



TETHYS RDR

Geowissenschaftlicher Datenverlag für Österreich

Datenpublikationsrichtlinien, Terms & Conditions, Anleitung

VIKTORIA HAIDER, JOHANNES REISCHER, WERNER STÖCKL,
THOMAS BRUS, CHRISTIAN CERMAK, THOMAS HOFMANN &
WERNER GESSELBAUER
viktorja.haider@geologie.ac.at

Version: 1.2

Geologische Bundesanstalt, Dienstag, 21. Juli 2020



i. Inhaltsverzeichnis

1. Datenpublikationsrichtlinien (Data Policy)	6
1.1 Leitlinie von Tethys RDR	6
1.1.1 Open Science Principles / Die offene Wissenschaft.....	6
1.1.2 FAIR-Prinzipien	7
1.1.3 Metadatenbeschreibung.....	8
1.1.3.1 Dublin Core.....	8
1.1.3.2 DataCite.....	9
1.1.3.3 Deskriptives Metadaten-set: „Dataset Type“ (Datensatzart).....	9
1.2 Die Richtlinien für das Publizieren von Forschungsdaten	11
1.2.1 Worin liegt der Mehrwert, die Forschungsdaten mit Hilfe des Tethys RDRs zu publizieren? ..	11
1.2.2 Welche Daten werden im Forschungsdatenrepositorium aufgenommen?	12
1.2.3 Welche Informationen müssen in der Datenpublikation vorhanden sein?	14
1.2.4 Können bereits im Tethys RDR publizierte Forschungsdaten wieder gelöscht werden?	14
1.2.5 Können extern erstellte Forschungsdaten publiziert werden?.....	15
1.2.6 Kann ich die Forschungsdaten publizieren, wenn mehrere Personen diese erstellt haben? ...	15
1.2.7 Welche Lizenzen stehen für die Datenpublikation zur Verfügung?	15
1.2.8 Was muss ich machen, wenn die Daten auf Fremddaten basieren?	17
1.2.9 In welchen Formaten darf ich die Daten publizieren?	17
1.2.10 In welcher Sprache kann ich die Daten publizieren?	19
1.2.11 Kann ich meine Datenpublikation einem Projekt zuordnen?	19
1.2.12 Kann ich ein Embargo setzen und was muss ich dazu alles berücksichtigen?	20
1.2.13 Muss ich meine Datenpublikation mit Keywords publizieren?.....	20
1.2.14 Gibt es ein Upload-Limit?	20
1.2.15 Kann ich die Datenpublikation auf bereits existierende Publikationen verweisen?	21
1.2.16 Wird die Datenpublikation überprüft (reviewed)?	22
1.2.17 Ich möchte verschiedene Datensatztypen gemeinsam publizieren. Ist das möglich?.....	23
1.2.18 Wie kann ich die Datenpublikation zitieren bzw. darauf verweisen?	23
1.3 Checkliste	23
2. Terms and Conditions	25
3. Technische Anleitung zur Datenpublikation	30
3.1 Zeichenerklärung	30
3.2 Eingabehilfe	30
3.2.1 Tethys Start	30
3.2.2 Login	31
3.2.3 PUBLISH NEW DATASET (Datenpublikation veröffentlichen)	31
3.2.3.1 Dataset Language (Sprache der Datenpublikation)	32
3.2.3.2 Mandatory Elements (Verpflichtende Angaben)	32
3.2.3.3 Recommended Elements (Empfohlene Angaben)	35
3.2.3.4 Other Elements (andere Angaben)	38
3.2.3.5 File Upload (Datei hochladen).....	39
3.2.3.6 Uploaded 1 file successfully (Hochladen erfolgreich)	40

3.2.4	RELEASE YOUR DATASET FOR EDITOR (Freigabe der Datenpublikation für den Editor).....	40
3.2.5	EDIT DATASET (Datenpublikation bearbeiten).....	41
3.2.6	RELEASE SAVED DATASETS (Gespeicherte Datenpublikationen freigeben).....	41
4.	<u>Definitionen</u>	43

ii. Abbildungsverzeichnis

Abbildung 1: Die sechs Prinzipien von Open Science.....	7
Abbildung 2: FAIR Daten Prinzipien adaptiert von FORCE11.	8
Abbildung 3: Deskriptive Auswahlliste „Dataset Type“.	11
Abbildung 4: Mögliche Lizenzvarianten, geordnet nach Grad der Offenheit.	16
Abbildung 5: Startseite des Tethys RDR.	30
Abbildung 6: Einstiegsfenster in den Publikationsbereich.	31
Abbildung 7: Erscheinungsbild nach dem Login.....	31
Abbildung 8: Auswahl der Hauptsprache.....	32
Abbildung 9: Verpflichtende Angaben zu der Datenpublikation.	33
Abbildung 10: Metadaten zu den Autoren der Datenpublikation.	34
Abbildung 11: Metadaten zu den Mitwirkenden der Datenpublikation.....	34
Abbildung 12: Name des Betreibers und das verpflichtend anzukreuzende Feld die terms and conditions betreffend.	35
Abbildung 13: Zweite Publikationsstufe zu den empfohlenen Angaben und die verpflichtende Angabe zu Geo Location.....	36
Abbildung 14: Ausdehnung der Datenpublikation.....	37
Abbildung 15: Möglichkeit zur Erfassung von Referenzen und die Angabe von Schlüsselworten.	37
Abbildung 16: Andere Angaben – Nutzungsbestimmungen.....	38
Abbildung 17: Hochladen der Datenpublikation 1.....	39
Abbildung 18: Hochladen der Datenpublikation 2.....	39
Abbildung 19: Edit, Release, Delete.	40
Abbildung 20: Hier besteht die Möglichkeit, einen Reviewer auszuwählen.....	40
Abbildung 21: Übersicht über die eigene Datenpublikationen während des Publikationsprozesses.....	41
Abbildung 22: Bedeutung der Farben bei „eigene Datenpublikationen“.	42
Abbildung 23: Forschungsdaten publizieren oder nicht publizieren, dass ist hier die Frage! (Bild gezeichnet von Ainsley Seago.)	44
Abbildung 24: Shortfilm (englisch) im Genre „Wissenschaftshorror“.	44

iii. Tabellenverzeichnis

Tabelle 1: Creative Commons Licences.....	16
Tabelle 2: Mögliche Lizenzvergaben bei Verwendung von Fremddaten.	17
Tabelle 3: Auflistung von unterstützten und geplant unterstützten Formaten.....	18
Tabelle 4: Ausgewählte Beispiele für nicht unterstützte Formate.....	19
Tabelle 5: Definition der zur Verfügung gestellten Identifier Types.....	21
Tabelle 6: Definitionen und Beispiele von möglichen Beziehungstypen.	22
Tabelle 7: Release saved datasets.....	42
Tabelle 8: Kurzdefinitionen zu spezifischen Begriffen	43

1. Datenpublikationsrichtlinien (Data Policy)

1.1 Leitlinie von Tethys RDR

Tethys RDR ist ein digitales geowissenschaftliches Forschungsdatenrepositorium der Geologischen Bundesanstalt (GBA).

In diesem digitalen Repositorium sollen in erster Linie digitale, georeferenzierte Forschungsdaten, die an bzw. im Auftrag der GBA generiert wurden, publiziert und 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 in deutscher Sprache bereitgestellt, eine englische Version ist vorgesehen.

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 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 über Rechte sowie Pflichten informiert.

Tethys ist zum jetzigen Zeitpunkt noch nicht als Forschungsdatenrepositorium zertifiziert. Die Grundlage dazu wurde bereits geschaffen. Wenn der Bedarf dazu besteht, kann eine Zertifizierung angestrebt werden.

Das Tethys RDR baut auf zwei in der Wissenschaftswelt bekannte und anerkannte Prinzipien auf, die nachhaltiges Archivieren sowie Publizieren von weiterverwertbaren wissenschaftlich relevanten Daten ermöglichen: Open Science Principles^{1,2} sowie FAIR Guiding Principles³.

1.1.1 Open Science Principles / Die offene Wissenschaft

Die Open Science Principles bestehen aus sechs Prinzipien (**Abbildung 1**). Besonderer Fokus liegt im speziellen auf Open Data⁴ und auf Open Access^{5,6}.

¹ https://de.wikibooks.org/wiki/Handbuch_Open_Science (abgerufen am 24.01.2020)

² https://en.wikipedia.org/wiki/Open_science (abgerufen am 24.01.2020)

³ 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>

⁴ https://de.wikipedia.org/wiki/Open_Data (abgerufen am 24.01.2020)

⁵ <http://open-access.net/AT-DE/informationen-zu-open-access/was-bedeutet-open-access/> (abgerufen am 07.03.2018)

⁶ https://openaccess.mpg.de/mission-statement_de (abgerufen am 07.03.2018)

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 und bei Bedarf die Verwendung von Copyleft⁷ bzw. CC BY-SA⁸ als einzige Einschränkung der Nutzungsrechte erlaubt.

Open Access steht für einen offenen, uneingeschränkten Zugang zu digitaler wissenschaftlicher Literatur. Im Fall des Tethys RDRs 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.

