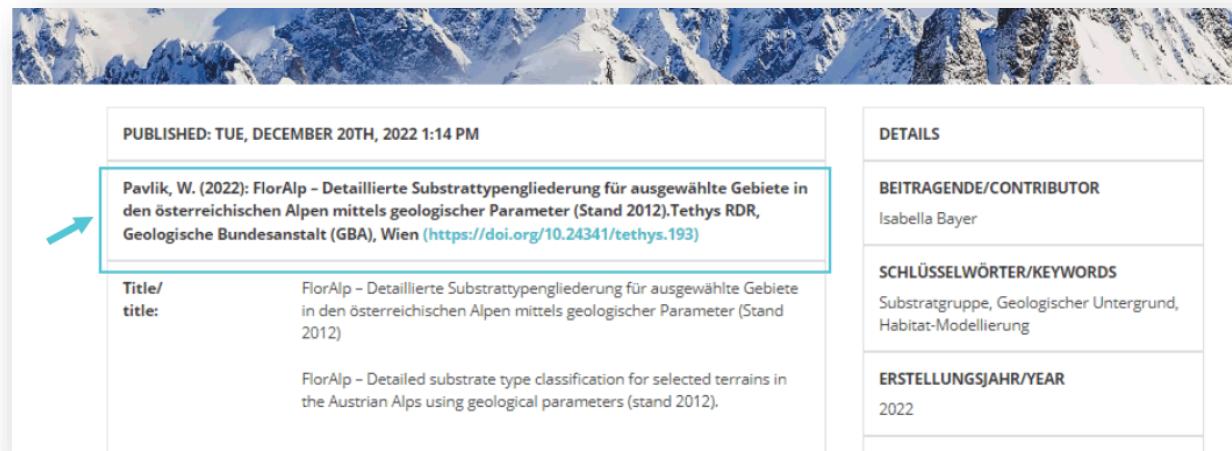
1.4.22 Wie kann ich die Datenpublikation zitieren bzw. darauf verweisen?

Die Datenpublikation kann über den „persistent identifier“, im Falle des Tethys RDRs über die DOI – eindeutig identifiziert und aufgerufen werden.

Tethys verwendet die DOI als persistente ID, die auch gleichzeitig die URI der betreffenden Datenpublikation ist. Die DOI ist einmalig und kann weder ausgetauscht noch verändert werden. Bei bekannter DOI kann man damit direkt die Landingpage der betroffenen Datenpublikation aufrufen bzw. darauf referenzieren.

Bei den einzelnen Datenpublikationen im Tethys RDR wird ein Zitervorschlag bereitgestellt (Abb. 6).

Abbildung 6: Beispiel eines Zitervorschlages (<https://doi.org/10.24341/tethys.193>).



1.5 Checkliste

Wer mit der Datenpublikation vertraut ist, kann die Fragen als Checkliste verwenden und benötigt nur vereinzelt die Unterstützung der ausführlichen Richtlinien für das Publizieren von Forschungsdaten.

- Eine nachträgliche Veränderung oder Löschung der publizierten Daten ist nicht vorgesehen. ([→ 1.4.4, 1.4.5](#))
- Es handelt sich um Forschungsdaten erstellt von bzw. in Kooperation mit der GeoSphere Austria. ([→ 1.4.2, 1.4.6](#))
- Die zu publizierenden Daten sind Tethys RDR konforme fachrelevante Forschungsdaten. ([→ 1.4.2](#))
- Daten sind wissenschaftlich korrekt strukturiert, konsistent und vollständig. ([→ 1.4.2, 1.4.3](#))
- Zusätzlich zu den Datensätzen wird eine selbsterklärende Beschreibung der verwendeten Parameter und die korrekte Nutzung bereitgestellt. ([→ 1.4.3](#))
- Alle beteiligten Ersteller wurden über die geplante Datenpublikation in Kenntnis gesetzt und es wurde das Einverständnis dazu eingeholt. ([→ 1.4.7](#))
- Es ist geklärt, welche Nutzungs Lizenz (Creative Commons) für die Datenpublikation vergeben wird. ([→ 1.4.8](#))
- Es sind gegebenenfalls die Rechte der verwendeten Fremddaten geklärt. ([→ 1.4.9](#))
- Die Datensätze liegen in offenen und nachhaltigen Formaten vor, die von Tethys RDR unterstützt werden. ([→ 1.4.10](#))
- Die Datenpublikation findet im Rahmen eines Projektes statt. Die Information dazu wurde bereits dem Editor gemeldet. ([→ 1.4.12](#))
- Es ist ein Embargo notwendig und wird bei der Publikation mitberücksichtigt. ([→ 1.4.13](#))
- Der Abstract wie auch der Titel der Datenpublikation liegt sowohl auf Englisch als auch auf Deutsch vor. ([→ 1.4.11](#))
- Gesamtgröße der hochgeladenen Datenfiles < 512 MB. ([→ 1.4.15](#))

2. Terms and Conditions

2.1 Webauftritt

Die Domain-Inhaberin (die GeoSphere Austria), explizit aber die Autoren der Website Tethys RDR (www.tethys.at), übernehmen keine Gewähr für die Aktualität, Richtigkeit und Vollständigkeit der bereitgestellten Informationen. Haftungsansprüche gegen die Autoren, die sich auf Schäden materieller oder ideeller Art beziehen, die durch die Nutzung oder Nichtnutzung der dargebotenen Informationen bzw. durch die Nutzung fehlerhafter und unvollständiger Informationen verursacht wurden, werden ausdrücklich ausgeschlossen, soweit nicht Vorsatz oder grobe Fahrlässigkeit der Autoren vorliegt. Die Autoren behalten es sich ausdrücklich vor, Teile der Seiten oder das gesamte Angebot ohne gesonderte Ankündigung zu verändern, zu ergänzen, zu löschen oder die Veröffentlichung zeitweise oder endgültig einzustellen.

2.1.1 Verweise/Links

Bei direkten oder indirekten Verweisen auf fremde Internetseiten (Links), die außerhalb des Verantwortungsbereichs der Autoren liegen, wird eine Haftung nur dann übernommen, wenn die Autoren von den Inhalten Kenntnis haben, dies zumutbar ist und diese über die technischen Mittel verfügen, um deren Verwendung bei rechtswidrigen Inhalten zu verhindern. Die Autoren erklären hiermit ausdrücklich, dass zum Zeitpunkt der Linksetzung keine illegalen Inhalte auf den zu verlinkenden Seiten erkennbar waren. Die Autoren haben keinerlei Einfluss auf die aktuelle und zukünftige Gestaltung und auf die Inhalte der verknüpften Seiten. Deshalb distanzieren sie sich hiermit ausdrücklich von allen Inhalten aller gelinkten/verknüpften Seiten, die nach der Linksetzung verändert wurden. Diese Feststellung gilt auch für alle innerhalb des eigenen Internetangebotes gesetzten Links und Verweise. Für illegale, fehlerhafte oder unvollständige Inhalte und insbesondere für Schäden, die aus der Nutzung oder Nichtnutzung solcherart dargebotener Informationen entstehen, haftet allein der Anbieter der Seite, auf die verwiesen wurde, nicht derjenige, der via Link auf die betreffende Seite verwiesen wurde.

2.1.2 Urheber- und Kennzeichenrecht

Die Autoren sind bestrebt, in allen Publikationen die Urheberrechte aller verwendeten Grafiken und Texte zu beachten, von ihnen selbst erstellte Grafiken und Texte zu nutzen oder auf lizenfreie Grafiken und Texte zurückzugreifen. Alle innerhalb des Internetangebotes genannten und von Dritten registrierten Marken- und Warenzeichen unterliegen uneingeschränkt den Bestimmungen des jeweils gültigen Kennzeichenrechts und den Besitzrechten der jeweiligen eingetragenen Eigentümer. Allein aufgrund der bloßen Nennung ist nicht der Schluss zu ziehen, dass Markenzeichen nicht den Rechten Dritter unterliegen. Sofern nicht anders angegeben, wird der Inhalt dieser Website unter einer Creative Commons Attribution 3.0-Lizenz lizenziert.

