

# Protokoll des Frühlings TIB DOI Konsortium online Workshop

Donnerstag, 11. Mai 2022, 10 - 11 Uhr

## Workshop Agenda

1. Begrüßung und neue Entwicklungen
2. TIB DOI Konsortium: Aktueller Stand
3. Best Practices Empfehlungen zum DOI-Metadatenfeld *relatedIdentifier*
4. Offene Feedback- und Fragerunde
5. Sonstiges

## Workshop Folien

<https://docs.google.com/presentation/d/1ZsOPCcvYoeuZhZfmLq8x6vKPUUnMy3lr1vnlT3hPMC/edit?usp=sharing>

## Linkempfehlungen

- DataCite Roadmap: <https://datacite.org/roadmap.html>
- Mitglieder des TIB DOI Konsortiums:  
<https://projects.tib.eu/pid-service/mitglieder-des-tib-doi-konsortiums/>
- Mailingliste: <https://www.listserv.dfn.de/sympa/subscribe/tib-doi-service>
- DataCite Fabrica Handbuch (Schnelleinstieg):  
<https://wiki.tib.eu/confluence/display/pid/DataCite+Fabrica+Handbuch+Startseite>
- Projekt OPTIMETA: <https://projects.tib.eu/optimeta/>
- Anmeldung zum Workshop "Journals in der OA-Transformation":  
<https://projects.tib.eu/koala/news/detail/anmeldung-zum-workshop-journals-in-der-oa-transformation/>

## Besprochene Fragen und vertiefte Themen

### Dokumentation Metadaten Schema

Anregungen zur Ergänzung/Änderung der Dokumentation des Metadaten Schemas können gerne an das Referat ([doi@tib.eu](mailto:doi@tib.eu)) oder über die [DataCite Roadmap](#) übermittelt werden.

### Use cases *relatedIdentifier*

1. Wird bei der Vergabe von *hasPart* die verlinkte Ressource automatisch um *IsPartOf* ergänzt oder ist das manuell durchzuführen?
  - a. Eine automatische Ergänzung ist bisher nicht möglich und erfordert zusätzlichen Kuratierungsaufwand bei den meisten *relationType*-Paaren.
2. Use Case 3: Über welche *relationTypes* können Versionen einer Ressource am besten abgebildet werden?

- a. Die Festlegung ist hier abhängig vom Ressourcentyp (Software, Zweitveröffentlichungen, Iterationen von Werken, Datenversionen) und eine Ermessensfrage. Hilfreich ist es in diesen Fällen, den Zweck des DOIs als unterstützendes Instrument zur Zitation zu betrachten und von diesem Zweck die Entscheidung für einen *relationType* abzuleiten, d.h. handelt es sich z. B. um veraltete Versionen, die nicht mehr genutzt werden können/sollten; um identische Zweitaufgaben; in Formatierung und Inhalt abweichende Preprints; marginal korrigierte Datensätze etc.?
3. Wie kann auf Synthese-Datensätze (aggregierte Daten aus mehreren anderen Datensätzen) referenziert werden?
  - a. Aggregationen können als Collection betrachtet werden, die einzelnen Datensätze könnten aber auch mit *IsPart/HasPart* mit einer DOI der Datensammlung verbunden werden. Wenn keine übergeordnete Instanz angelegt werden soll, wäre auch eine Verbindung mit *IsDerivedFrom* denkbar.
4. Wie kann die Beziehung zwischen Elementen von Zeitreihen ideal abgebildet werden, z. B. Elemente wie "Indicator\_A\_2021" und "Indicator\_A\_2022"?
  - a. Auch hier gibt es keinen klaren Favoriten: Möglich ist der Gebrauch von *hasPart/IsPartOf*, *Continues/IsContinuedBy*, je nachdem, wie bedeutsam die zeitliche Abfolge ist. Letzteres Paar wird häufig im Bereich Zeitschriftenfolgen genutzt, wird aber auch für die Abbildung der Chronologie von Events diskutiert.
5. Gibt es eine Option für einen "unspecific link"? Etwas, dass universell eingesetzt werden kann?
  - a. Es gibt keinen *relationType*, der zwei Ressourcen miteinander in Verbindung setzt, ohne diese Verbindung näher zu spezifizieren. In vielen Fällen kann *IsVersion/HasVersion* als Workaround dienen.
6. Wie kann die Versionierung von Daten abgebildet werden?
  - a. Pauschal ist das schwer zu beantworten, weil es auf die Differenzen der verschiedenen Versionen ankommt. Handelt es sich um eine Messdatenreihe, bei der Daten in verschiedenen Abständen gemessen werden, ist es sicher sinnvoll, diese Reihe mit *IsPartOf/HasPart* abzubilden, handelt es sich um Aktualisierungen, würden wir *IsNewVersionOf/IsPreviousVersionOf* empfehlen. *IsVersionOf/HasVersion* würden wir empfehlen, wenn die Unterschiede in den Daten im Aufbereitungsgrad oder der -form liegen.
7. Welchen *relationType* gebe ich an, wenn ich nicht genau weiß, ob das andere Dokument genau das selbe ist (*IsIdenticalTo*) oder ob es eine andere Form ist (*IsVariantFormOf*)?
  - a. Wenn nicht klar ist, ob es sich um ein identisches Dokument handelt, empfehlen wir *IsVariantFormOf*.