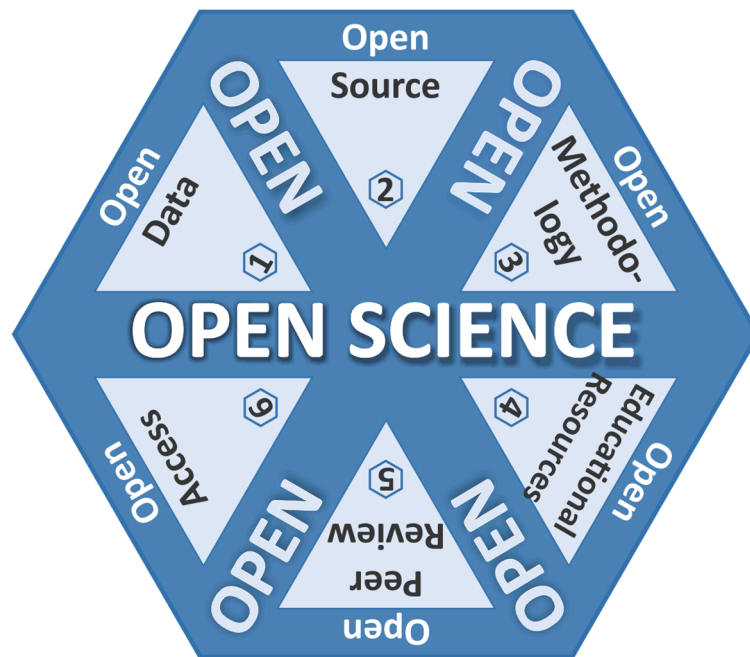


Abbildung 1: Die sechs Prinzipien von Open Science.

1.1.2 FAIR-Prinzipien

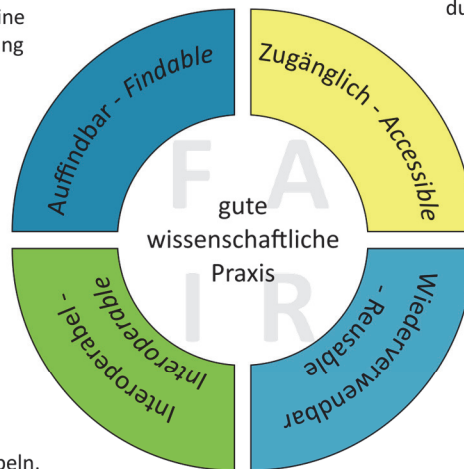
Die Prinzipien nach FAIR, die ein Akronym für *Findability* (Auffindbarkeit), *Accessibility* (Zugänglichkeit), *Interoperability* (Kompatibilität) und *Re-usability* (Wiederverwendbarkeit) sind (Abbildung 2), unterstützen eine nachhaltige wissenschaftliche Datenpflege sowie deren Verwaltung. Erst durch 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 Wert.

⁷ <https://de.wikipedia.org/wiki/Copyleft> (abgerufen am 25.03.2020)

⁸ <https://creativecommons.org/licenses/by-sa/3.0/at/> (abgerufen am 25.03.2020)

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

- Die (Meta-)Daten sind einem weltweit einzigartigen sowie unvergänglichen Identifier zugeordnet
- Die Daten sind mit umfangreichen Metadaten beschrieben
 - Die (Meta-)Daten sind in einem durchsuchbaren Resource registriert bzw. indiziert.
 - Die (Meta-)Daten sind fest verknüpft mit dem Identifier



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

- Die (Meta-)Daten haben eine Vielzahl von genauer und relevante Attribute aufzuweisen
- Die (Meta-)Daten werden mit einer klaren und zugänglichen Nutzungsrechtlizenz veröffentlicht
 - Die (Meta-)Daten sind mit der Datenherkunft verbunden
- Die (Meta-)Daten erfüllen den in der Forschungsgemeinschaft übliche Standards

content of the graphic: <https://www.force11.org/group/fairgroup/fairprinciples>

Abbildung 2: FAIR Daten Prinzipien adaptiert von FORCE11.⁹

1.1.3 Metadatenbeschreibung

Metadatenbeschreibungen sind essenziell, um die Datenpublikationen sowohl innerhalb als auch außerhalb des Research Data Repositorys 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 zwei weitverbreitete Metadatenstandards (Dublin Core¹⁰ und DataCite¹¹) sowie ein deskriptives Metadatenelement (dataset type) verwendet.

1.1.3.1 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“ (DCMI¹²) entwickelt und umfasst 15 Kernfelder, welche eine grundsätzliche Beschreibung der Objekte zulassen. Die genaue Definition der einzelnen Elemente können unter folgender Adresse nachgelesen werden: <https://www.dublincore.org/specifications/dublin-core/dcmi-terms/> (abgerufen am 25.03.2020) bzw. https://de.wikipedia.org/wiki/Dublin_Core (abgerufen am 25.03.2020).

⁹ <https://www.force11.org/group/fairgroup/fairprinciples> (abgerufen am 07.03.2018)

¹⁰ <https://www.dublincore.org/> (abgerufen am 25.11.2019)

¹¹ <https://datacite.org/> (abgerufen am 25.11.2019)

¹² <https://www.dublincore.org/> (abgerufen am 25.03.2020)

1.1.3.2 DataCite

DataCite, ebenfalls ein internationales Konsortium, stellt ein sehr umfangreiches Metadaten-Schema¹³ 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 und man unterwirft sich der strengen Überprüfung.

Die genaue Definition der einzelnen Elemente kann unter folgender Adresse nachgelesen werden:

<https://schema.datacite.org/> bzw. <https://de.wikipedia.org/wiki/DataCite>

Im Tethys RDR wurde der DataCite Metadatenstandard implementiert und bis auf die Elemente betreffend der DOI vollumfänglich verwendet. Dies ist die Grundlage, um in weiterer Folge für die publizierten Forschungsdaten die DOI-Vergabe anzustreben.

1.1.3.3 Deskriptives Metadatenet: „Dataset Type“ (Datensatzart)

Nachdem Tethys RDR nicht nur supplementäre Forschungsdaten zu Textpublikationen veröffentlicht, sondern auch andere Datentypen, wurde ein weiteres, rein deskriptives Metadatenet („Dataset Type“) eingeführt. Darunter fallen die sieben Metadatenelemente: Analysis, Measurements, Monitoring, Remote sensing, Geoinformationssystem, Models und Mixed Type (Abbildung 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.

1.1.3.3.1 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.

Beispiele:

- Geochemische Zusammensetzung von Gesteinen.
- Altersdatierungen.
- Zusammensetzung von thermalen Wässern.

1.1.3.3.2 Measurements

„Eine Messung ist das Ausführen von geplanten Tätigkeiten zu einer quantitativen Aussage über eine Messgröße durch Vergleich mit einer Einheit.“¹⁴

Beispiele:

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

1.1.3.3.3 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 reproduzierbar sind.

Beispiele:

¹³ <https://schema.datacite.org/> (abgerufen am 25.03.2020)

¹⁴ Definition entnommen aus Wikipedia: <https://de.wikipedia.org/wiki/Messung> (abgerufen am 27.11.2019)

- 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.

1.1.3.3.4 Remote sensing (RS)

Im Rahmen der Fernerkundung werden Daten über die Erdoberfläche durch berührungsfreie Erkundung generiert. Als Grundlage und für weitere Messungen und Interpretationen dienen reflektierende elektromagnetische Wellen oder Schallwellen.

Beispiele:

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

1.1.3.3.5 Geoinformationssystem (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.

Beispiele:

- Objektarten im Sinne von GeoSciML (Geologische Einheiten, Strukturlinien, Hydrogeologische Objekte, Profile, Bohrungen, Abbaue etc.).
- Geometrie- und Sachdaten zu geologischen Karten.

1.1.3.3.6 Models

Ein Modell ist ein vereinfachtes Abbild der Realität, in der unter anderem auch Veränderungen simuliert werden können. Durch das Quantifizieren von Modelldaten können für die Wissenschaft relevante Daten entstehen.

Beispiele:

- 3D-Modelle.
- Modellierete geologische Profile.
- Tiefenmodellierungen.
- Versickerungsmodelle.

1.1.3.3.7 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:

- 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.

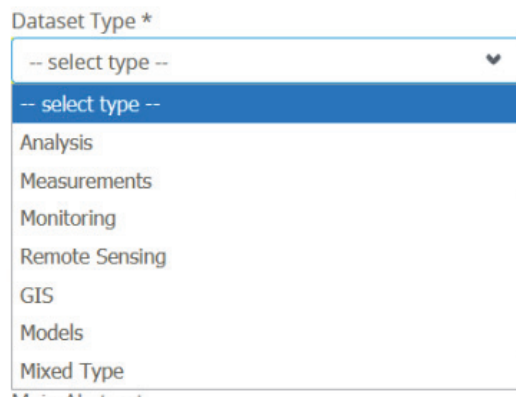


Abbildung 3: Deskriptive Auswahlliste „Dataset Type“.

1.2 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 zu ermöglichen, wurden die Richtlinien so aufgebaut, dass ein Überblick sehr schnell geschaffen werden kann.

Die Tethys RDR-Richtlinie wurde in Form eines Fragenkatalogs aufgebaut. 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 ausführlichere Beantwortung für alle, die noch mehr Hintergrundwissen benötigen oder daran interessiert sind.

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.

1.2.1 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.

- Interessierte können weltweit auf die Daten zugreifen und im Rahmen der bereitgestellten Lizenzbedingungen diese nachvollziehbar und nach wissenschaftlicher „Best Praxis“ 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.
- Durch das Zitieren der Datenpublikation kann der eigenen Leistung gezollt werden.
- Die Daten verschwinden nicht in der lokalen oder institutionellen Computerperipherie und gewinnen dadurch auch an Wert.
- Datenpublikationen sind maschinenlesbar abgelegt und können 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 GBA gehostet.
- Neben den Datensätzen können ergänzende Daten publiziert werden.
- Das Publizieren der Daten ist kostenfrei.