2.1.3 Datenschutz-Bestimmungen

Datenschutz ist der GeoSphere Austria ein wichtiges Anliegen. Mit dieser Datenschutzerklärung möchte unsere gastgebende Einrichtung, die GeoSphere Austria, die Öffentlichkeit über Art, Umfang und Zweck der von uns erhobenen, verwendeten und verarbeiteten personenbezogenen Daten informieren. Darüber hinaus werden betroffene Personen mittels dieser Datenschutzerklärung über die ihnen zustehenden Rechte informiert.

Die Nutzung der Internetseiten von Tethys RDR ist ohne Angabe personenbezogener Daten möglich. Wenn eine betroffene Person jedoch unsere Dienste wie das Übermitteln von Datensätzen nutzen oder mit uns in Kontakt treten möchte, ist die Verarbeitung personenbezogener Daten erforderlich. Wenn die Verarbeitung personenbezogener Daten notwendig ist und keine gesetzliche Grundlage für eine solche Verarbeitung besteht, holen wir in der Regel die Zustimmung der betroffenen Person ein. Wir erfassen Informationen von Ihnen, wenn Sie sich auf unserer Website anmelden.

Die Verarbeitung personenbezogener Daten von datenpublizierenden Erstellern, deren Co-Erststeller sowie Beitragende, wie Name, E-Mail-Adresse und optional die ORCID ID erfolgt stets durch die datenpublizierenden Ersteller im Einklang mit der Allgemeinen Datenschutzgrundverordnung (DSGVO) nach Artikel 6, Absatz 1b. Diese Informationen werden zum Zwecke der Anmeldung der datenpublizierenden Ersteller und zum Zwecke der Umsetzung der guten wissenschaftlichen Praxis gesammelt. Damit einhergeht, dass nach erfolgreichem Publizieren der Daten keine Löschung oder Veränderung der Daten und Metadaten inklusive aller personenbezogenen Daten möglich ist.

Als für die Verarbeitung Verantwortlicher hat Tethys RDR zahlreiche technische und organisatorische Maßnahmen getroffen, um den bestmöglichen Schutz der über diese Website verarbeiteten personenbezogenen Daten sicherzustellen. Internetbasierte Datenübertragungen können jedoch grundsätzlich Sicherheitslücken aufweisen, sodass ein absoluter Schutz möglicherweise nicht gewährleistet ist.

Sie haben grundsätzlich das Recht auf Auskunft, Berichtigung, Löschung, Einschränkung, Datenübertragbarkeit und Widerspruch.

Verantwortlich für die Verarbeitung personenbezogener Daten im Sinne der DSGVO ist die Generaldirektion der GeoSphere Austria:

Kontaktdaten

GeoSphere Austria
Hohe Warte 38, 1190 Wien
Telefonnummer: +43 1 36026
Fax: +43 1 3691233

Datenschutzbeauftragter

CoreTEC IT Security Solutions GmbH
Ernst-Melchior-Gasse 24/DG, 1020 Wien
E-Mail: datenschutz@geosphere.at
Fax: +43 1 3691233

Wenn Sie glauben, dass die Verarbeitung Ihrer Daten gegen das Datenschutzrecht verstößt oder Ihre datenschutzrechtlichen Ansprüche sonst in einer Weise verletzt worden sind, können Sie bei der dafür zuständigen Aufsichtsbehörde eine Beschwerde einlegen:

Österreichische Datenschutzbehörde
Barichgasse 40–42, 1030 Wien
Telefon: +43 1 52 152-0
E-Mail: dsb@dsb.gv.at
Website: <http://www.dsb.gv.at/>

2.1.4 Erfassung allgemeiner Daten und Informationen

Die Website von Tethys RDR sammelt eine Reihe von allgemeinen Daten und Informationen, wenn eine betroffene Person oder ein automatisiertes System die Website aufruft. Diese allgemeinen Daten und Informationen werden in den Server-Protokolldateien gespeichert. Gesammelt werden (1) die verwendeten Browsetypen und -versionen, (2) das vom zugreifenden System verwendete Betriebssystem, (3) die Website, von der aus ein zugreifendes System auf unsere Website gelangt (sogenannte Referrer), (4) die Sub-Sites, (5) Datum und Uhrzeit des Zugriffs auf die Internetseite, (6) eine Internetprotokolladresse (IP-Adresse), (7) der Internetdienstanbieter des zugreifenden Systems und (8) alle anderen ähnlichen Daten und Informationen, die im Falle von Angriffen auf unsere Informationstechnologiesysteme verwendet werden können.

Bei Verwendung dieser allgemeinen Daten und Informationen kann Tethys RDR keine Rückschlüsse auf die betroffene Person ziehen. Diese Informationen werden vielmehr benötigt, um (1) den Inhalt unserer Website korrekt bereitzustellen, (2) den Inhalt unserer Website sowie deren Werbung zu optimieren, (3) die langfristige Überlebensfähigkeit unserer Informationstechnologiesysteme und der Website-Technologie sicherzustellen und (4) den Strafverfolgungsbehörden, die für die strafrechtliche Verfolgung im Falle eines Cyberangriffs erforderlichen Informationen zur Verfügung stellen. Daher analysiert Tethys RDR anonym erhobene Daten und Informationen statistisch mit dem Ziel, den Datenschutz und die Datensicherheit unserer Institution zu erhöhen und ein optimales Schutzniveau für die von uns verarbeiteten personenbezogenen Daten zu gewährleisten. Die anonymen Daten der Server-Logfiles werden getrennt von allen personenbezogenen Daten einer betroffenen Person gespeichert.

2.1.5 Anmeldung/Registrierung auf unserer Website

Die betroffene Person hat die Möglichkeit, sich auf der Website des für die Verarbeitung Verantwortlichen unter Angabe personenbezogener Daten anzumelden (zu registrieren). Welche personenbezogenen Daten an die Steuerung übermittelt werden, bestimmt die jeweilige Eingabemaske des Anmeldeformulars. Die von der betroffenen Person eingegebenen personenbezogenen Daten werden ausschließlich für den internen Gebrauch durch den für die Verarbeitung Verantwortlichen und für eigene Zwecke erhoben und gespeichert. Der Controller kann die Übertragung an einen oder mehrere Prozessoren (z.B. einen wissenschaftlichen Herausgeber) anfordern, die personenbezogene Daten auch für einen internen Zweck verwenden, der dem Controller zuzuordnen ist.

Durch die Registrierung auf der Website des Controllers werden auch die vom Internet Service Provider (ISP) zugewiesene und vom Betroffenen verwendete IP-Adresse – Datum und Uhrzeit der Registrierung – gespeichert. Die Speicherung dieser Daten erfolgt vor dem Hintergrund, dass nur so ein Missbrauch unserer Dienste verhindert und gegebenenfalls eine Aufklärung der begangenen Verstöße ermöglicht wird. Insofern ist die Speicherung dieser Daten erforderlich, um die Steuerung abzusichern. Diese Daten werden ohne Ihre ausdrückliche Zustimmung nicht an Dritte weitergegeben, es sei denn, es besteht eine gesetzliche Verpflichtung zur Weitergabe der Daten oder die Übermittlung dient der strafrechtlichen Verfolgung.

Die Registrierung der betroffenen Person mit der freiwilligen Angabe personenbezogener Daten soll es dem für die Verarbeitung Verantwortlichen ermöglichen, die betroffenen Inhalte oder Dienste anzubieten, die aufgrund der Art der betreffenden Angelegenheit nur registrierten Nutzern angeboten werden dürfen. Registrierte Personen können die bei der Registrierung angegebenen personenbezogenen Daten jederzeit ändern oder vollständig aus dem Datenbestand des Verantwortlichen löschen lassen.

Der für die Verarbeitung Verantwortliche teilt jeder betroffenen Person auf Anfrage jederzeit mit, welche personenbezogenen Daten über die betroffene Person gespeichert sind. Darüber hinaus berichtigt oder löscht der für die Datenverarbeitung Verantwortliche personenbezogene Daten auf Verlangen oder unter Angabe der betroffenen Person, sofern keine gesetzlichen Aufbewahrungspflichten bestehen. Ein in dieser Datenschutzerklärung ausdrücklich benannter Datenschutzbeauftragter sowie die gesamten Mitarbeiter des für die Verarbeitung Verantwortlichen stehen der betroffenen Person als Ansprechpartner zur Verfügung.

