

DAVID BÖHM  
ALEXANDER GROSSMANN  
MICHAEL REICHE

SELINA SANNEMANN  
DIANA TILLMANN

# Akademischer Ansatz zum strukturierten Austausch von Inhalten und Metainformationen über die Herstellungs- und Distributionsprozesse von Open-Access-Publikationen

Forschungsbericht



DAVID BÖHM  
ALEXANDER GROSSMANN  
MICHAEL REICHE  
SELINA SANNEMANN  
DIANA TILLMANN

Akademischer Ansatz zum strukturierten Austausch von Inhalten und  
Metainformationen über die Herstellungs- und Distributionsprozesse von  
Open-Access-Publikationen

Forschungsbericht



DAVID BÖHM  
ALEXANDER GROSSMANN  
MICHAEL REICHE  
SELINA SANNEMANN  
DIANA TILLMANN

# Akademischer Ansatz zum strukturierten Austausch von Inhalten und Metainformationen über die Herstellungs- und Distributionsprozesse von Open-Access-Publikationen

Forschungsbericht

**HITWK**

Hochschule für Technik,  
Wirtschaft und Kultur Leipzig

## Über das Forschungsprojekt

Im Projekt OPEN-ACCESS-Strukturierte-Kommunikation (OA-STRUKTKOMM) wurden auf Grundlage des im Vorgängerprojekt OA-HVerlag entwickelten Open-Access-Publikationsworkflows für akademische Bücher Probleme der Standardisierung der Metadaten-Kommunikation innerhalb von (Open-Access-) Publikationsprozessen erforscht. Mit den Ergebnissen soll zum Aufbau eines offenen und integrierten (Open-Access-) Verlagsökosystems und damit zur Vereinfachung der Inhaltsverwaltung beigetragen werden.

Das Forschungsprojekt wurde vom 15. Februar 2021 bis 30. September 2023 an der HTWK Leipzig durchgeführt und durch das Bundesministerium für Bildung und Forschung (BMBF) gefördert.

Der Forschungsbericht wurde maßgeblich von Prof. Dr.-Ing. Michael Reiche und Diana Tillmann mit Beiträgen von Prof. Dr. rer. nat. Alexander Grossmann und David Böhm sowie unter Verwendung von Teilen der Masterarbeit von Selina Sannemann verfasst.

## Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie. Detaillierte bibliografische Daten sind im Internet unter <http://dnb.de> abrufbar.



Der Text dieses Werks ist unter der Creative-Commons-Lizenz CC BY 4.0 International veröffentlicht. Den Vertragstext der Lizenz finden Sie unter <https://creativecommons.org/licenses/by/4.0/>. Die Abbildungen sind von dieser Lizenz ausgenommen, hier liegt das Urheberrecht beim jeweiligen Rechteinhaber.



Die Online-Version dieser Publikation ist abrufbar unter <http://doi.org/10.33968/9783966270557-00>

© 2023 David Böhm, Alexander Grossmann, Michael Reiche, Selina Sannemann, Diana Tillmann

## Herausgeber

Open-Access-Hochschulverlag  
Hochschule für Technik, Wirtschaft und Kultur Leipzig  
Karl-Liebknecht-Str. 143  
04277 Leipzig, Deutschland

Druck & Bindung in Deutschland und den Niederlanden  
Gedruckt auf säurefreiem Papier

ISBN (Hardcover) 978-3-96627-056-4

ISBN (PDF) 978-3-96627-055-7

## Kurzfassung

Die Kommunikation von wissenschaftlichen Texten und deren Metadaten in Workflows zur Erstellung von Open-Access-Publikationen ist ein aktuelles Diskussionsthema in der Medienlandschaft und Gegenstand des vorliegenden Forschungsberichtes. Ausgehend von der Sachlage, dass im Feld mannigfaltige Standardisierungsansätze zu verschiedenen Teilabschnitten des Gesamtworkflows vorliegen, kann einerseits geschlussfolgert werden, dass nahezu alle Metadaterme bereits einen Standardisierungsprozess durchlaufen haben, was aber andererseits zu einer großen Heterogenität der vorliegenden Standards und Quasistandards geführt hat. Die Standards weisen bereits Schnittmengen auf, es existiert aber nicht derjenige Metadatenatz, der alle notwendigen Metadatenelemente enthält. Die Schnittstellenproblematiken der am Gesamtworkflow beteiligten Infrastrukturen würden durch einen solchen aber reduziert werden können. Der Forschungsbericht diskutiert zwei Lösungsansätze. Einerseits wird der Ansatz eines allgemeinen Metadatenstandards, der alle Metadatenelemente enthält, untersucht, andererseits wird als Vorzugslösung das Denkmodell eines virtuellen Superstandards vorgestellt, mit dem im Sinne eines Containerformats verschiedene Metadaten-Instanzen in einer Kommunikationslösung integriert werden können. Dieser akademische Ansatz soll als ein Keimpunkt einer Entwicklung einer dynamischen Kommunikationsstruktur verstanden werden, welche in einer Community aus Anwendern, wie Bibliotheken und Hochschulverlagen, mit Dienstleistern und akademischen Projektbeteiligten weiterentwickelt werden kann.