1.2.2 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, 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 Nutzbarkeit von Forschungsergebnissen.“¹⁵

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 arbeiten 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

¹⁵ https://www.dfg.de/foerderung/info_wissenschaft/2015/info_wissenschaft_15_66/index.html (abgerufen am 23.01.2020)

Erlangung wissenschaftlicher Kenntnis sein, also noch unprozessierte Datensätze. Sie können aber auch bereits reproduzierbar verarbeitet sein und auf die 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 weitere Datensupplements publiziert werden.

Neu: Das 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 Relation zu dem Mehrwert für den Nutzer stehen.

Fachrelevant: Tethys RDR startet als Forschungsdatenrepositorium für die Geologische Bundesanstalt. Daher werden alle geologischen sowie geophysikalischen Forschungsdaten publiziert, die eine naheliegende fachliche Relevanz für die Geologische Bundesanstalt haben. Forschungsdaten, die im Zuge einer Kooperation mit der GBA entstehen, können ebenfalls über Tethys RDR publiziert werden.

Georeferenziert: Alle in 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 der klassische Nutzer die publizierten Forschungsdaten mit der dazugehörigen Datenerläuterung und den Supplementen verstehen und verwenden kann, ohne die Daten zuerst aufbereiten zu müssen. Die Einholung zusätzlicher Informationen von den Datenautoren ist ebenfalls nicht nötig.

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 Wissenschaftlern 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, 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 es 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 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 in Tethys RDR verlinkt werden.

1.2.3 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 Datenerläuterung. Der dritte Teil besteht aus den eigentlichen Datensätzen.

Metadaten: Die Metadaten sind das Herzstück einer jeden digitalen Publikation. Sie haben die essenzielle Funktion, 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 hinausgehen. Sie ist damit auch ein essenzieller Bestandteil der Publikation. Ziel der Datenerläuterung ist, 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 den 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 bestehen¹⁶ und diese ohne weitere händisch oder/und technisch durchgeführte Zwischenschritte erst als solche aufbereitet werden müssen.

1.2.4 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, allerdings bleiben die Metadaten ausnahmslos erhalten, da die Zitierfähigkeit sowie die wissenschaftliche Nachvollziehbarkeit zu jeder Zeit gewährleistet sein soll.

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 komplett gesperrt werden, sodass diese nicht mehr verbreitet werden können.

¹⁶ <https://www.techfacts.de/ratgeber/was-sind-daten> (abgerufen am 16.01.2020)

1.2.5 Können extern erstellte Forschungsdaten publiziert werden?

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

Zum jetzigen Zeitpunkt (März 2020) 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.

1.2.6 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, wäre dies möglich.

Der Submitter übernimmt die Verantwortung über die Datenpublikation und ist somit auch dazu angehalten, von seinen Co-Autoren eine Einverständniserklärung einzuholen. Der Tethys RDR-Editor nimmt sich das Recht heraus, diese Einverständniserklärung vorlegen zu lassen bzw. die mitangeführten Co-Autoren persönlich über die bevorstehende Datenpublikation in Kenntnis zu setzen.

1.2.7 Welche Lizenzen stehen für die Datenpublikation zur Verfügung?

Für die Datenpublikation stehen ausschließlich die Lizenzen nach Creative Commons¹⁷ (

Tabelle 1) zur Verfügung. Das Konzept von Creative Commons umfasst sieben verschiedene Lizenztypen, welche die Daten unterschiedlich weit öffnen (**Abbildung 4**). 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 gibt es weitere Alternativen. Die Details dazu sind im unten angeführten Text enthalten.

Achtung! Werden Fremddaten mitpubliziert, so ist zu prüfen, ob die vorhandene Nutzungslizenz mit der zu vergebenen Lizenz kompatibel ist (→**1.2.8**).

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.

CC BY 4.0 und CC BY-SA (

Tabelle 1) sind jene Lizenzen, die mit den Open Data Prinzipien konform sind¹⁸, uneingeschränkte Wissenstransfers fördert 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.

¹⁷ <https://creativecommons.org> (abgerufen am 21.01.2020)

¹⁸ https://de.wikipedia.org/wiki/Open_Data (abgerufen am 24.01.2020)

Die Rechte können bei Bedarf und mit guter Begründung weiter eingeschränkt werden, allerdings verlieren die publizierten Daten den Status von Open Science bzw. Open Access. Dies könnte auch in Konflikt mit den Anforderungen seitens des Forschungsförderers stehen und sollte vor der Datenpublikation zweifelsfrei geklärt sein. Mit jeder weiteren Einschränkung durch zusätzliche Bedingungen wie SA, NC und ND wird naturgemäß auch die Weiternutzung der Datenpublikation eingegrenzt.

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
CC BY-NC	http://creativecommons.org/licenses/by-nc/4.0	Namensnennung; keine kommerzielle Nutzung
CC-BY-ND	http://creativecommons.org/licenses/by-nd/4.0	Namensnennung; keine Veränderung erlaubt
CC-BY-NC-SA	http://creativecommons.org/licenses/by-nc-sa/4.0	Namensnennung; keine kommerzielle Nutzung; Weitergabe unter gleicher Bedingung
CC-BY-NC-ND	http://creativecommons.org/licenses/by-nc-nd/4.0	Namensnennung; keine kommerzielle Nutzung; keine Veränderung erlaubt

Tabelle 1: Creative Commons Licences.

Nach erfolgreicher Veröffentlichung der Forschungsdatenpublikation können einmal vergebene Lizenzen nicht mehr verändert werden. Daher ist im Vorfeld gründlich zu prüfen (Abbildung 4), 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 (siehe dazu auch Tabelle 2).

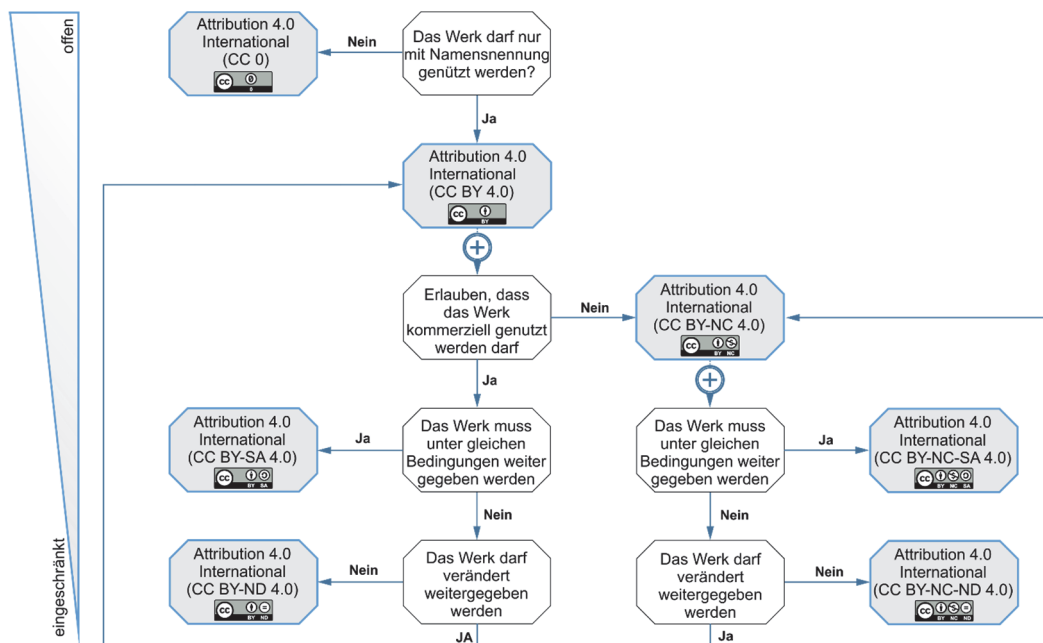


Abbildung 4: Mögliche Lizenzvarianten, geordnet nach Grad der Offenheit.

1.2.8 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 (Tabelle 2).

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 eines Rechtsexperten zu klären.

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

		Mögliche CC Lizenzen für die Datenpublikation						
		Tethys RDR konform	CC-BY	CC-BY-SA	CC BY-NC	CC-BY-ND	CC-BY-NC-SA	CC-BY-NC-ND
CC Lizenzen Fremddaten	CC-00	Ja	✓	✓	✓	✓	✓	✓
	CC-BY	Ja	✓	✓	✓	✓	✓	✓
	CC-BY-SA	Ja	✗	✓	✗	✗	✗	✗
	CC BY-NC	Ja	✗	✗	✗	✗	✓	✓
	CC-BY-ND	Nein	✗	✗	✗	✗	✗	✗
	CC-BY-NC-SA	Ja	✗	✗	✗	✗	✓	✗
	CC-BY-NC-ND	Nein	✗	✗	✗	✗	✗	✗

Tabelle 2: Mögliche Lizenzvergaben bei Verwendung von Fremddaten.

1.2.9 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 Nutzbarkeit digitaler Daten sicherzustellen haben sich im Laufe der letzten Jahre einige Formate bewährt.“¹⁹

Damit Daten auch noch 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

¹⁹ <https://www.forschungsdaten.org/index.php/Dateiformat> (abgerufen am 13.02.2020)

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 weiter unten aufgelistet.