2.1.6 Kontaktmöglichkeit über die Website

Die Website von Tethys RDR enthält Informationen, die einen schnellen elektronischen Kontakt zu unserer Einrichtung sowie eine direkte Kommunikation mit uns ermöglichen. Dazu gehört auch eine allgemeine Adresse der sogenannten elektronischen Post (E-Mail-Adresse). Wenn eine betroffene Person den für die Verarbeitung Verantwortlichen per E-Mail über das Kontaktformular kontaktiert, werden die von der betroffenen Person übermittelten personenbezogenen Daten automatisch gespeichert. Diese von einer betroffenen Person freiwillig übermittelten personenbezogenen Daten werden zum Zwecke der Verarbeitung oder Kontaktaufnahme mit der betroffenen Person gespeichert. Bei Nutzung des Kontaktformulars erfolgt keine Weitergabe dieser personenbezogenen Daten an Dritte.

2.1.7 Übermittlung wissenschaftlicher Daten über die Website

Die Website von Tethys RDR enthält eine Webanwendung zum Einreichen von wissenschaftlichen Datensätzen, die in der Tethys RDR-Datenbank gespeichert werden, um sie zu veröffentlichen. Wenn eine betroffene Person mit dem Antrag auf Einreichung wissenschaftlicher Daten Kontakt zum für die Verarbeitung Verantwortlichen aufnimmt, werden die von der betroffenen Person übermittelten personenbezogenen Daten automatisch gespeichert.

Um wissenschaftliche Datensätze erfolgreich bei Tethys RDR einzureichen, muss die betroffene Person einige zusätzliche personenbezogene Daten (z.B. Name des Einreichers, des Erstellers und Beitragende) bereitstellen, damit diese Datensätze in der wissenschaftlichen Gemeinschaft korrekt zitiert werden können. Diese Informationen werden der Öffentlichkeit nach dem Einreichungsprozess über Datensatz-Metadaten (im XML- oder JSON-Format und über die Tethys RDR-Website) zur Verfügung gestellt. Dies ist eine Voraussetzung für das wissenschaftliche Publizieren. Veröffentlichungen zu wissenschaftlichen Daten, einschließlich der oben genannten personenbezogenen Daten, können von Dritten (z.B. Bibliotheken, Datenportalen) unter Verwendung der Metadaten und Datendienste von Tethys RDR verwendet werden.

2.1.8 Nutzung von Content Delivery Networks (CDN)

Auf dieser Website hat der Controller Javascript, Schriftarten und Bilder integriert, die von Content Delivery Networks bereitgestellt werden. Ein Content Delivery Network (CDN) ist ein geografisch verteiltes Netzwerk von Proxy-Servern und deren Rechenzentren. Ziel ist es, den Service räumlich auf die Endbenutzer zu verteilen, um eine hohe Verfügbarkeit und Leistung zu gewährleisten. Während dieses technischen Verfahrens können Dritte Kenntnis von personenbezogenen Daten, wie der IP-Adresse der betroffenen Person, erlangen.

Die bei Tethys verwendeten CDN-Dienste wurden vom Controller auf DSGVO-Konformität geprüft.

2.1.9 Routinemäßige Löschung und Sperrung personenbezogener Daten

Der für die Verarbeitung Verantwortliche verarbeitet und speichert die personenbezogenen Daten der betroffenen Person nur für den Zeitraum, der zur Erreichung des Zwecks der Speicherung erforderlich ist, oder, soweit dies vom europäischen Gesetzgeber oder anderen Gesetzgebern in Gesetzen oder Verordnungen, denen der für die Verarbeitung Verantwortliche unterliegt, gewährt wird.

Ist der Speicherzweck nicht anwendbar oder läuft eine vom europäischen Gesetzgeber oder einem anderen zuständigen Gesetzgeber festgelegte Speicherfrist ab, werden die personenbezogenen Daten gemäß den gesetzlichen Bestimmungen routinemäßig gesperrt oder gelöscht.

Bitte beachten Sie auch unsere rechtlichen Hinweise zu Nutzung, Haftungsausschluss und Haftungsbeschränkungen für die Nutzung unserer Website.

2.1.10 OpenStreetMap

Diese Seite nutzt über eine Programmierschnittstelle (Application Programming Interface, API) das Open Source Mapping-Werkzeug „OpenStreetMap“ (OSM). Anbieter ist die OpenStreetMap Foundation. Zur Nutzung der Funktionen von OpenStreetMap ist es notwendig, Ihre IP-Adresse zu speichern. Diese Informationen werden in der Regel an einen Server von OpenStreetMap übertragen und dort gespeichert. Der Anbieter dieser Seite hat keinen Einfluss auf diese Datenübertragung. Die Nutzung von OpenStreetMap erfolgt im Interesse einer ansprechenden Darstellung unserer Online-Angebote und an einer leichten Auffindbarkeit der von uns auf der Website angegebenen Orte. Dies stellt ein berechtigtes Interesse im Sinne von Artikel 6, Absatz 1 lit. f DSGVO dar. Mehr Informationen zum Umgang mit Nutzerdaten finden Sie auf der Datenschutzseite von OpenStreetMap und auf http://wiki.openstreetmap.org/wiki/Legal_FAQ.

2.1.11 Rechtswirksamkeit dieses Haftungsausschlusses

Dieser Haftungsausschluss ist als Teil des Internetangebotes zu betrachten, von dem aus auf diese Seite verwiesen wurde. Sollten Teile des Textes oder eines Wortlautes nicht, nicht vollständig oder nicht mehr dem geltenden Recht entsprechen, so wird hierdurch die Gültigkeit oder der Inhalt der übrigen Teile des Dokumentes nicht berührt.

3. Technische Anleitung zur Datenpublikation

3.1 Zeichenerklärung



Ein- und Ausblenden eines Hilfetextes.



Die mit Nummern gekennzeichneten Begriffe werden im Text näher erläutert.



Löschen, Papierkorb.

3.2 Eingabehilfe

Die vorliegende Hilfe erklärt die schrittweise Eingabe bzw. das Hochladen von Metadaten einer Datenpublikation in das Tethys RDR. Auf die Erklärung der einzelnen Metadatenelemente wird in dieser Anleitung nicht eingegangen. Diese sind in der Datenpublikationsrichtlinie (DPR) erklärt. Es wird auch empfohlen, die DPR zu lesen, bevor der Publikationsprozess gestartet wird.

3.2.1 Tethys Start

Über www.tethys.at wird die Startseite des Tethys RDR erreicht (Abb. 7).



Abbildung 7: Startseite des Tethys RDR.

Die Eingabe der im Tethys RDR zu publizierenden Daten beginnt mit dem **LOGIN** (1) auf der Startseite rechts oben (Abb. 7). Der Account für das Tethys RDR muss bei dem für das RDR zuständigen Administrator beantragt werden. Derzeit dafür zuständig ist Viktoria Haider (repository@geosphere.at).

3.2.2 Login

Für das Login (Abb. 8) werden **E-Mail-Adresse** (1) und **Password** (2) benötigt. Danach auf das Icon **Login** klicken. Durch Anhaken von **Remember me** werden die Anmeldedaten bei Bedarf gespeichert, sodass sie beim nächsten Login nicht mehr eingegeben werden müssen.

The screenshot shows the Tethys Login Form. At the top, there is a green header bar with the TETHYS logo and links for Home, Intro, Search, Services, and Help. To the right of the header are 'Login' and 'OAI' buttons. The main page title is 'Tethys Login Form'. Below the title, a note says: 'Login is only required for access to data under embargo or for submitting new datasets.' There are two input fields: one for 'Email' containing 'johannes.reischer@geologie.ac.at' (marked with a red circle and number 1), and one for 'Password' containing '*****' (marked with a red circle and number 2). Below these fields is a 'Remember me' checkbox. At the bottom is a large blue 'LOGIN' button.