## Abstract

The communication of scientific texts and their metadata in workflows for the creation of Open Access publications is a current topic of discussion in the media landscape and the subject of this research report. Based on the fact that there are many different approaches to standardization in the field for different parts of the overall workflow, it can be concluded on the one hand that almost all metadata terms have already undergone a standardization process, but on the other hand that this has led to a great heterogeneity of the existing standards and quasi-standards. These already have intersections, but there is not one metadata set that contains all the necessary metadata elements. However, the interface problems of the infrastructures involved in the overall workflow could be reduced by such a set. The research report discusses two approaches to a solution. On the one hand, the approach of a general metadata standard that contains all metadata elements is examined; on the other hand, the concept of a virtual super-standard is presented as preferred solution, where various metadata instances can be integrated in a communication system utilizing a container format. This academic approach is to be understood as a nucleus for developing a dynamic communication structure, which can be further enhanced in a community of users, such as libraries and university publishers, with service providers and academic project participants.



# Inhaltsverzeichnis

---

|  |           |
|--|-----------|
| Abbildungsverzeichnis .....                                      | IX        |
| Tabellenverzeichnis .....  | XI        |
| Abkürzungsverzeichnis .....                                      | XII       |
| <b>1 Einleitung .....</b>  | <b>15</b> |
| 1.1 Hintergrund .....  | 16        |
| 1.2 Forschungsprojekt OA-STRUKTKOMM .....                        | 18        |
| <b>2 Theoretische Grundlagen .....</b>                           | <b>19</b> |
| 2.1 Kommunikation und Schnittstellen .....                       | 19        |
| 2.1.1 Kommunikationsmodelle .....                                | 20        |
| 2.1.2 Schnittstellen und deren Spezifik .....                    | 22        |
| 2.1.3 Interoperabilität .....                                    | 25        |
| 2.2 Metadatenstandards .....                                     | 29        |
| 2.2.1 Heterogenität von Standards .....                          | 29        |
| 2.2.2 Transformation .....                                       | 39        |
| 2.3 Offene Standards im wissenschaftlichen Publizieren .....     | 43        |
| 2.4 Anwendung und Analyse im Workflowmodell .....                | 45        |
| 2.4.1 Schnittstellen .....                                       | 45        |
| 2.4.2 Metadatenmapping .....                                     | 50        |
| <b>3 Einführung in die Kommunikationsstruktur .....</b>          | <b>53</b> |
| 3.1 Rahmenbedingungen für die Diskussion .....                   | 53        |
| 3.2 Diskussion möglicher Strategien .....                        | 54        |
| 3.2.1 Diskussion von S1 .....                                    | 55        |
| 3.2.2 Diskussion von S2 .....                                    | 57        |
| 3.2.3 Konklusion und Extraktion der Anforderungen .....          | 58        |
| <b>4 Akademischer Ansatz eines Containerformats .....</b>        | <b>63</b> |
| 4.1 Aufriss der Idee .....                                       | 63        |
| 4.2 Aufbau und Funktionsweise .....                              | 66        |
| 4.3 Anwendungsszenarien .....                                    | 69        |
| 4.3.1 Produktlebenszyklus .....                                  | 69        |
| 4.3.2 Herausziehen eines Metadatensatzes aus dem Container ..... | 70        |
| 4.3.3 Anlegen und Exportieren eines Containers .....             | 70        |