Unterstützte Formate	Anmerkung	Referenz
<i>Textfiles</i>		
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/
dxg	AutoCAD mit ASCII Support	https://en.wikipedia.org/wiki/AutoCAD_DXF http://www.crlf.de/Verlag/DXF-intern/DXF-intern.html
<i>Dokumente</i>		
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
<i>Bilder</i>		
JPEG	ISO10918-1:1994	http://en.wikipedia.org/wiki/Jpeg http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=18902
PNG	ISO/IEC 15948	https://en.wikipedia.org/wiki/Portable_Network_Graphics
<i>Geodaten</i>		
Shapefile (shp)		http://www.esri.com/library/whitepapers/pdfs/shapefile.pdf https://de.wikipedia.org/wiki/Shapefile
GeoPackage (gpkg)		https://de.wikipedia.org/wiki/GeoPackage
<i>Models</i>		
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

Tabelle 3: Auflistung von unterstützten und geplant unterstützten Formaten.

Folgende Tabellen geben einen Überblick über die Eignung betreffend Langzeitarchivierung (Tabelle 3) bzw. Nichteignung von Formaten (Tabelle 4). Diese Tabellen haben keinen Anspruch auf Vollständigkeit. Sie werden auch laufend an den aktuellen Stand der Technologie angepasst.

Nicht unterstützte Formate (ausgewählte Beispiele)

Document Format ArcGIS (mxd)
FileGeodatabase (gdb)
MS Excel < 2013 (xls)
PersonalGeodatabase (mdb)
Projektdatei ArcGIS Pro (aprx)
QGIS (qgz)
Skriptsprachen

Tabelle 4: Ausgewählte Beispiele für nicht unterstützte Formate.

1.2.10 In welcher Sprache kann ich die Daten publizieren?

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

Grundsätzlich wird empfohlen, 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 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 soll die Publikation einheitlich in einer Sprache publiziert werden. Es besteht die zusätzliche Möglichkeit, den Titel in zwei Sprachen zu publizieren.

Neben dem klassischen Abstract können zusätzlich noch eine Methoden- und/oder technische Beschreibung hinzugefügt werden. Es wird jedoch ausdrücklich empfohlen, detailliertere Beschreibungen in Form eines pdf zusammen mit den Datensätzen zu publizieren.

1.2.11 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 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 Uploadprozess der Datenpublikation das Projekt in der Auswahlliste gefunden und ausgewählt werden.

Nach der erfolgreichen Datenpublikation kann im Frontend unter „Search“ nach der Projektbezeichnung gefiltert werden.

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

Die Datenpublikation kann bis zu maximal zwei Jahre vor Freigabe 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 in Widerspruch mit den Interessen des Fördergeldgebers steht.

1.2.13 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. Zusätzlich wird eine Tabelle mit kontrolliertem Vokabular bereitgestellt, aus der standardisierte Keywords in der jeweiligen Hauptsprache ausgewählt werden können. 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.

1.2.14 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 den Anwender zu gewährleisten.

Der Download der Datenpublikation erfolgt nicht nur innerhalb der GBA, sondern weltweit mit unterschiedlichen technischen Voraussetzungen. Um einen schnellen und unkomplizierten Download zu ermöglichen, wurde der Datenumfang auf eine Gesamtgröße von 512 MB limitiert.

1.2.15 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 (Tabelle 5 bzw. Tabelle 6).

Dabei ist zu beachten, dass nur auf nachhaltige Referenzen Bezug genommen werden kann.

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 (Tabelle 5).

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. ²⁰ Beispiel: https://doi.org/10.1594/PANGAEA.701578
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. ²¹ Beispiel: https://hdl.handle.net/20.500.11756/dd4424fd
ISBN	Die Internationale Standardbuchnummer (ISBN) ist ein eindeutiges Nummernkennzeichen von Büchern und anderen redaktionierten 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. ²² Als URL (Abk.: Uniform Resource Locator) versteht man eine definierte Adresse, welche auf die Position einer Datei auf einem Server zeigt und diese abrufen. URLs werden in einem Webbrowser eingegeben, um auf Dokumente im Web zuzugreifen oder werden als Hyperlinks innerhalb eines Dokumentes eingebettet. ²³ Beispiel: http://www.domainname.at/verzeichnis1/datei1.html
URN	Ein Uniform Resource Name (URN) ist ein dauerhafter, ortsunabhängiger Bezeichner für eine Ressource (z.B. Datenpublikation). ²⁴ Das URN-Schema: <i>urn:<NID>:<NID-spezifischer Teil></i>

Tabelle 5: Definition der zur Verfügung gestellten Identifier Types.

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 (Tabelle 6).

²⁰ <https://www.forschungsdaten.org/index.php/DOI> (abgerufen am 05.12.2019)

²¹ <https://www.forschungsdaten.org/index.php/Handle> (abgerufen am 05.12.2019)

²² <https://fischerlaender.de/lexikon/url-uri-urn> (abgerufen am 06.12.2019)

²³ <https://de.ryte.com/wiki/URL> (abgerufen am 06.12.2019)

²⁴ https://de.wikipedia.org/wiki/Uniform_Resource_Name (abgerufen am 05.12.2019)

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 Autoren 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 Autoren der verwendeten Fremddaten nicht verletzt werden.
IsVariantFormOf	Werden Datensätze kompiliert und neu strukturiert, so kann dieser Beziehungstyp verwendet werden.	

Tabelle 6: Definitionen und Beispiele von möglichen Beziehungstypen.

1.2.16 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 bereitgestellte 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.

1.2.17 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 (→ 1.2.15) aufeinander zu verweisen.

Datensatztypen: Dieses Metadatenelement dient zur Klassifizierung vom 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 abzuwägen, 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.

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

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

Derzeit ist eine persistente ID, die auch gleichzeitig die URI ist, mit der Datenpublikation verbunden. Diese URI ist einmalig und kann weder ausgetauscht noch verändert werden. Bei bekannter URI kann man damit direkt die Landingpage der betroffenen Datenpublikation aufrufen bzw. darauf referenzieren.

Vorgeschlagener Zitationsstyle für die Datenpublikation: Autor/Autoren (Jahr der Veröffentlichung): Titel der Datenpublikation. Tethys RDR, URI (z.B.: <https://tethys.at/id/1407>).

1.3 Checkliste

Wer mit der Datenpublikation vertraut ist, kann die erste Frage als Checkliste verwenden und benötigt nicht oder nur vereinzelt die Unterstützung der restlichen ausführlicheren Fragen.

- Eine nachträgliche Veränderung oder Löschung der publizierten Daten ist nicht vorgesehen. (→ 1.2.3)
- Es handelt sich um Forschungsdaten erstellt von bzw. in Kooperation mit der Geologischen Bundesanstalt. (→ 1.2.5)
- Die zu publizierten Daten sind Tethys RDR konforme fachrelevante Forschungsdaten. (→ 1.2.2)
- Daten sind wissenschaftlich korrekt strukturiert, konsistent und vollständig. (→ 1.2.2)

- Zusätzlich zu den Datensätzen wird eine selbsterklärende Beschreibung der verwendeten Parameter und die korrekte Nutzung bereitgestellt. ([→ 1.2.3](#))
- Alle beteiligten Autoren wurden über die geplante Datenpublikation in Kenntnis gesetzt und es wurde das Einverständnis dazu eingeholt. ([→ 1.2.6](#))
- Es ist geklärt, welche Nutzungslizenz (Creative Commons) für die Datenpublikation vergeben wird. ([→ 1.2.7](#))
- Es sind gegebenenfalls die Rechte der verwendeten Fremddaten geklärt. ([→ 1.2.8](#))
- Die Datensätze liegen in offenen und nachhaltigen Formaten vor, die von Tethys RDR unterstützt werden. ([→ 1.2.9](#))
- Die Datenpublikation findet im Rahmen eines Projektes statt. Die Information dazu wurde bereits dem Editor gemeldet. ([→ 1.2.11](#))
- Es ist ein Embargo notwendig und wird bei der Publikation mitberücksichtigt. ([→ 1.2.12](#))
- Der Abstract der Datenpublikation liegt sowohl auf Englisch als auch auf Deutsch vor. ([→ 1.2.10](#))
- Gesamtgröße der hochgeladenen Datenfiles < 512 MB. ([→ 1.2.14](#))

2. Terms and Conditions

Die Domain-Inhaberin (die GBA), 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.

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.

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 lizenzfreie 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.

Datenschutz-Bestimmungen

Mit dieser Datenschutzerklärung möchte unsere gastgebende Einrichtung, die „Geologische Bundesanstalt“, 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 Autoren, deren Co-Autoren sowie Beitragende, wie Name, E-Mail-Adresse und optional die ORCID ID erfolgt stets durch die datenpublizierenden Autoren im Einklang mit der Allgemeinen Datenschutzgrundverordnung (DSGVO) nach Artikel 6, Absatz 1b. Diese Informationen werden zum Zwecke der Anmeldung der datenpublizierenden Autoren und zum Zwecke der Umsetzung der guten wissenschaftlichen Praxis gesammelt. Damit einher geht, 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, so dass ein absoluter Schutz möglicherweise nicht gewährleistet ist.