Abbildung 8: Einstiegsfenster in den Publikationsbereich.

3.2.3 Publish new dataset (Datenpublikation veröffentlichen)

Um eine Datenpublikation publizieren zu können, muss im Menü links unter der Überschrift **Publish „+ Create“** (1) angeklickt werden (Abb. 9). Direkt im Dashboard (2) steht das Handbuch (Datenpublikationsrichtlinien, Terms & Conditions, Anleitung) und eine Anleitung zu GeoPackage (Hilfe zu GeoPackage) zum Download bereit.

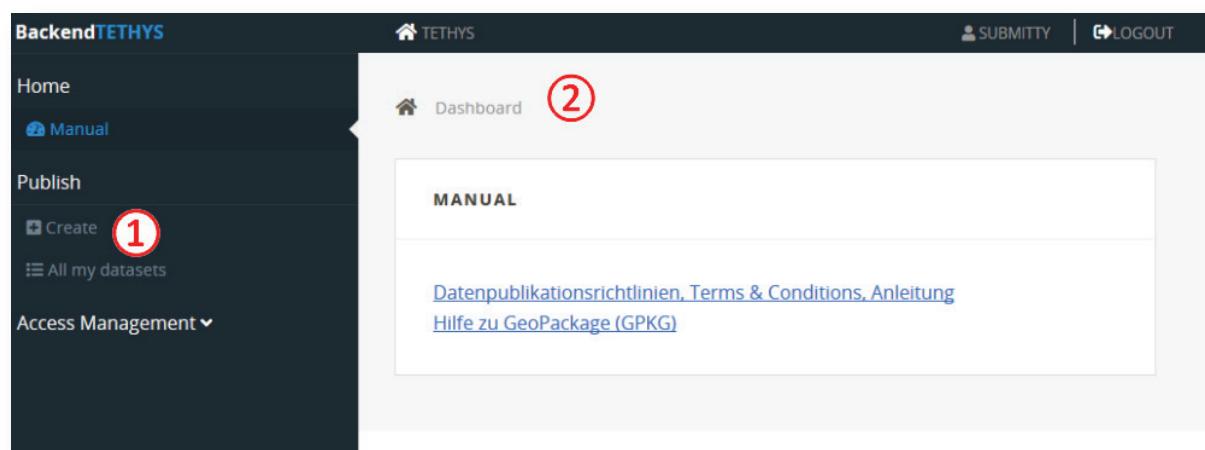


Abbildung 9: Erscheinungsbild nach dem Login.

3.2.3.1 Dataset Language (Sprache der Datenpublikation)

Bei **Language** ① muss die Hauptsprache der Datenpublikation ausgewählt werden (Abb. 10). Zur Auswahl stehen „de“ (Deutsch) und „en“ (Englisch). Danach auf -> **Continue** (Weiter) klicken. Falls keine Sprache ausgewählt wird, erscheint die Fehlermeldung „The language field is required“.

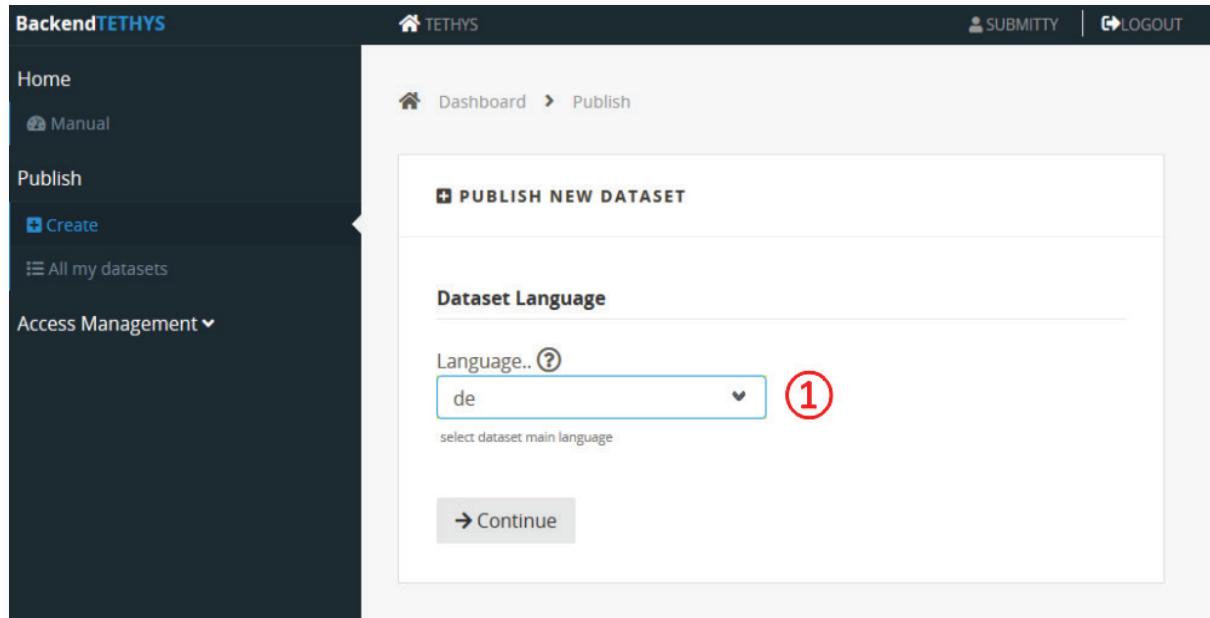


Abbildung 10: Auswahl der Hauptsprache.

3.2.3.2 Mandatory Elements (Verpflichtende Angaben)

3.2.3.2.1 General (Allgemein)

Bei **Dataset Type** (Typ der Datenpublikation) ① muss aus einer Liste von vorgegebenen Typen (Kategorien) ausgewählt werden (Abb. 11). Die genaue Beschreibung der Dataset Types kann in der Datenpublikationsrichtlinie nachgelesen werden.

Die Auswahl eines Dataset Types ist verpflichtend.

3.2.3.2.2 Title (Titel)

Unter **Main Title** (Haupttitel) ② muss der Titel der Datenpublikation eingetragen werden. Die Eingabe verlangt mindestens vier Zeichen. Die **Title Language** ③ ist durch die Auswahl der Dataset Language von voriger Seite vorgegeben.

Die Eingabe eines Titels ist verpflichtend.

Mit **Add additional title(s)** (zusätzliche Titel hinzufügen) ④ können Untertitel, Alternative Titel und dergleichen hinzugefügt werden. Der Titel wird unter **Title** ⑤ eingetragen. Für die Art kann bei **Type** unter verschiedenen Kategorien gewählt werden. Die Sprache (**Language**) kann in diesem Fall auch frei gewählt werden. Beispielsweise kann beim zusätzlichen Titel ein Titel in der jeweiligen anderen Sprache eingegeben werden, bei **Type** wäre dann *Translated* auszuwählen. Es können mehrere zusätzliche Titel wie *Subtitel* oder *Alternativer Titel* ausgewählt werden. Sollen auch diese zweisprachig zur Verfügung gestellt werden, so ist darauf zu achten, dass bei **Type** in beiden der gleiche ausgewählt

wird und keinesfalls *Translated*. Mit können der zusätzliche Titel und alle Angaben dazu wieder gelöscht werden.

PUBLISH NEW DATASET

Step One: Mandatory Elements

General

Bitte wählen Sie einen Datensatztyp aus der Liste aus.