|  |            |
|--|------------|
| 4.3.4 Hinzufügen und Entfernen von Daten .....                                       | 70         |
| 4.3.5 Importieren eines Containers .....   | 70         |
| 4.3.6 Dateninkonsistenzen und Auflösung von Konflikten .....                         | 71         |
| 4.3.7 Benutzerinterface .....  | 71         |
| 4.3.8 Programminterface .....  | 71         |
| 4.3.9 Validierung .....  | 71         |
| 4.3.10 Optionale Anwendungsfälle .....   | 72         |
| 4.4 Beispielhafte Umsetzung .....  | 72         |
| 4.4.1 Vorschlag für die Struktur der Containerinstanz .....                          | 73         |
| 4.4.2 Validieren der Struktur(en) innerhalb eines Containers .....                   | 81         |
| 4.4.3 Anwendungsfall: Herausziehen eines Metadatensatzes aus dem Container .....     | 82         |
| 4.4.3.1 Generierung eines Formulars zur Unterstützung der DOI-Aktivierung .....      | 82         |
| 4.4.3.2 Erzeugung einer Upload-Instanz zur Registrierung von Peer-Review-Daten ..... | 84         |
| <b>5 Schlussbetrachtung .....</b>  | <b>85</b>  |
| Literaturverzeichnis .....   | 93         |
| <b>Anhang .....</b>  | <b>101</b> |
| <b>Anhang A Offene Standards im wissenschaftlichen Publizieren .....</b>             | <b>102</b> |
| <b>Anhang B Externe Kommunikationsteilnehmer .....</b>                               | <b>104</b> |
| <b>Anhang C Metadaten im OA-HVerlag .....</b>  | <b>106</b> |
| C.1 Metadaten einer Open-Access-Publikation im OA-HVerlag .....                      | 107        |
| C.2 Glossar der Metadaten einer Open-Access-Publikation im OA-HVerlag .....          | 108        |
| <b>Anhang D Transformationsregeln .....</b>  | <b>116</b> |
| <b>Anhang E Projektsteuerdatei projekt.xml .....</b>                                 | <b>128</b> |
| <b>Anhang F Generierung des Formulars zur DOI-Aktivierung .....</b>                  | <b>134</b> |
| F.1 Stylesheet .....   | 135        |
| F.2 Formular zur DOI-Aktivierung .....   | 141        |
| <b>Anhang G Ergebnisinstanz für die Registrierung von Peer-Review-Daten .....</b>    | <b>144</b> |

# Abbildungsverzeichnis

---

|                |   |    |
|----------------|---|----|
| <b>Abb. 1</b>  | Eigene Darstellung eines Kommunikationssystems nach Shannon und Weaver, Karl-Dieter Bunting, Karl-Heinz Wagner .....  | 21 |
| <b>Abb. 2</b>  | Bidirektionale Schnittstellen (in Anlehnung an Mara Kasubke) .....  | 23 |
| <b>Abb. 3</b>  | Spezifik von Schnittstellen .....   | 24 |
| <b>Abb. 4</b>  | Gliederung des Interoperabilitätsbegriffs (in Anlehnung an Peter Staub und opengeoedu.de) .....   | 26 |
| <b>Abb. 5</b>  | Operabilitätsgrade nach vereinfachter und semantischer Betrachtung (obere Grafiken in Anlehnung an Matthias Weiss) .....  | 27 |
| <b>Abb. 6</b>  | Eigene Darstellung der Arten von Heterogenität nach Leser/Naumann .....   | 30 |
| <b>Abb. 7</b>  | Darstellungsformen eines objektorientierten Datenmodells .....  | 31 |
| <b>Abb. 8</b>  | Darstellung eines RDF-Graphen nach Enoksson et al. ....   | 32 |
| <b>Abb. 9</b>  | MARC21 Record (links) und BITS-Metadaten (rechts) eines Buchtitels .....  | 33 |
| <b>Abb. 10</b> | Darstellung des Metadatum Autor-Name in Dublin Core (links) und BITS (rechts) .....   | 34 |
| <b>Abb. 11</b> | Übersetzung eines Metadatum höherer Granularität (links) in ein Metadatum niedriger Granularität (rechts) am Beispiel des Metadatum Autornamens in Quellschema BITS (links) zu Zielschema DC (rechts) .....               | 34 |
| <b>Abb. 12</b> | Übersetzung eines Metadatum geringerer Granularität (links) in ein Metadatum höherer Granularität (rechts) am Beispiel des Metadatum „Publikationsdatum“ im Quellschema MARC21 (links) zum Zielschema BITS (rechts) ..... | 35 |
| <b>Abb. 13</b> | Heterogenitäten in der Occurence und Obligation zwischen den Regelngen von BITS (links) und Crossref (rechts) zum Metadatum „Publikationsdatum“ ...   | 36 |
| <b>Abb. 14</b> | Darstellung möglicher Konstellationen im Vergleich der Eigenschaften zweier Standards .....   | 37 |
| <b>Abb. 15</b> | Beispielhafte Darstellung schematische Heterogenität hinsichtlich der Darstellung der Werte als Elemente, Kindelemente und Attribute (oben und hinsichtlich der Relation (unten) .....                                    | 37 |
| <b>Abb. 16</b> | Relation Autor-Bankverbindung .....   | 40 |
| <b>Abb. 17</b> | Mögliche Beziehungen zwischen Metadatenbeschreibungen verschiedener Standard .....  | 41 |
| <b>Abb. 18</b> | Poster „Überblick über offene Standards im wissenschaftlichen Publizieren“ ...  | 44 |
| <b>Abb. 19</b> | Beispiel einer Prozessschnittstelle im allgemeingültigen Workflowmodell .....   | 45 |
| <b>Abb. 20</b> | Darstellung der Schnittstellen-Spezifik beim Prozess P-EP-100: Produkt auf primärer Veröffentlichungsplattform veröffentlichen .....  | 47 |