Die Anregung, dass bei der Einführung künftiger *relationTypes* die Nachvollziehbarkeit der Definition und ihrer Anwendbarkeit für die Nutzer:innen berücksichtigt werden sollte, wird an die DataCite Metadata Working Group weitergegeben und ist ein generelles Anliegen bei der Überarbeitung des Metadaten Schemas und seiner Dokumentation.

## Generelle Fragen

1. Unsere Verbindung zu TIB DOI läuft über OPUS - muss/sollte ich die Beziehungen dann händisch in DataCite Fabrica nacharbeiten?
  - a. Falls eine Konfiguration der OPUS-Version die Erfassung des *relatedIdentifier*-Properties nicht zulässt, müsste es über Fabrica bzw. eine andere API-Anbindung eingetragen werden.
2. Ist es notwendig/wichtig/wünschenswert, dass Datei-Metadaten (bspw. Verfasser, Keywords, CC-Lizenz, DOI) mit den Metadaten im Datensatz genau übereinstimmen, auch in der Form? Kommt es bei Nicht-Übereinstimmung zu Fehlermeldungen bei der OAI-Schnittstelle?
  - a. Es sollte nicht zu technischen Problemen kommen, allerdings steuern unterschiedliche Services die OAI-Schnittstelle bzw. die DataCite-API an. In Ihrem Interesse sollte daher die Übereinstimmung der Metadateninhalte möglichst groß sein.
3. Inwiefern wird es möglich sein, im Zuge der neuen Partnerschaft zwischen DataCite und IGSN (International Generic Sample Number für physische Proben) künftig eine Registrierung von IGSNs im Rahmen des TIB Konsortiums vornehmen zu können?
  - a. Informationen zur Integration sind im DataCite-Blogartikel zu finden: <https://blog.datacite.org/igsn-2040-vision-to-execution/> und in folgender Video-Präsentation: <https://www.youtube.com/watch?v=lgKpCYGoGBM>. DataCite plant mit der IGSN-Registrierung Ende des Jahres zu starten. Auch für Konsortialmitglieder wird die IGSN-Registrierung möglich sein. Wir werden Sie auf dem Laufenden halten.

## Ergebnisse der im Workshop durchgeführten Umfragen

### Umfrage zum *relatedIdentifier*

- 1. Kannten Sie das *relatedIdentifier*-Feld vor dem Workshop?

○ A. Kannte ich noch nicht.	12/75 ( 16%)
○ B. Ich habe schon einmal davon gehört.	14/75 ( 19%)
○ C. War mir bekannt.	20/75 ( 27%)
○ D. Nutzen wir bereits.	12/75 ( 16%)
○ Keine Antwort	19/75 ( 25%)
  
- 2. Werden Sie zukünftig das *relatedIdentifier*-Feld befüllen?

○ A. Nein, wir sehen nicht den Mehrwert.	0/75 ( 0%)
○ B. Nein, uns liegen keine Informationen vor.	1/75 ( 1%)
○ C. Nein, die Informationen sind unverhältnismäßig schwer zu erhalten.	4/75 ( 5%)
○ D. Ja, planen wir.	25/75 ( 33%)
○ E. Ja, werden wir weiterhin tun.	16/75 ( 21%)
○ Keine Antwort	30/75 ( 40%)
  
- 3. Was würde Ihnen helfen das Feld *relatedIdentifier* zu befüllen?

- A. Automatisierte Kontrolle und Ergänzung, wenn auf mein Objekt verwiesen wird. 28/75 ( 37%)
- B. Extraktion von Beziehungen aus dem Literaturverzeichnis. (Vgl. OPTIMETA) 8/75 ( 11%)
- C. Andere. (Bitte im Chat ausführen). 1/75 ( 1%)
- Keine Antwort 41/75 ( 55%)

## Feedback zum Workshop

- 1. Haben Sie inhaltlich etwas mitgenommen?
  - A. Ja, ich habe etwas für meine Arbeit gelernt. 46/68 ( 68%)
  - B. Ja, auch wenn es keinen unmittelbaren Bezug zu meiner Arbeit hat. 13/68 ( 19%)
  - C. Nein, mir war schon fast alles bekannt. 1/68 ( 1%)
  - D. Nein, die Inhalte waren für mich nicht relevant. 0/68 ( 0%)
  - Keine Antwort 10/68 ( 15%)
- 2. Was wünschen Sie sich für weitere Workshops?
  - A. Mehr praktische Beispiele. 42/68 ( 62%)
  - B. Mehr Hintergrundinformationen (Theorie). 20/68 ( 29%)
  - C. Mehr Austausch/Diskussionen. 11/68 ( 16%)
  - Keine Antwort 20/68 ( 29%)
- 3. Finden Sie die Länge des Workshops angemessen?
  - A. Ja, finde ich angemessen. 46/68 ( 68%)
  - B. Nein, ich fände kürzer besser. 0/68 ( 0%)
  - C. Nein, ich fände länger besser. 14/68 ( 21%)
  - Keine Antwort 10/68 ( 15%)