Dataset Type * ①
— select type —

Main Title ②
②

Title Language..
③ de

Add additional title(s) ④
+ ④

Title	Type	Language
[TITLE]	[titleType] ▾	[language] ▾

5 [TITLE] [titleType] [language]

Description

Main Abstract ⑦
⑦

Abstract Language..
⑥ de

Add additional descriptions(s) ⑧
+ ⑧

Description	Type	Language
[DESCRIPTION]	[descriptionType] Methods Series_Information Technical_Info Translated Other	[language] ▾

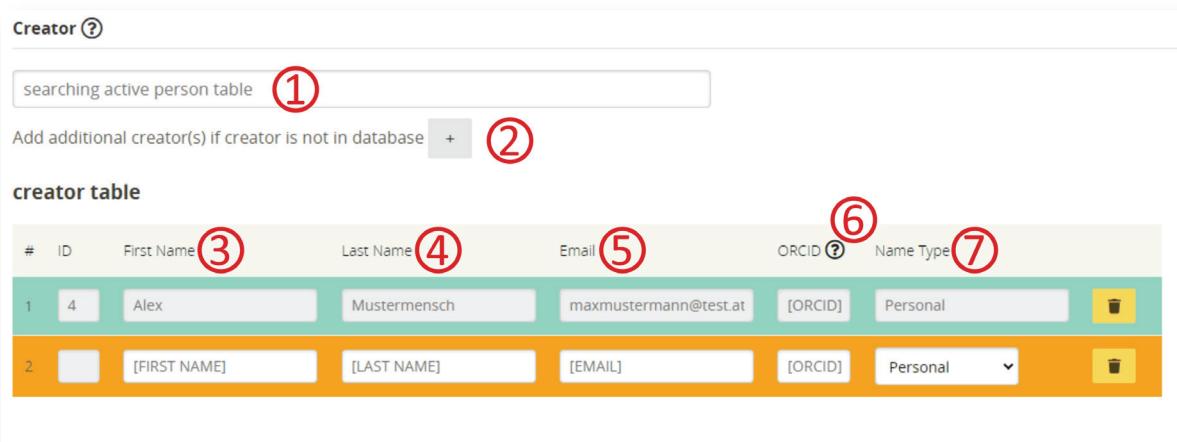
Abbildung 11: Verpflichtende Angaben zur Datenpublikation.

3.2.3.2.3 Description (Beschreibung)

Bei **Main Abstract** (Zusammenfassung) ⑦ wird verpflichtend erwartet, dass eine Zusammenfassung in der Hauptsprache ⑥ und mit **Add additional description(s)** (zusätzliche Beschreibung) ⑧ zusätzlich die übersetzte *Zusammenfassung (Translated)* in der alternativen Sprache verfasst wird. Optional können hier noch die *Methodenbeschreibung (Methods)*, Serien-Informationen (*Series_information*), *Technische Beschreibung (Technical_info)* und *Sonstiges (Other)* abgelegt werden. Werden die optionalen Beschreibungen zweisprachig zur Verfügung gestellt, so wird hier als Type nicht *Translated* verwendet, sondern der gleiche **Type** wie in der Hauptsprache. Die maximale Zeichenlänge beträgt jeweils 2.500 Zeichen.

3.2.3.2.4 Creator (Ersteller)

In Feld ① kann aus einer Liste von bereits vorhandenen Erstellern ausgewählt werden (Abb. 12). Sobald der passende Name ausgewählt wurde, scheint der Eintrag mit hellblauem Hintergrund in der Tabelle (**creator table**) unterhalb des Suchfeldes auf. Ein Autor kann auch eine Organisation sein. Um zusätzliche Ersteller einzutragen oder Ersteller, die noch nicht in der Datenbank abgelegt wurden, auf + ② klicken. Dann öffnet sich für den neuen anzugebenen Ersteller eine Zeile in der Tabelle (**creator table**), die orange hinterlegt ist. Für jeden weiteren zusätzlichen Ersteller muss abermals auf + geklickt werden. **First Name** (Vorname) ③, **Last Name** (Nachname) ④ und **E-Mail** ⑤ müssen ausgefüllt werden. Die **ORCID** ⑥ ist eine dauerhafte digitale Kennung für Ersteller (Forschende) und kann optional ausgefüllt werden. Unter **Name Type** ⑦ muss ausgewählt werden, ob es sich um eine Person (Personal) oder um eine Organisation (Organisational) handelt. Dieses Feld kann nur bei neu angelegten Erstellern befüllt oder geändert werden. Die ID wird vom System vergeben und kann nicht verändert werden. Der Eintrag zumindest eines Erstellers ist verpflichtend.



The screenshot shows the 'Creator' section of a form. At the top is a search bar containing 'searching active person table' with a red circle around it (1). Below the search bar is a button labeled 'Add additional creator(s) if creator is not in database' followed by a plus sign (+) and a red circle around it (2). The main area is titled 'creator table'. It contains a table with columns: #, ID, First Name (3), Last Name (4), Email (5), ORCID (6), and Name Type (7). There are two rows. Row 1 has an ID of 4, first name 'Alex', last name 'Mustermensch', email 'maxmustermann@test.at', an ORCID field, and 'Personal' selected in the Name Type dropdown. Row 2 is a new entry with a blank ID field, 'FIRST NAME' placeholder, 'LAST NAME' placeholder, 'EMAIL' placeholder, an empty ORCID field, and 'Personal' selected in the Name Type dropdown. Both rows have a yellow trash can icon at the end.

Abbildung 12: Weitere verpflichtende Angaben zu der Datenpublikation.

3.2.3.2.5 Contributor (Beitragender)

Hier ① sollen jene Personen oder Organisationen eingetragen werden (Abb. 13), die zur Erstellung der Dateninhalte beigetragen haben oder den Erstellern bei der Datengenerierung unterstützt haben. Um zusätzliche Mitwirkende oder Beitragende einzutragen, die noch nicht in der Datenbank erfasst sind, zunächst auf + ② klicken. Für jeden weiteren zusätzlichen Mitwirkenden muss erneut auf + geklickt werden.



The screenshot shows the 'Contributor' section of a form. At the top is a search bar containing 'searching active person table' with a red circle around it (1). Below the search bar is a button labeled 'Add additional contributor(s) if contributor is not in database' followed by a plus sign (+) and a red circle around it (2). The main area is titled 'contributor table'. It contains a table with columns: #, ID, First Name (3), Last Name (4), Email (5), ORCID (6), Name Type (7), and Type. There are two rows. Row 1 has an ID of 11, first name 'Jane', last name 'Public', email 'jane.public@mail.aw', an ORCID field, 'Personal' selected in the Name Type dropdown, and 'Other' selected in the Type dropdown. Row 2 is a new entry with a blank ID field, 'FIRST NAME' placeholder, 'LAST NAME' placeholder, 'EMAIL' placeholder, an empty ORCID field, and 'Personal' selected in the Name Type dropdown. Both rows have a yellow trash can icon at the end.

Abbildung 13: Metadaten zu den Mitwirkenden der Datenpublikation.

Das Eintragen der Daten von den Mitwirkenden verhält sich ähnlich wie bei den Erstellern. Zusätzlich wird **Type** abgefragt. Hier hat der Einreicher die Möglichkeit, aus einer Liste von verschiedenen Funktionen eine dem Ersteller zuzuordnen (Abb. 14). In der Regel wird hier „Other“ ausgewählt.

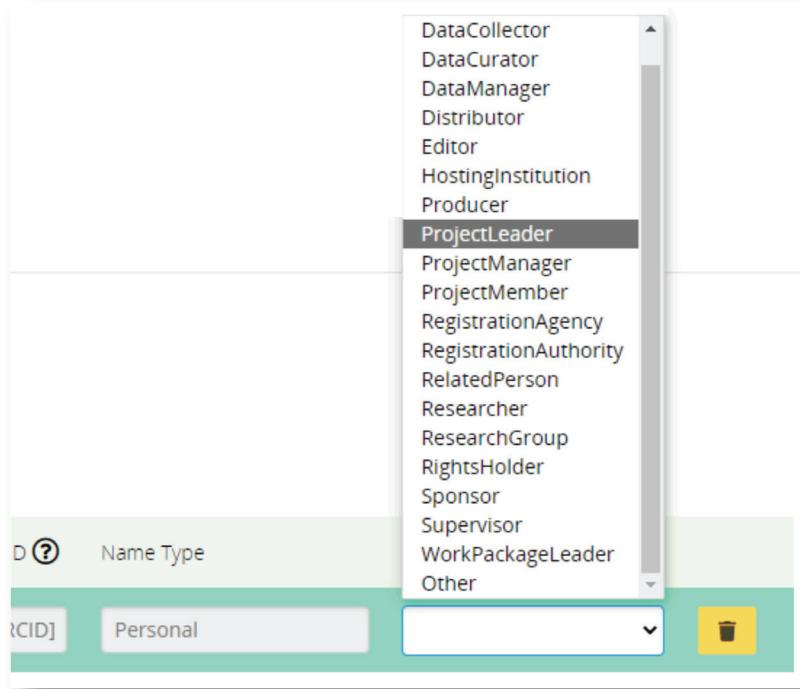


Abbildung 14: Auswahlfeld für das Metadatenelement „Type“.