|                |   |    |
|----------------|---|----|
| <b>Abb. 21</b> | Darstellung der Schnittstellen-Spezifika beim Prozess P-CiP-160: Freigabe vom Content-Urheber einholen .....                    | 48 |
| <b>Abb. 22</b> | Darstellung der Schnittstellen-Spezifika beim Prozess I-Qe-130: Peer-Review beauftragen .....                                   | 49 |
| <b>Abb. 23</b> | Grundsätzlicher Ansatz einer auf einer „Superstruktur“ basierenden Kommunikationsstruktur .....                                 | 64 |
| <b>Abb. 24</b> | Einsatz von Regeln bei der Verknüpfung semantischer Elemente .....  | 65 |
| <b>Abb. 25</b> | Aufbau der Kommunikationsstruktur als Containerformat .....   | 66 |
| <b>Abb. 26</b> | Produktlebenszyklus einer OA-Publikation .....  | 69 |
| <b>Abb. 27</b> | Einsatz von Validierungswerkzeugen zur Qualitätssicherung der Kommunikationsstruktur .....                                      | 72 |
| <b>Abb. 28</b> | Grundsätzliche Struktur eines Containers für die strukturierte Kommunikation von Daten und Metadaten für OA-Publikationen ..... | 73 |
| <b>Abb. 29</b> | Musterinstanzen im Content-Verzeichnis des Containers .....   | 74 |
| <b>Abb. 30</b> | Verzeichnis für nicht standardisierte Daten .....   | 74 |
| <b>Abb. 31</b> | Verzeichnis für Mediendaten .....   | 75 |
| <b>Abb. 32</b> | Verzeichnis für Stylesheets .....   | 75 |
| <b>Abb. 33</b> | Verzeichnis für DTDs und Schemata .....   | 76 |
| <b>Abb. 34</b> | Aufruf des Formatierungsstylesheets zur Generierung des DOI-Formulars .....   | 83 |
| <b>Abb. 35</b> | Ausschnitt aus dem DOI-Formular .....   | 84 |
| <b>Abb. 36</b> | Transformation zur Hinterlegung von Peer-Review-Daten .....   | 85 |

# Tabellenverzeichnis

---

|               |  |    |
|---------------|--|----|
| <b>Tab. 1</b> | Regeln für die Transformation von Elementen zwischen zwei standardisierten<br>Strukturen ..... | 42 |
|---------------|--|----|

## Abkürzungsverzeichnis

---

|         |  |
|---------|--|
| C       | Code   |
| DCF     | Design rule for Camera File system             |
| DIN     | Deutsches Institut für Normung                 |
| DTD     | Document Type Definitions                      |
| FOAF    | Friend of a Friend                             |
| HTML    | Hypertext Markup Language                      |
| I       | Input  |
| IOP     | Interoperabilität                              |
| ISO     | International Organization for Standardisation |
| K       | Kommunikationsteilnehmer                       |
| KIM     | Kompetenzzentrum Interoperable Metadaten       |
| O       | Output   |
| OA      | Open Access                                    |
| OAI-PMH | Open Archives Protocol for Metadata Harvesting |
| OPF     | Open Packaging Format                          |
| PDF     | Portable Document Format                       |
| RDF     | Resource Description Framework                 |
| URI     | Uniform Resource Identifier                    |
| URL     | Uniform Resource Locator                       |
| XSD     | XML Schema Definitions                         |