Sie haben grundsätzlich das Recht auf Auskunft, Berichtigung, Löschung, Einschränkung, Datenübertragbarkeit und Widerspruch. Dafür wenden Sie sich bitte an die Datenschutzbeauftragte:

Dr. Viktoria Haider

E-Mail: datenschutz@geologie.ac.at

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/>

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 Browsertypen 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-websites, (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, welche 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.

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.

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.

Übermittlung wissenschaftlicher Daten über die Website

Die Website von Tethys RDR enthält eine Webanwendung zum Einreichen von wissenschaftlichen Datensätzen, die in der Tethys -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. Diese von einer betroffenen Person freiwillig übermittelten personenbezogenen Daten werden zum Zwecke der Verarbeitung oder Kontaktaufnahme mit der betroffenen Person gespeichert.

Um wissenschaftliche Datensätze erfolgreich bei Tethys RDR einzureichen, muss die betroffene Person einige zusätzliche personenbezogene Daten (z.B. Name der Autoren, der Mitautoren 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-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.

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.

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.

Uns kontaktieren

Bei Fragen zu dieser Datenschutzrichtlinie wenden Sie sich bitte an uns:

Geologische Bundesanstalt
Neulinggasse 38
1030 Wien
+43-1-712 56 74
repository@geologie.ac.at

Änderungen unserer Datenschutzerklärung

Wenn wir uns entscheiden, unsere Datenschutzrichtlinie zu ändern, werden wir diese Änderungen auf dieser Seite veröffentlichen. Diese Richtlinie wurde zuletzt am 25.03.2020 geändert.

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 in der Datenschutzseite von OpenStreetMap und auf http://wiki.openstreetmap.org/wiki/Legal_FAQ.

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 Research Data Repository (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 (Abbildung 5).



Abbildung 5: Startseite des Tethys RDR.

Die Eingabe einer im Tethys RDR zu publizierenden Datenpublikation beginnt mit dem **LOGIN** ① auf der Startseite (Abbildung 5). Der Account für das Tethys RDR muss bei dem für das RDR zuständigen Administrator beantragt werden. Derzeit dafür zuständig ist Viktoria Haider.

3.2.2 Login

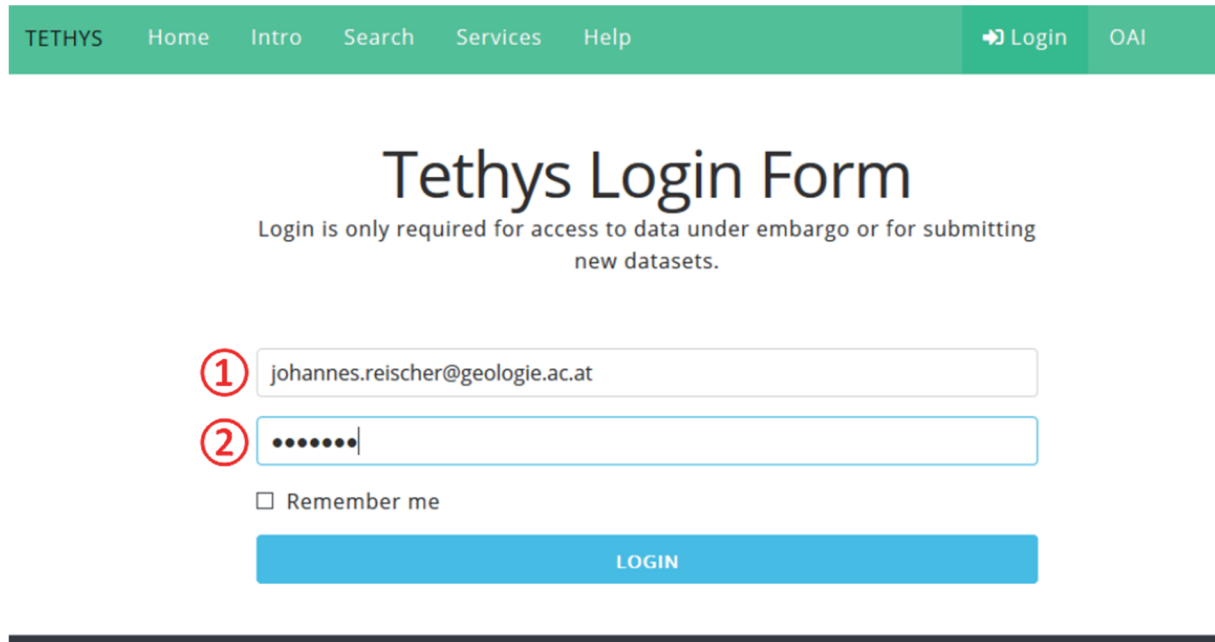


Abbildung 6: Einstiegsfenster in den Publikationsbereich.

Für das Login (Abbildung 6) werden **E-Mail Adress** (E-Mail-Adresse) ① und **Password** (Passwort) 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.

3.2.3 PUBLISH NEW DATASET (Datenpublikation veröffentlichen)

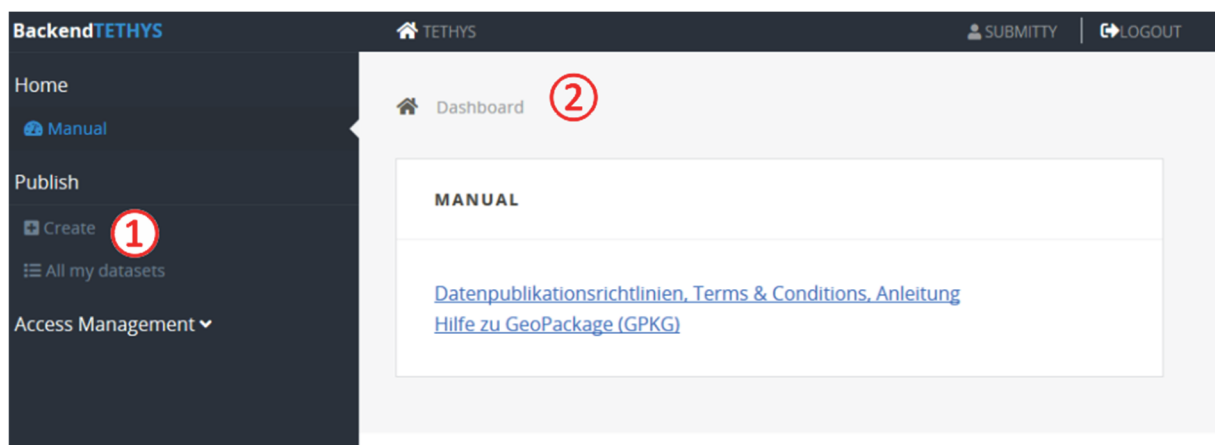


Abbildung 7: Erscheinungsbild nach dem Login

Um eine Datenpublikation publizieren zu können, muss im Menü links unter der Überschrift **Publish+ Create** ① angeklickt werden (Abbildung 7).

Direkt im Dashboard ② steht das Handbuch (Datenpublikationsrichtlinien, Terms & Conditions, Anleitung) und eine Anleitung zu GeoPackage (Hilfe zu GeoPackage) zum Download bereit.

3.2.3.1 Dataset Language (Sprache der Datenpublikation)

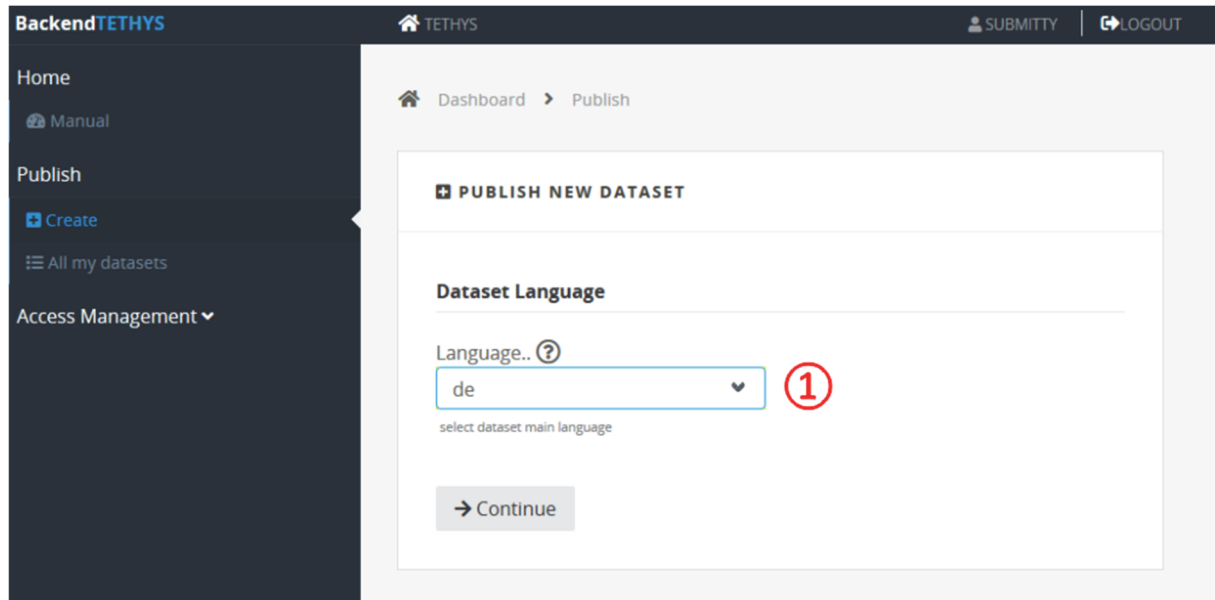


Abbildung 8: Auswahl der Hauptsprache.