3.2.3.2.6 Creating Corporation (Betreiber)

Als **Corporation Name** ① ist Tethys RDR voreingestellt. Hier ist keine eigene Eingabe möglich. Bevor der Publikationsprozess fortgesetzt wird, müssen die **terms and conditions** (Rechte) ② akzeptiert und dies durch das Setzen eines Häkchens bestätigt werden (Abb. 15).

Abbildung 14: Name des Betreibers und die Checkbox für terms and conditions.

Nachdem alle Felder ausgefüllt wurden auf -> **Continue** (Weiter) ③ klicken. Falls ein verpflichtendes Feld nicht ausgefüllt wurde, erscheint am Ende der Seite eine Fehlermeldung wie z.B. „**The Type field is required**“. Sind alle verpflichtenden Felder ausgefüllt geht es zur nächsten Seite.

3.2.3.3 Recommended Elements (*Empfohlene Angaben*)

3.2.3.3.1 Project

Bei **Project** (1) kann das Projekt ausgewählt werden, in dem oder für das die Datenpublikation erstellt wurde (Abb. 16). Der Projektname muss vorher an den Tethys RDR-Administrator kommuniziert werden und kann erst dann aus der Liste gewählt werden.

3.2.3.3.2 Embargo Date (Datum der Veröffentlichung)

Das **Embargo Date** (2) ist jenes Datum, an dem die Datenpublikation veröffentlicht wird. Bei Angabe eines Embargo Dates werden die Metadaten zum Lesen freigegeben, die mitpublizierten Datensätze und Dokumente sind allerdings bis zu diesem Datum für den Download gesperrt und können nicht heruntergeladen und eingesehen werden.

Das Datum muss in der Form tt.mm.jjjj eingegeben werden, z.B. 14.09.2023.

3.2.3.3.3 Geo Location (Lage)

In der Karte kann mit dem Werkzeug (4) ein Rechteck aufgezogen werden, das die Lage der Daten der Datenpublikation umgrenzt. Die Koordinaten werden als Geographische Koordinaten (WGS84) in Dezimalgrad in die Felder (6) unter der Karte vom System eingetragen. Mit dem Werkzeug Papierkorb (5) können das Rechteck und die Koordinaten wieder gelöscht werden. Durch erneutes Klicken auf das Werkzeug (4) wird ein neues Rechteck aufgezogen und die Koordinaten des alten Rechtecks werden überschrieben. Die Koordinaten können auch per Tastatur in die Felder eingegeben werden.

Step Two: Recommended Elements

Project

Project.. (1)

project is optional

Date(s)

Embargo Date.. (2)

tt.mm.jjjj

EmbargoDate is optional

Geo Location (3)



xmin: (6)

ymax:

ymin:

xmax:

Abbildung 15: Zweite Publikationsstufe zu den empfohlenen Angaben.

Der Eintrag der Lage ist verpflichtend!

Bei **validate coordinates** (Überprüfung der Koordinaten) geht es um die Überprüfung der Richtigkeit bezüglich der Koordinateneingabe. Hier wird überprüft, ob die Koordinaten auch ein Rechteck bilden. Auf das umgrenzte Gebiet wird gezoomt.

3.2.3.3.4 Coverage (Bereich)

Unter **Coverage** (Abb. 17) können Angaben zu Höhe ①, Tiefe ② und Zeit ③ gemacht werden. Höhe und Tiefe sind in Meter (m) anzugeben, das Datum mit Zeit in yyyy-MM-dd HH:mm:ss also z.B.: 2022-04-01 10:42:00.

Achtung: Die Datumsangabe ist hier anders als beim Embargo Date. Die meisten Browser bieten nach Klick in das Datumsfeld einen Kalender an, aus dem das Datum und die Zeit ausgewählt werden können.

Coverage ②

○ absolut elevation (m)
● elevation range (m) ①
○ no elevation

elevation min:
1000

elevation max:
3000

○ absolut depth (m)
○ depth range (m)
○ no depth

depth absolut:
-3000

○ absolut time (yyyy-MM-dd HH:mm:ss)
● time range (yyyy-MM-dd HH:mm:ss) ③
○ no time

time absolut:
2020-01-01 10:42:00

Abbildung 16: Ausdehnung der Datenpublikation.

3.2.3.3.5 Dataset References (Referenzen)

Dataset References ②

Add Reference ①

Value of the identifier ② [RELATED IDENTIFIER] ② Type ③ [relatedIdentifierType] ③ Relation ④ [relationType] ④ Label ⑤ []

Dataset Keywords ②

Add Keyword ⑥ [] Schlüsselwörter Empfehlung ⑦

Keyword Value ⑧ [KEYWORD VALUE] ⑧ Keyword Type ⑨ [uncontrolled] ⑨ Language de []

← Back → Continue ⑩

Abbildung 17: Möglichkeit zur Erfassung von Referenzen und die Angabe von Schlüsselworten.

Mit **Add Reference** (Referenz hinzufügen) ① wird eine Zeile geöffnet, in die ganz links der **Value of the identifier** (Identifikator) ② eingetragen werden kann (Abb. 18). Das ist ein dauerhafter und eindeutiger digitaler Identifikator wie z.B. eine DOI oder ISBN. Daneben wird der **related identifier Type** (Typ des Identifikators) ③ aus einer Liste ausgewählt. Im Feld **Label** ⑤ muss die Reference näher beschrieben werden. An dieser Stelle kann auch eine Bezeichnung vergeben werden.

Die genaue Beschreibung der Relation ④ kann in der Datenpublikationsrichtlinie nachgelesen werden.

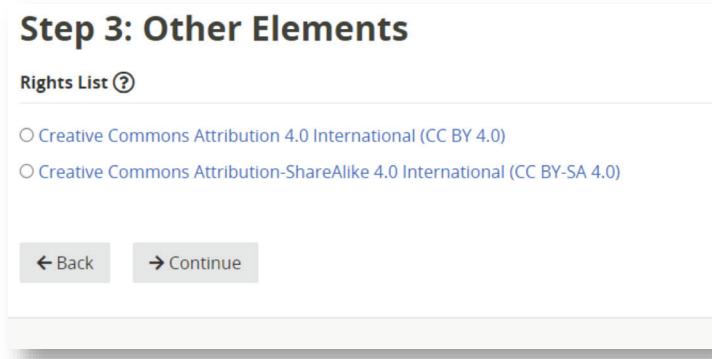
3.2.3.3.6 Dataset Keywords (Schlüsselwörter)

Es müssen insgesamt mindestens drei unterschiedliche Schlüsselwörter eingegeben werden. Idealerweise sind diese Schlüsselwörter nicht im Titel enthalten. Durch Klick auf **Add Keyword** ⑥ werden weitere Zeilen geöffnet. Es steht eine Schlüsselwörter-Empfehlung ⑦ von einem kontrollierten und standardisierten Vokabular in der jeweiligen Hauptsprache zur Verfügung. In das Feld ⑧ wird das Schlüsselwort eingegeben. **Keyword Type** und **Language** sind vorgegeben.

Nachdem alle Felder ausgefüllt wurden auf -> **Continue** (Weiter) ⑩ klicken. Falls ein verpflichtendes Feld nicht ausgefüllt wurde, erscheint am Ende der Seite eine Fehlermeldung.

3.2.3.4 Other Elements (andere Angaben)

Unter **Other Elements** (andere Angaben) gilt es nun, aus der **Rights List** eine passende Nutzungsbestimmung nach Creative Commons auszuwählen (Abb. 19). Bei Klick auf die jeweilige Nutzungsbestimmung wird die Seite <https://creativecommons.org> aufgerufen. Hier finden sich die Erläuterungen zu den jeweiligen Lizzenzen.