# 1 Einleitung

---

Kommunikation ist eine der wichtigsten Triebkräfte der menschlichen Entwicklung. Nach der Entwicklung der Sprachfähigkeit, die uns Menschen sogar physiologisch formte, trat mit der partiellen Sesshaftwerdung und der damit verbundenen Organisation des Menschen in größeren Gruppen die Notwendigkeit der Verschriftlichung von Sprache auf. Schrift und Schreiben wurden, wohl mehrfach und in verschiedenen Weltgegenden, erfunden. Seitdem wird nach Indizien gesucht, ob es eine frühere Ursprache gegeben haben könnte, die durch verschiedene Ursachen, seien es die babylonische Sprachverwirrung oder doch eher verschiedene Entwicklungsbedingungen, in verschiedene Sprachen unterschiedlichster Komplexität und Leistungsfähigkeit aufgespalten wurde. Seit historisch kurzer Zeit kommen zu den vielen natürlichen Sprachen, die auf der Welt gesprochen werden, die artifiziellen hinzu. Dazu gehören vor allem Programmiersprachen, Metamarkupsprachen bzw. die mit letzteren konstruierten konkreten Markupsprachen. Artifizielle Sprachen erweitern damit die Artikulationsmöglichkeiten des Menschen in den von ihm kontrollierten technischen Bereich. Eine der zu leistenden Aufgaben ist dabei die Formalisierung der dialogorientiert angelegten menschlichen Kommunikation mit vielen inhärenten Freiheitsgraden auf möglichst eindeutige Formulierungen, die für die artifiziellen Sprachen notwendigerweise mit eingeschränkten Freiheitsgraden verbunden ist.

Mit der menschlichen Artikulationsfähigkeit erschien auch der Drang zur Beschreibung der Welt, vorrangig als Versuch, deren Komplexität zu begegnen. Lange war man der Meinung, dass sich die Welt als Ganzes beschreiben ließe, wende man nur das richtige Ordnungsprinzip an. Dem Ziel, dieses zu finden, widmeten sich Generationen von Philosophen. Letztlich setzte sich die Erkenntnis durch, dass das immer progressiver wachsende Menschheitswissen keinen Zeitpunkt kennt, an dem es als vollständig erkannt (bzw. beschrieben) definiert werden kann. Dieser pessimistischen Erkenntnis setzen die Wissenschaften neue Methoden, wie Data Mining und KI, entgegen.

Gleichsam synonym zur Entwicklung der natürlichen Sprachen von einfachsten Artikulationen zu komplexesten Grammatiken vollzieht sich jene der artifiziellen parallel zur Leistungsfähigkeit der Rechentechnik. Lange war Speicherplatz rar, die Variablen der Programmiersprachen in ihrer Länge beschnitten und die Möglichkeiten von Markupsprachen beschränkt. SGML<sup>1</sup> wurde als mächtige Metamarkupsprache entwickelt, war aber auf der zeitgenössischen Technik nicht flächendeckend ausführbar. Erst das verschlankte XML konnte etwas später auf leistungsfähigeren Computern seinen

---

1 vgl. Library of Congress 2023

Durchbruch feiern. Gleiches gilt für die Entwicklung der konkreten Markupssprachen zur Beschreibung von Metadaten. Wurde 1994 im amerikanischen Dublin noch nach fünfzehn Termen gesucht<sup>2</sup>, um möglichst jede Ressource in einem elektronischen Metadatensatz beschreiben zu können, sehen wir uns heute, bei nahezu unbeschränkten Kapazitäten zur Datenspeicherung und -übertragung, einer unüberschaubaren Vielzahl von Metadatenstandards gegenüber. Hier geschah wohl in historisch kurzer Zeit, wofür die natürlichen Sprachen Jahrtausende benötigten: mit wachsenden Ressourcen und in unterschiedlichen Domains bildeten sich lokale Sprachen heraus, welche in dem, was sie beschreiben, einen hohen semantischen Überdeckungsgrad, aber auch technische Unterschiede aufweisen. Wie bei verschiedenen natürlichen Sprachen den Kommunikationsproblemen mit einem Dolmetscher abgeholfen werden kann, wird zwischen den artifiziiellen Sprachen mittels Transformationsschnittstellen kommuniziert. Aber, sowohl die Menschen als auch die Maschinen (bzw. die Ingenieure, welche diese Maschinen entwickeln), verspüren tief in ihrem Inneren die Sehnsucht nach dieser einen Sprache, welche die Kommunikation dramatisch vereinfachen, Missverständnisse vermeiden und den Aufwand reduzieren würde. Diese Sehnsucht kann ein Leitmotiv für neue Forschungsansätze sein.