Bei **Language** ① muss die Hauptsprache der Datenpublikation ausgewählt werden (Abbildung 8). 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“.

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 (Abbildung 9). 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 eingegeben werden. 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 --

Title(s)

Main Title [?] 2 Title Language.. 3
 de

Add additional title(s) [?] + 4

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

Description

Main Abstract [?] 7 Abstract Language.. 6
 de

Add additional descriptions(s) [?] + 8

Description	Type	Language
[DESCRIPTION]	[descriptionType]	[language]

[descriptionType] dropdown options:
 Methods
 Series_information
 Technical_info
 Translated
 Other

Abbildung 9: Verpflichtende Angaben zu der 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. Die maximale Zeichenlänge beträgt jeweils 2.500 Zeichen.

3.2.3.2.4 Creator (Autorin/Autor)

Creator

searching active person table **1**

Add additional creator(s) if creator is not in database + **2**

creator table

#	ID	First Name	Last Name	Email	Orcid
1	3	[FIRST NAME] 4	[LAST NAME] 5	[EMAIL] 6	[ORCID optional] 7

Abbildung 10: Metadaten zu den Autoren der Datenpublikation.

In Feld **1** kann aus einer Liste von bereits vorhandenen Autoren ausgewählt werden (Abbildung 10). Ein Autor kann auch eine Organisation sein. Um zusätzliche Autoren einzutragen oder Autoren, die noch nicht in der Datenbank abgelegt wurden, bitte auf + **2** klicken. Für jeden weiteren zusätzlichen Autor muss abermals auf + geklickt werden. Dann öffnet die Applikation die **creator table**. **First Name** (Vorname) **4**, **Last Name** (Nachname) **5** und **E-Mail** **6** müssen ausgefüllt werden. Die **ORCID** **7** ist eine dauerhafte digitale Kennung für Autoren (Forschende) und kann optional und, wenn bekannt, ausgefüllt werden. Die ID **3** wird vom System vergeben.

Der Eintrag zumindest eines Autors ist verpflichtend.

3.2.3.2.5 Contributor (Mitwirkende, Beitragende)

Contributor

searching active person table **1**

Add additional contributor(s) if contributor is not in database + **2**

contributor table

#	ID	First Name	Last Name	Email	Orcid
1	3	[FIRST NAME] 4	[LAST NAME] 5	[EMAIL] 6	[ORCID optional] 7

Abbildung 11: Metadaten zu den Mitwirkenden der Datenpublikation.

Hier **1** sollen jene Personen oder Organisationen eingetragen werden (Abbildung 11), die zur Erstellung der Dateninhalte beigetragen haben oder die Autoren bei der Datengenerierung unterstützt haben. Um zusätzliche Mitwirkende oder Beitragende einzutragen, welche noch nicht in der Datenbank erfasst sind, zunächst auf + **2** klicken. Für jeden weiteren zusätzlichen Mitwirkenden muss erneut auf + geklickt werden. Dann öffnet die Applikation die **contributor table**. **First Name** (Vorname) **4**, **Last Name** (Nachname) **5** und **Email** **6** müssen ausgefüllt werden. Die **ORCID** **7** ist eine dauerhafte digitale Kennung für Autorinnen und Autoren (Forschende) und kann, wenn bekannt, ausgefüllt werden. Die ID **3** wird vom System vergeben.

3.2.3.2.6 Creating Corporation (Betreiber)

Creating Corporation

Corporation Name ?
TETHYS RDR

terms and conditions i
You must agree to continue

Abbildung 12: Name des Betreibers und das verpflichtend anzukreuzende Feld die terms and conditions betreffend.

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 werden und dies durch das Setzen eines Häkchens bestätigt werden (Abbildung 12).

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 (Abbildung 13). 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. 14.09.2020.

3.2.3.3.3 Geo Location (Lage)

In der Karte kann mit dem Werkzeug ④ ein Rechteck aufgezoogen werden, das die Lage der Daten der Datenpublikation umgrenzt. Die Koordinaten werden als Geographische Koordinaten (WGS84) im 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 aufgezoogen und die Koordinaten des alten Rechtecks werden überschrieben. Die Koordinaten können auch per Tastatur in die Felder eingegeben werden.

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.

Step Two: Recommended Elements

Project

Project.. ?

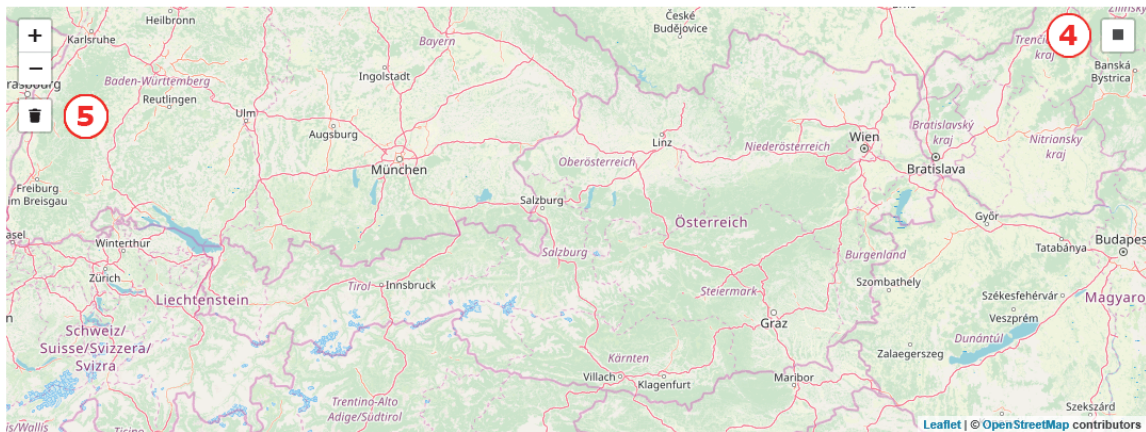
project is optional

Date(s)

Embargo Date.. ?

EmbargoDate is optional

Geo Location ?



xmin:

ymin:

max:

ymax:

Abbildung 13: Zweite Publikationsstufe zu den empfohlenen Angaben und die verpflichtende Angabe zu Geo Location.

Coverage (Bereich)

Unter **Coverage** (Abbildung 14) 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.: 2020-04-01 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 ?

absolut elevation (m)
 elevation range (m) **1**
 no elevation
 elevation min:
 elevation max:

absolut depth (m)
 depth range (m) **2**
 no depth
 depth absolut:

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

Abbildung 14: Ausdehnung der Datenpublikation.

3.2.3.3.4 Dataset References (Referenzen)

Mit **Add Reference** (Referenz hinzufügen) **1** wird eine Zeile geöffnet, in die ganz links der **Value of the identifier** (Identifikator) **2** eingetragen werden kann (Abbildung 15). Das ist ein dauerhafter und eindeutiger digitaler Identifikator wie z.B. eine DOI oder ISSN. 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.

Dataset References ?

Add Reference **1**

Value of the identifier ?	Type ?	Relation ?	Label ?
[RELATED IDENTIFIER] 2	[relatedIdentifierType] 3	[relationType] 4	<input type="text"/> 5

Dataset Keywords ?

Add Keyword **6** Schlüsselwörter Empfehlung **7**

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

Abbildung 15: Möglichkeit zur Erfassung von Referenzen und die Angabe von Schlüsselwörtern.

3.2.3.3.5 Dataset Keywords (Schlüsselwörter)

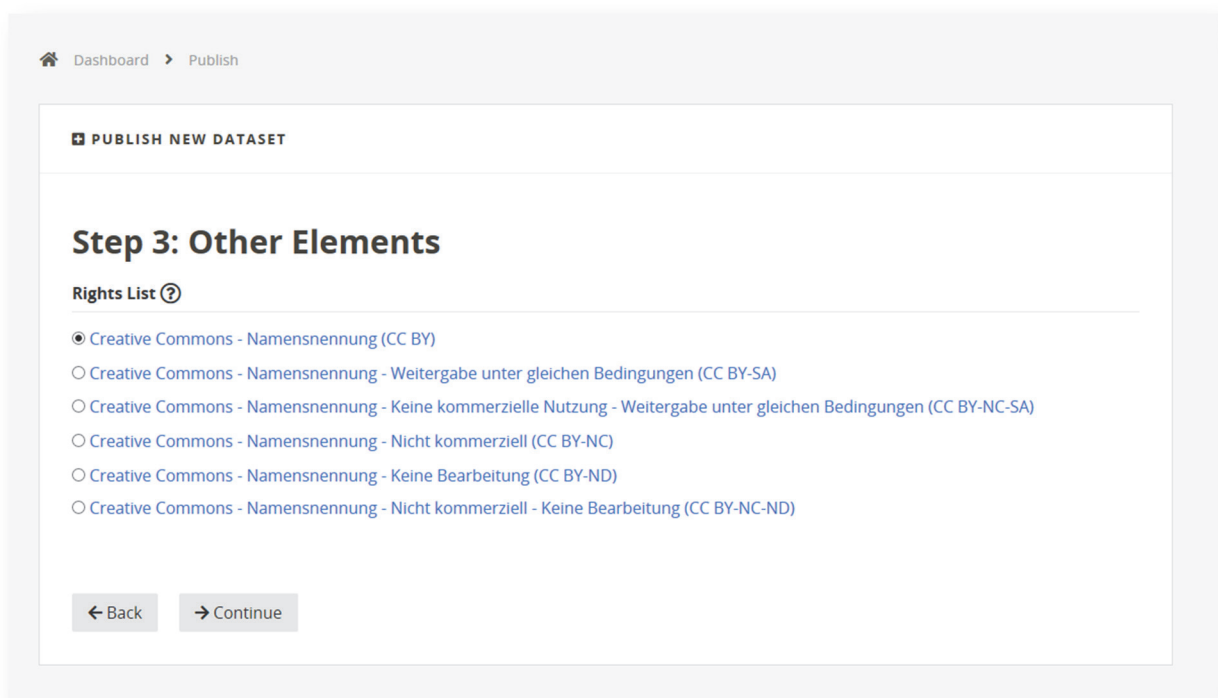
Es müssen insgesamt mindestens drei unterschiedliche Schlüsselwörter eingegeben werden. Idealerweise sind diese Schlüsselwörter nicht im Titel enthalten. Durch Klick auf **Add Keyword** **6** werden weitere Zeilen geöffnet. Es steht eine Schlüsselwörter-Empfehlung **7** von einem

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 (Abbildung 16). Bei Klick auf die jeweilige Nutzungsbestimmung wird die Seite <https://creativecommons.org> aufgerufen. Hier finden sich die Erläuterungen zu den jeweiligen Lizenzen.