Step 3: Other Elements

Rights List

Creative Commons Attribution 4.0 International (CC BY 4.0)
 Creative Commons Attribution-ShareAlike 4.0 International (CC BY-SA 4.0)

Back Continue

Abbildung 18: Andere Angaben – Nutzungsbestimmungen.

3.2.3.5 File Upload (Datei hochladen)

Der Schritt **File Upload** (Datei hochladen) ist das eigentliche Herzstück der Applikation (Abb. 20). Nun können die Daten der Datenpublikation hochgeladen werden.

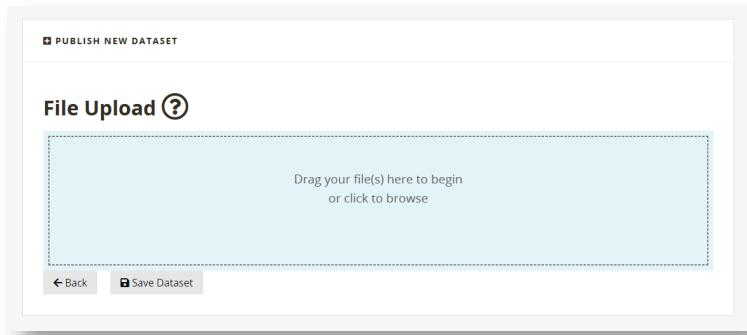


Abbildung 19: Hochladen der Datenpublikation 1.

In **File** (Datei) ① steht der Name der hochgeladenen Datei (Abb. 21). Dieser wird auch automatisch in **Label** (Text, Beschreibung) ② notiert. Der Text in **Label** kann aber verändert bzw. ergänzt werden.

Es besteht die Möglichkeit, mehrere Dateien gleichzeitig hochzuladen. Dabei können aber nur bestimmte Datentypen hochgeladen werden. In der Datenpublikationsrichtlinie können diese nachgelesen werden.

Mit **Save Dataset** ③ werden die Metadaten und die Daten im Tethys RDR gespeichert.

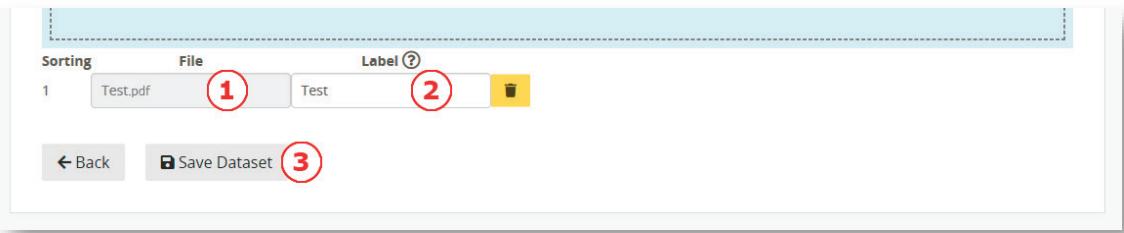


Abbildung 20: Hochladen der Datenpublikation 2.

3.2.3.6 Uploaded 1 file successfully (Hochladen erfolgreich)

Nach erfolgreichem Save bietet das Tethys RDR folgende drei Möglichkeiten (Abb. 22):

- [1] Die Metadaten können mit **Edit** (Bearbeiten) ① weiterbearbeitet werden. Die hochgeladene Datenpublikation kann allerdings **nicht** mehr gelöscht oder verändert werden.
- [2] Mit **Release** (Freigabe) ② werden die Metadaten und die Datenpublikation für den Editor freigegeben.
- [3] Mit **Delete** (Löschen) ③ werden der Uploadprozess gestoppt und die Inhalte gelöscht.

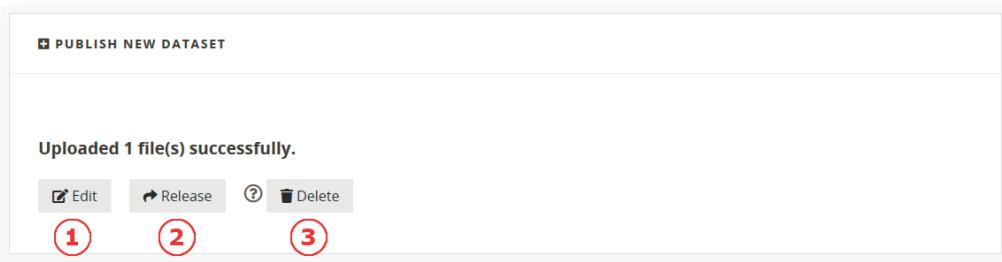
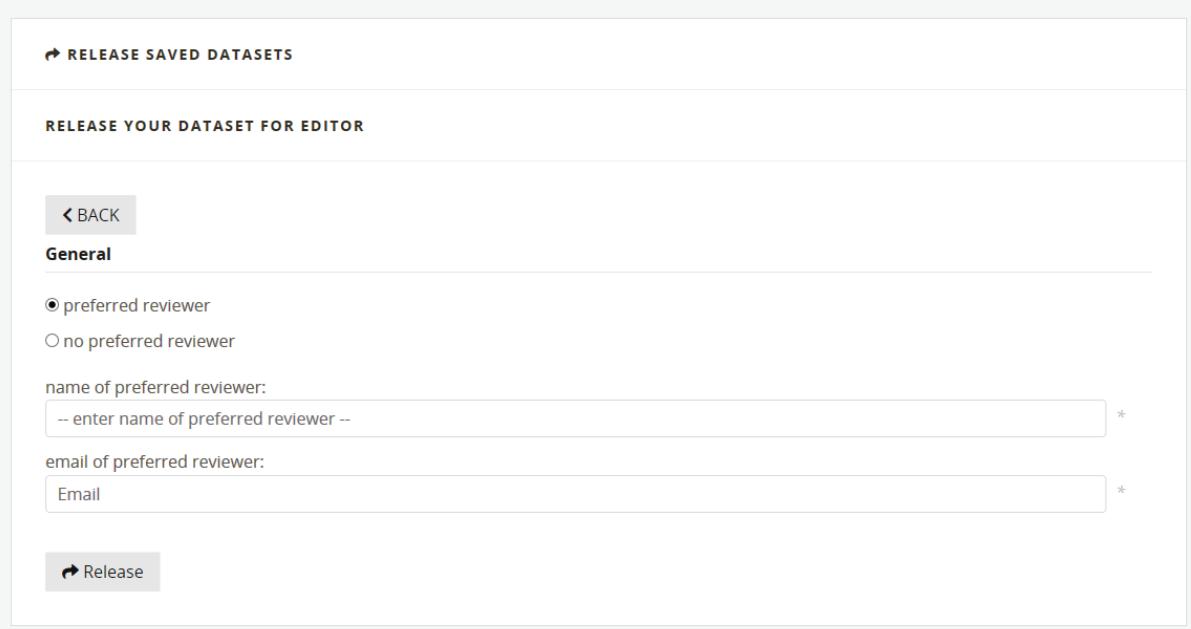


Abbildung 21: Edit, Release, Delete.

3.2.4 Release your dataset for editor (Freigabe der Datenpublikation für den Editor)

Nach Klick auf **Release** besteht die theoretische Möglichkeit, einen bevorzugten **Reviewer** (Begutachter) auszuwählen (Abb. 23). Zum derzeitigen Stand ist das Benennen eines Reviewers obsolet, da der Reviewprozess noch nicht vollständig in das Tethys RDR implementiert wurde. Daher hat es noch keine Auswirkung. Es empfiehlt sich hier **no preferred reviewer** auszuwählen. Anschließend **Release** anklicken.