## 1.1 Hintergrund

In der wissenschaftlichen Publikationslandschaft und in den unterschiedlichsten Bereichen der gesamten Wertschöpfungskette von wissenschaftlichen Werken wurde unter den beschriebenen Prämissen bereits eine Vielzahl existierender Standards, Normen und Spezifikationen entwickelt und in die Praxis überführt. Mit der fortschreitenden Digitalisierung und Automatisierung von Publikationsprozessen ist die Verwendung von Standards unumgänglich geworden, um Prozesse ökonomisch und effizient zu gestalten. Doch auch der Austausch und die Weiterverarbeitung von Informationen über Systemgrenzen hinweg macht die Verwendung von Standards notwendig.

Diese Notwendigkeit hat zu den verschiedensten Standardisierungskampagnen geführt. Wie sich die natürlichen Sprachen territorial entwickelt haben, geschah und geschieht dies auch für bestimmte abgegrenzte Bereiche, Domains, die sich hier nicht örtlich aber thematisch abgrenzen. Und wie die natürlichen Sprachen die Kommunikation nur dahingehend regeln, was die regionalen Bedingungen verlangen, werden die artifiziiellen Sprachen nur die Elemente enthalten, die dem zu vereinheitlichenden Zweck dienen. Dies führt zu verschiedenen Arten von Standards: technischen Standards, bibliographischen Standards, herstellerepezifischen Standards, um nur beim Herstellungs- und Publikationsworkflow von wissenschaftlichen Monografien zu bleiben. Und wie bei den natürlichen Sprachen lässt es sich nicht vermeiden, dass die entstandenen

---

2 vgl. Dublin Core 2023a

Standards idealerweise aufeinander aufbauen bzw. sich ergänzen, oft aber einander durchdringen und nur in Schnittmengen übereinstimmen. Zudem gilt für artifizielle wie für natürliche Sprachen, dass sie sich entwickeln und ihren „Duden“ regelmäßig überarbeiten. Notwendigerweise führt dies zu einer heterogenen Landschaft an Standards. Besonders im wissenschaftlichen Publizieren verstärkt sich die Diversität an Standards, da die zu publizierenden Werke aus den verschiedensten Wissenschaftsdisziplinen kommen und sich daher inhaltlich und strukturell stark voneinander unterscheiden. Dabei verfügt jede Wissenschaftsdisziplin über spezifische Besonderheiten und Anforderungen, sodass sich oftmals ein eigener Standard im jeweiligen Bereich entwickelt und etabliert hat. Die Komplexität wird zusätzlich gesteigert, da zwischen den verschiedenen Standards Schnittmengen bestehen, die beispielsweise dadurch entstehen, wenn sich mehrere Standards aus einem gemeinsamen „Mutter“-Standard herausgebildet haben und sich auf diesen beziehen. All dies ist nicht neu und soll auch nicht kritisiert werden. Vielmehr ist die Frage zu stellen, wie mit der gegebenen Situation umzugehen ist.

Die bestehenden Standards wurden von Fachleuten in den jeweiligen Domains erstellt und sind daher durch deren spezifische Anforderungen bestimmt. So existieren bibliografische Standards wie MARC, welche Werke auf der Metaebene beschreiben, Standards, welche die Werkstruktur vereinheitlichen, wie DocBook, JATS und BITS, sowie technische Standards, allen voran der Metamarkup-Standard<sup>3</sup> XML, mit welchem die meisten der anderen Standards konstruiert wurden. Da im Veröffentlichungsprozess sowohl technische, als auch bibliografische und herstellungsbezogene Daten verwendet werden, müssen Standards aus all diesen Domains verwendet werden. Wenn man zusätzlich davon ausgeht, dass werks- bzw. unternehmensspezifische Workflows für die Herstellung und Veröffentlichung durchlaufen werden, ist abzuleiten, dass sich offensichtlich projektabhängig verschiedene Sichten auf die zu verwendenden Schnittmengen der benutzten Standards ergeben, die dann nur einen jeweiligen Teilausschnitt aus dem gesamten Sprachumfang abbilden. Letztendlich sind die erwähnten Sichten projekt- bzw. workflowspezifische Sprachen, die sich aus Teilmengen anderer Sprachen, bzw. Standards, zusammensetzen. Das Problem der Konstruktion dieser Sprachen kann mit unterschiedlichen Strategien angegangen werden. Die Untersuchung und Herausarbeitung der Strategien ist Gegenstand des Forschungsberichtes.