Dashboard > Publish

PUBLISH NEW DATASET

Step 3: Other Elements

Rights List ?

- Creative Commons - Namensnennung (CC BY)
- Creative Commons - Namensnennung - Weitergabe unter gleichen Bedingungen (CC BY-SA)
- Creative Commons - Namensnennung - Keine kommerzielle Nutzung - Weitergabe unter gleichen Bedingungen (CC BY-NC-SA)
- Creative Commons - Namensnennung - Nicht kommerziell (CC BY-NC)
- Creative Commons - Namensnennung - Keine Bearbeitung (CC BY-ND)
- Creative Commons - Namensnennung - Nicht kommerziell - Keine Bearbeitung (CC BY-NC-ND)

Abbildung 16: Andere Angaben – Nutzungsbestimmungen.

3.2.3.5 File Upload (Datei hochladen)

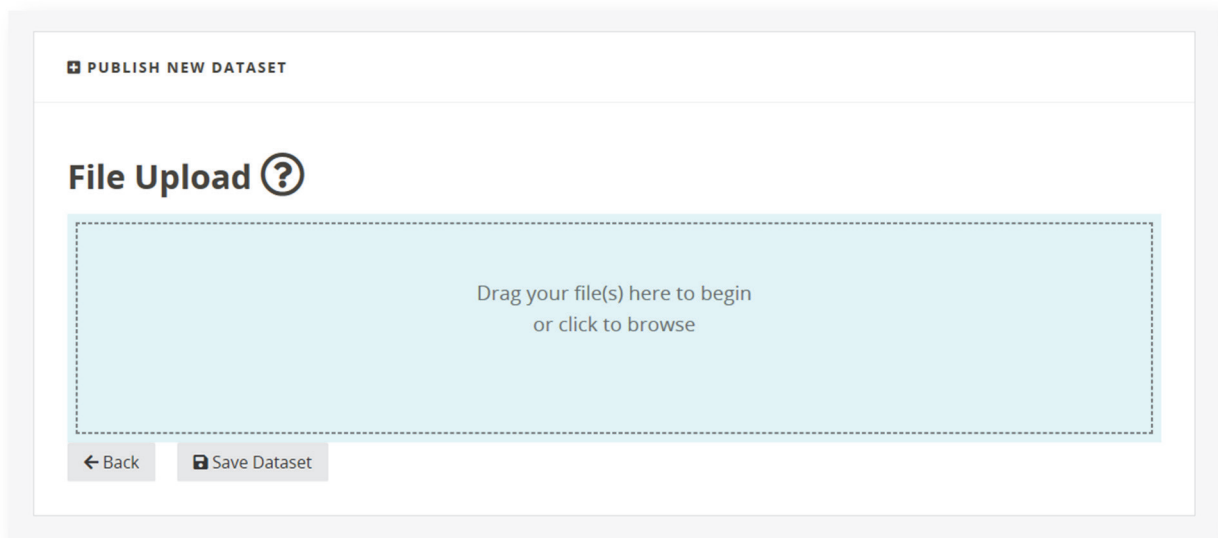


Abbildung 17: Hochladen der Datenpublikation 1.

Der Schritt **File Upload** (Datei hochladen) ist das eigentliche Herzstück der Applikation (Abbildung 17). Nun können die Daten der Datenpublikation hochgeladen werden.

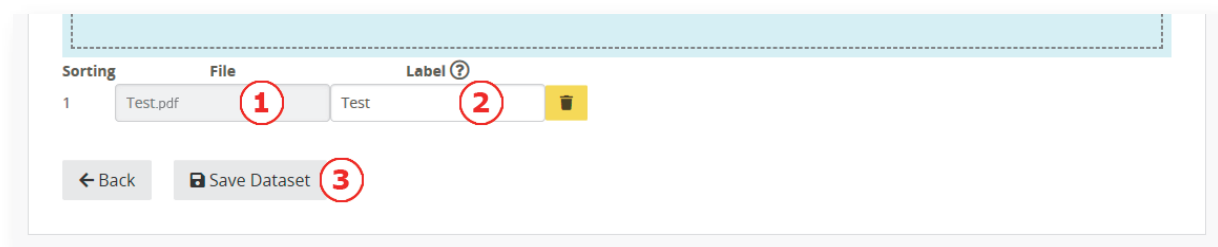


Abbildung 18: Hochladen der Datenpublikation 2.

In **File** (Datei) ① steht der Name der hochgeladenen Datei (Abbildung 18). 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.

3.2.3.6 Uploaded 1 file successfully (Hochladen erfolgreich)

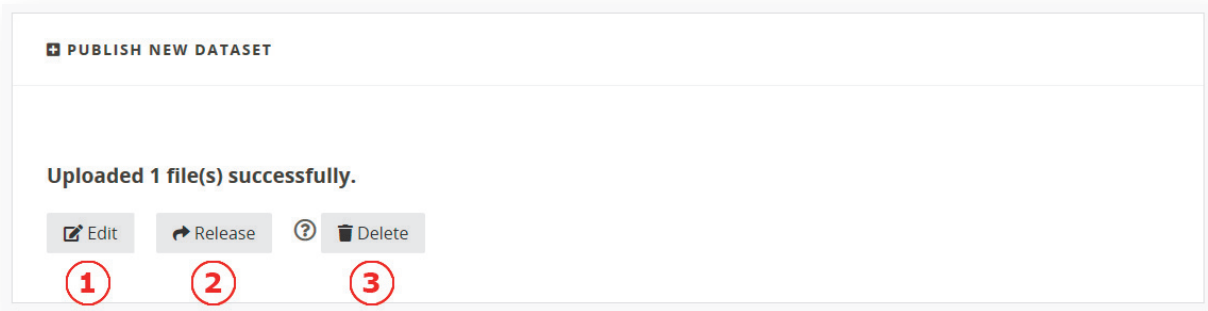


Abbildung 19: Edit, Release, Delete.

Nach erfolgreichem Save bietet das Tethys RDR folgende drei Möglichkeiten (Abbildung 19):

Die Metadaten können mit **Edit** (Bearbeiten) ① weiterbearbeitet werden. Die hochgeladene Datenpublikation kann allerdings **nicht** mehr gelöscht oder verändert werden.

Mit **Release** (Freigabe) ② werden die Metadaten und die Datenpublikation für den Editor freigegeben.

Mit **Delete** (Löschen) ③ werden der Uploadprozess gestoppt und die Inhalte gelöscht.