RELEASE SAVED DATASETS

RELEASE YOUR DATASET FOR EDITOR

< BACK

General

preferred reviewer
 no preferred reviewer

name of preferred reviewer:
-- enter name of preferred reviewer -- *

email of preferred reviewer:
Email *

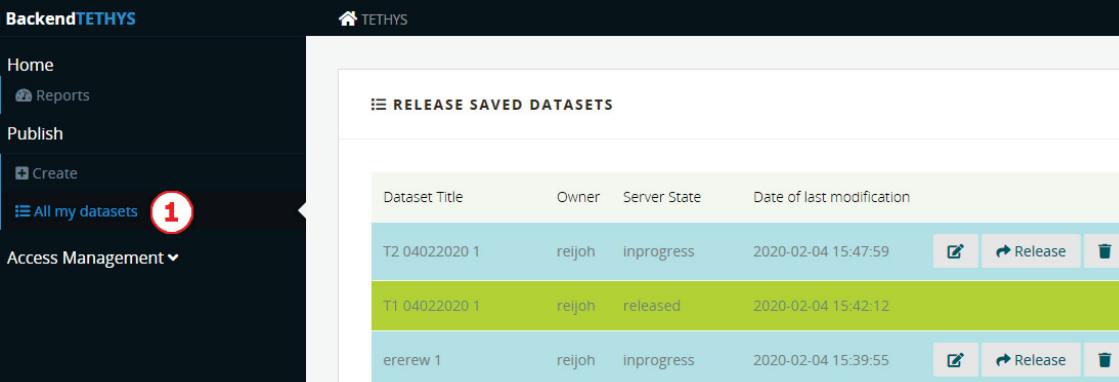
Release

Abbildung 22: Hier besteht die Möglichkeit, einen Reviewer auszuwählen.

3.2.5 Edit dataset (Datenpublikation bearbeiten)

Die Seite **EDIT DATASET** bietet eine gute Übersicht über die eingegebenen Metadaten. Hier lassen sich die Daten auch noch editieren, sofern die Datenpublikation noch nicht hochgeladen wurde. Diese können nicht mehr verändert werden.

3.2.6 Release saved datasets (Gespeicherte Datenpublikationen freigeben)



BackendTETHYS

TETHYS

RELEASE SAVED DATASETS			
Dataset Title	Owner	Server State	Date of last modification
T2 04022020 1	reijoh	inprogress	2020-02-04 15:47:59
T1 04022020 1	reijoh	released	2020-02-04 15:42:12
ererew 1	reijoh	inprogress	2020-02-04 15:39:55

Abbildung 23: Übersicht über die eigene Datenpublikationen während des Publikationsprozesses.

RELEASE SAVED DATASETS bietet einen Überblick über die eigenen eingegebenen Datenpublikationen und deren aktuellen Status (Abb. 24).

Es gibt mehrere Möglichkeiten, um zur Ansicht **Release saved datasets** zu gelangen. Eine davon ist im linken Menübereich unter **All my datasets** ① auszuwählen. Die anderen Möglichkeiten werden in Tabelle 6 weiter unten angeführt.

Tabelle 6: Bedeutung der Farben bei „eigene Datenpublikationen“.

①	hellgrün ■	released – freigegeben	Der Uploadprozess ist abgeschlossen und der Editor kann die Überprüfung der Datenpublikation starten.
②	hellblau ■	in progress – in Bearbeitung	Die Datenpublikation befindet sich noch in der Entwurfsphase. Der Uploadprozess ist noch nicht abgeschlossen. Vom Submitter können noch Änderungen der Metadaten vorgenommen werden oder auch der Entwurf komplett gelöscht werden.
③	blau ■	editor_accepted – vom Editor angenommen	Der Editor hat die Datenpublikation entgegengenommen und startet den Reviewprozess. Sobald die Datenpublikation erfolgreich publiziert ist, verschwindet diese aus der Tabelle.
④	orange ■	editor_rejected – vom Editor zurückgewiesen	Ist die Datenpublikation unvollständig oder in der Form noch nicht publikationsreif, so gibt der Editor die Datenpublikation mit einem dazugehörigen Kommentar zurück an den Submitter. Dieser kann die Publikation wieder bearbeiten.

In **RELEASE SAVED DATASETS** werden die eigenen Datensätze je nach Veröffentlichungsstadium in verschiedenen Farben angezeigt (Abb. 25).

The screenshot shows a table titled "RELEASE SAVED DATASETS" with the following data:

Dataset Title	Owner	Server State	Date of last modification	Actions
Test 5.29_November 1	reijoh	released	2020-01-31 13:01:09	1
ererew 1	reijoh	inprogress	2020-01-31 12:45:31	2
Test_20200114 1	reijoh	released	2020-01-14 09:08:50	
Test 20200112 1	reijoh	released	2020-01-13 13:14:19	
models 1	reijoh	editor_accepted	2019-11-26 10:18:50	3
Geologische 1	reijoh	editor_accepted	2019-11-26 10:14:46	
Test 20191125 1	reijoh	rejected_editor	2019-11-26 10:09:22	4
xxxxx 1	reijoh	editor_accepted	2019-11-18 09:32:11	

Callouts numbered 1 through 4 point to specific rows in the table:

- Callout 1 points to the first row (green background, released state).
- Callout 2 points to the second row (light blue background, inprogress state).
- Callout 3 points to the fifth row (light blue background, editor_accepted state).
- Callout 4 points to the eighth row (orange background, rejected_editor state).

Abbildung 24: Release saved datasets.

4. Definitionen

Viele Begriffe werden je nach Bezug und Fachdisziplin sehr unterschiedlich verwendet und auch verstanden. Damit schon im Vorfeld Missverständnisse vermieden werden können, sind in [Tabelle 7](#) einige relevante Begriffe, Kurzdefinitionen und dazugehörige Referenzen aufgeführt.

Tabelle 7: Kurzdefinitionen zu spezifischen Begriffen.

Wort	Kurzdefinition	Link/Referenzen
Datenrepository (data repository)	Ist ein verwaltetes Verzeichnis zur Speicherung und Beschreibung von digitalen Objekten für ein digitales Archiv.	https://de.wikipedia.org/wiki/Repository
DOI	Digital Object Identifier.	https://www.doi.org/ https://de.wikipedia.org/wiki/Digital_Object_Identifier
Forschungsdaten (research data)	Unter digitalen Forschungsdaten verstehen wir alle digital vorliegenden Daten, die während des Forschungsprozesses entstehen oder dessen Ergebnis sind. Der Forschungsprozess umfasst dabei verschiedene Schritte, angefangen mit der Nutzung bereits verfügbarer Forschungsdaten oder ihrer Generierung, wobei verschiedene Formen wie ein Experiment in den Naturwissenschaften, eine dokumentierte Beobachtung in einer Kulturwissenschaft oder eine empirische Studie in den Sozialwissenschaften denkbar sind. Weiter umfasst der Forschungsprozess die Aufbereitung sowie die Analyse und Auswertung von Forschungsdaten, der sich die Distribution eines Forschungsergebnisses (klassischerweise in Form einer Publikation) anschließt sowie die Datenpublikation und die Datenarchivierung der Forschungsdaten.	KINDLING, M., SCHIRMBACHER P. & SIMUKOVIC, E.: (2013): Forschungsdatenmanagement an Hochschulen: das Beispiel der Humboldt-Universität zu Berlin. LIBREAS. Library Ideas 23 (2013). urn:nbn:de:kobv:11-100212700; https://libreas.eu/ausgabe23/07kindling/
ID	Eindeutige Identifikationsnummer oder Zeichenfolge.	https://en.wikipedia.org/wiki/Identifier
Interoperabilität	Fähigkeit der Zusammenarbeit von verschiedenen Systemen.	https://de.wiktionary.org/wiki/Interoperabilit%C3%A4t
ORCID	Open Researcher and Contributor Identifier. Das ist jener Identifier, der eindeutig eine Person identifiziert. Eine gültige ORCID kann mit weiteren gültigen IDs (ISBN, DOIs, URIs,...) verknüpft werden, sodass die dazugehörige Information eindeutig der Person zuzuordnen ist.	https://orcid.org/
Datenerläuterung	Eine Datenbeschreibung, welche die Daten soweit erklärt, sodass die User ohne weitere Unterstützung die Daten nach bester wissenschaftlicher Praxis weiterverwenden können. Im Rahmen der Datenerläuterung für das Tethys RDR soll auch eine Datenspezifikation vorhanden sein.	https://www.wortbedeutung.info/Erl%C3%A4uterung/