---

3 Zumindest im Fall von XML ist der Begriff Standard nicht korrekt. Richtigerweise werden vom W3C Recommendations veröffentlicht. Diese erhalten qua Akzeptanz den Rang von Quasi-Standards. Daher soll der kritische Begriff trotzdem weiterverwendet werden.

## 1.2 Forschungsprojekt OA-STRUKTKOMM

Die Entwicklung eines akademischen Ansatzes zum strukturierten Austausch von Daten und Informationen innerhalb von Herstellungsprozessen wissenschaftlicher Werke im Open Access (OA) stellte ein wesentliches Forschungsziel des Forschungsprojektes *OA-STRUKTKOMM* (*Open-Access-Strukturierte-Kommunikation*)<sup>4</sup> dar. Grundlage für die Forschungen war der Open-Access-Publikationsworkflow für akademische Bücher<sup>5</sup>, welcher durch das 2020 abgeschlossene Forschungsprojekt *OA-HVerlag*<sup>6</sup> entwickelt und veröffentlicht wurde. Mit dem Workflowmodell wurden alle Prozesse, die in den verschiedensten Szenarien von OA-Publikationsprozessen auftreten können identifiziert, abstrahiert und beschrieben, und ermöglichen somit die Modellierung individueller Workflowmodelle. Innerhalb dieser Modelle sind einzelne Prozesse über Schnittstellen zum Informations- und Datenaustausch miteinander verbunden und konsumieren Daten aus vorhergehenden Prozessen und geben diese (durch Prozesse veränderte Daten) an nachfolgende weiter. Mit Hilfe des Workflowmodells kann ein qualitatives Datenflussmodell formuliert und dieses als Grundlage für die Erarbeitung konkreter Datenmodelle genutzt werden. Aktuell ist dieser Schritt noch weitestgehend durch individuelle Abstimmung zwischen einzelnen Anbietern von den Systemen im Workflowmodell abhängig. Da in der Regel aber die gleichen Daten ausgetauscht werden und einzelne dieser Datensegmente bereits standardisiert sind, liegt es nahe, für Open-Access-Publikationsworkflows und darüber hinaus ein Schema für die Kommunikation zwischen Publikationssystemen zu entwickeln. Auf Grundlage des Workflowmodells beschäftigte sich das Forschungsteam daher unter anderem mit der Frage, wie der Datenaustausch zwischen den workflowbeteiligten Systemen robuster und kompatibel gestaltet werden kann und wie OA-Publikationsinfrastrukturen durch zusätzliche Bausteine unterstützt und damit weiter vorangebracht werden können.

Neben dem Schwerpunkt der Entwicklung eines Ansatzes zur strukturierten Kommunikation verfolgte das Forschungsprojekt weitere Ziele: Der Wissenstransfer und die Etablierung des entwickelten Workflowmodells sowie die Weiterentwicklung eines aus dem Vorgängerprojekt hervorgegangenen Produktionssystems, welches die Möglichkeit bietet ohne spezifisches herstellerisches Wissen OA-Publikation herzustellen. Der Forschungsbericht konzentriert sich ausschließlich auf die Entwicklung eines akademischen Ansatzes einer Kommunikationsstruktur. Die Ergebnisse hier nicht betrachteter Forschungsschwerpunkte sind der Projektwebsite zu entnehmen.

---

4 OA-STRUKTKOMM 2023

5 Böhm et al. 2020

6 HTWK FIM 2023

## 2 Theoretische Grundlagen

---

Um die Kommunikation zwischen den am Publikationsprozess beteiligten Systemen zu strukturieren und sich damit einer Standardisierung von OA-Publikationsinfrastrukturen anzunähern, wurde sich mit den Fragen beschäftigt, welche offenen, technischen Standards bei der Publikation wissenschaftlicher Werke zur Anwendung kommen und wie sich bestehende Standards kategorisieren lassen. Darüber hinaus war die Auseinandersetzung und Beschreibung von Kommunikationsprozessen an entsprechenden Schnittstellen im wissenschaftlichen Publikationsprozess für die Entwicklung eines Ansatzes einer Kommunikationsstruktur von wesentlicher Bedeutung.