3.2.4 RELEASE YOUR DATASET FOR EDITOR (Freigabe der Datenpublikation für den Editor)

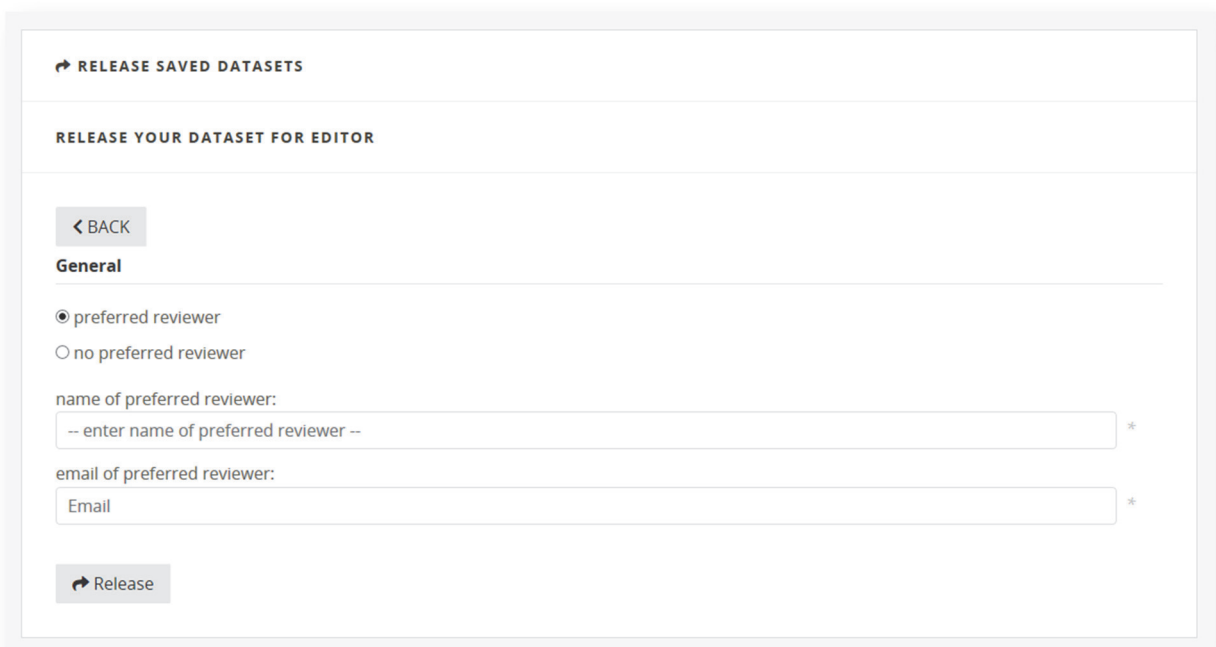
The screenshot shows a web interface for releasing saved datasets. At the top, it says 'RELEASE SAVED DATASETS'. Below that, the main heading is 'RELEASE YOUR DATASET FOR EDITOR'. There is a '< BACK' button. Under the heading 'General', there are two radio buttons: 'preferred reviewer' (which is selected) and 'no preferred reviewer'. Below these are two input fields: 'name of preferred reviewer:' with a placeholder '-- enter name of preferred reviewer --' and 'email of preferred reviewer:' with a placeholder 'Email'. Both input fields have an asterisk (*) indicating they are required. At the bottom, there is a 'Release' button with a right-pointing arrow icon.

Abbildung 20: Hier besteht die Möglichkeit, einen Reviewer auszuwählen.

Nach Klick auf **Release** besteht die theoretische Möglichkeit, einen bevorzugten **Reviewer** (Begutachter) auszuwählen (Abbildung 20). 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.

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)

RELEASE SAVED DATASETS bietet einen Überblick über die eigenen eingegebenen Datenpublikationen und deren aktuellen Status (Abbildung 21).

Es gibt mehrere Möglichkeiten, um zur Ansicht **Release saved datasets** zu gelangen.

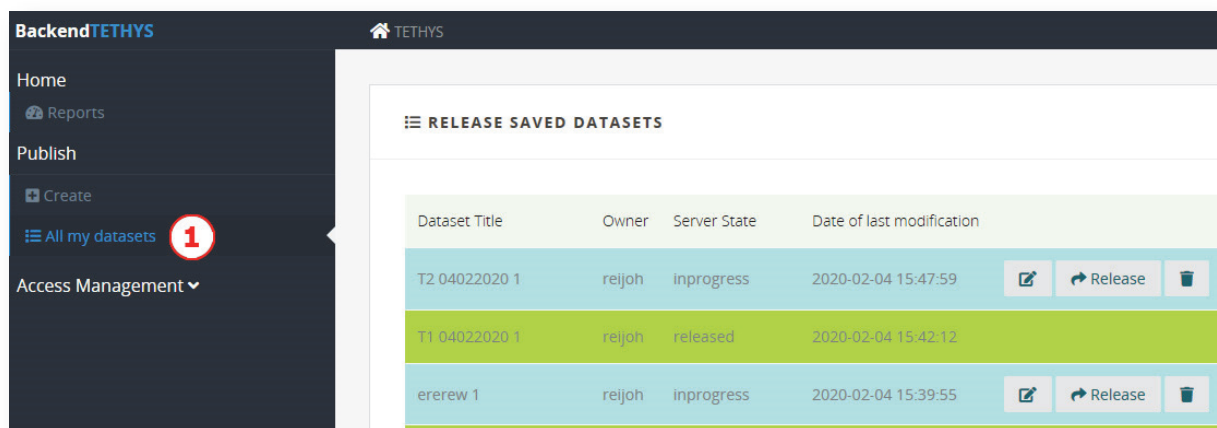


Abbildung 21: Übersicht über die eigene Datenpublikationen während des Publikationsprozesses.

Eine davon ist im linken Menübereich unter **All my datasets** ① auszuwählen. Die anderen Möglichkeiten werden in Tabelle 7 weiter unten angeführt.

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.
--	--	--	---

Tabelle 7: Release saved datasets.

In **RELEASE SAVED DATASETS** werden die eigenen Datensätze je nach Veröffentlichungsstadium in verschiedenen Farben angezeigt (Abbildung 22).

The screenshot shows a table titled "RELEASE SAVED DATASETS" with the following columns: Dataset Title, Owner, Server State, and Date of last modification. The rows are color-coded based on their state, and callouts 1-4 point to specific rows.

Dataset Title	Owner	Server State	Date of last modification	Callout
Test 5 29. November 1	reijoh	released	2020-01-31 13:01:09	1
erew 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	

Abbildung 22: Bedeutung der Farben bei „eigene Datenpublikationen“.

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 nachfolgend (Tabelle 8) einige relevante Begriffe, Kurzdefinitionen und dazugehörige Referenzen aufgeführt.

Wort	Kurzdefinition	Link/Referenzen
Datenrepositorium (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 et al. (2013): Forschungsdatenmanagement an Hochschulen: das Beispiel der Humboldt-Universität zu Berlin. LIBREAS. Library Ideas 23 (2013). urn:nbn:de:kobv:11-100212700
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/

Tabelle 8: Kurzdefinitionen zu spezifischen Begriffen.

Forschungsdaten publizieren ist ein internationales und sehr präsenten Thema in der Wissenschaftsgemeinschaft (Abbildung 23,Abbildung 24):

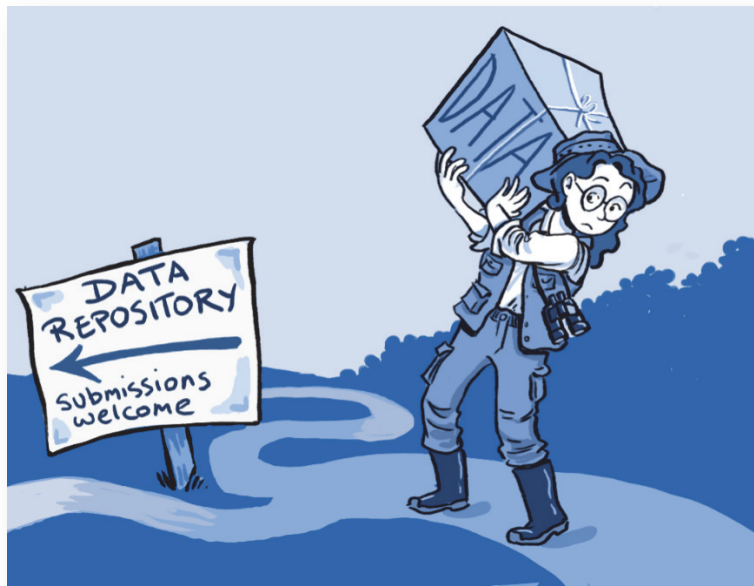


Abbildung 23: Forschungsdaten publizieren oder nicht publizieren, das ist hier die Frage! (Bild gezeichnet von Ainsley Seago.²⁵)

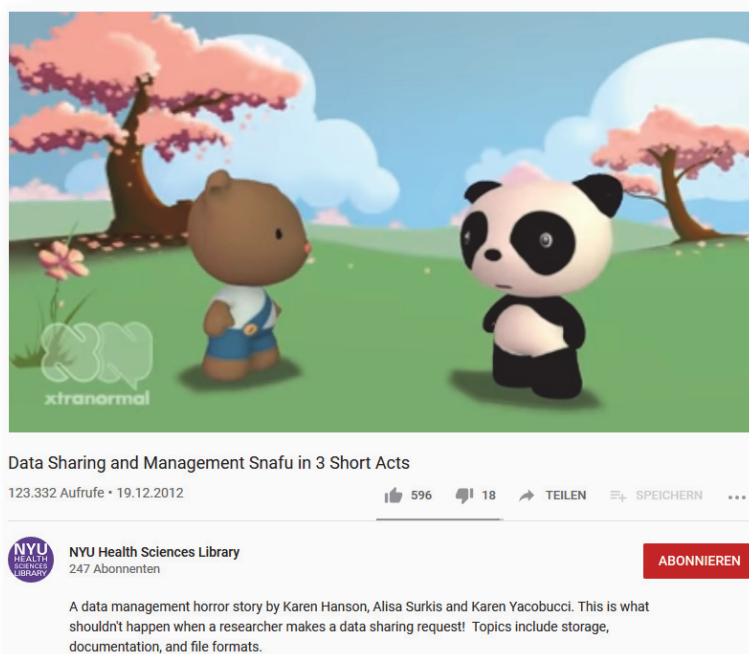


Abbildung 24: Shortfilm²⁶ (englisch) im Genre „Wissenschaftshorror“.

²⁵ ROCHE et al. (2014): Troubleshooting Public Data Archiving: Suggestions to Increase Participation. PLoS Biol 12(1): e1001779: <https://doi.org/10.1371/journal.pbio.1001779>; Illustration von Ainsley Seago: <https://doi.org/10.1371/journal.pbio.1001779.g001> (abgerufen am 24.01.2020)

²⁶ <https://www.youtube.com/watch?v=N2zK3sAtr-4&feature=youtu.be> (abgerufen am 12.02.2020)