Mit der Basis des Workflowmodells sowie der theoretischen Grundlagen wurden die Integrationsmöglichkeiten und -potentiale von Daten und Standards im Kontext des zugrundeliegenden OA-Publikationsworkflow untersucht. Dabei wurde den Fragen nachgegangen, welche kommunizierenden Daten im Workflowmodell mit welchen (technischen) Systemen ausgetauscht werden und welche Prozesse daran beteiligt sind.

### 2.1 Kommunikation und Schnittstellen

Bei der Publikation wissenschaftlicher Werke im Open Access werden eine Vielzahl verschiedener Prozesse von unterschiedlichen Akteuren im Produktionsprozess realisiert. Durch das Zusammenwirken und die Kommunikation aller Beteiligten an den auszuführenden Prozessen bilden sich diverse Schnittstellen aus. An diesen Schnittstellen kann eine standardisierte Kommunikationsstruktur den Austausch von Informationen standardisierter und leistungsfähiger organisieren.

Die menschliche Sinneswahrnehmung zeigt auf ganz natürliche Weise das Vorhandensein von Schnittstellen, die es uns letztlich ermöglichen, Erfahrungen außerhalb unseres Körpers wahrzunehmen, um mit der Außenwelt zu kommunizieren und interagieren. Dabei bilden unsere Sinnesorgane diejenigen Verbindungsstellen, an der Reize (Informationen) aus der Umwelt übertragen und zur Weiterverarbeitung an das jeweilige gekoppelte Sinnessystem weitergeleitet werden.<sup>7</sup> Schnittstellen finden sich also überall dort, wo Informationen zwischen zwei oder mehreren Systemen, welche der Kommunikation dienen, übertragen und/oder ausgetauscht werden. Voraussetzung für das Vorhandensein von Schnittstellen ist somit in gewisser Hinsicht ein Kommunikationsprozess.

---

7 vgl. Kasubke 2020

### 2.1.1 Kommunikationsmodelle

Seit seiner Betrachtung und Beschreibung unterliegt der Kommunikationsbegriff einem stetigen Wandel. In den verschiedenen Disziplinen wird Kommunikation (lat. communication „Mitteilung“, communicare „teilhaben“, communis „gemeinsam“)<sup>8</sup> aus ganz unterschiedlichen Perspektiven betrachtet. Daher findet sich in der Literatur eine Vielzahl an Definitionen. Betrachtet man das grundlegende Prinzip der Informationsübertragung, so übermittelt ein Sender/Quelle (Person, System) eine Information (Reiz, Signal) über einen Kanal (Medium) an einen Empfänger/Senke (Person, System). Die Grundlage für das Verständnis zwischen Sender und Empfänger und somit einer möglichst verlustfreien Übertragung aller Informationen (eigentlich Signalen) bildet eine gemeinsame allgemeine Symbolik ( $A \wedge B$ ). Diese spezifiziert beispielsweise durch Sprache, Schrift, Mimik oder Gestik das zu übermittelnde Signal. In welche Richtung eine Kommunikation verläuft, kann durch die Direktionalität beschrieben werden: Bei einer unidirektionalen Kommunikation erfolgt die Signalübertragung in nur eine Richtung, d. h. ausschließlich eine Person oder ein System sendet Informationen an das Gegenüber, wobei kein Rückkanal vorgesehen ist, um auf die Information zu reagieren. Bei der bidirektionalen Kommunikation hingegen verläuft der Austausch in beide Richtungen, sodass Sender auch Empfänger und umgekehrt sein können. Die zeitliche Abfolge eines Kommunikationsprozesses wird durch die Synchronität konkretisiert. Dabei wird zwischen einer synchronen Kommunikation, bei der die Informationsübertragung zeitgleich stattfindet, und einer asynchronen Kommunikation, bei der es sich um eine zeitversetzte Kommunikation handelt, indem Sender und Empfänger zeitlich voneinander entkoppelt sind, unterschieden.<sup>9</sup> Konsequenterweise muss auch die Lokalität einer Kommunikation beachtet werden, die beschreibt, ob Sender und Empfänger an einem Ort lokalisiert sind oder ob Kanäle etabliert werden müssen, welche Entfernungen überbrücken, mit zusätzlichen Einflüssen, wie Zeitverzögerungen oder Dämpfung.

Ausgehend vom beschriebenen Grundprinzip der Informationsübertragung soll für die weitere Betrachtung des Kommunikationsprozesses das *Sender-Empfänger-Modell* von Shannon und Weaver<sup>10</sup> herangezogen werden. Das Modell entstand innerhalb der Entwicklung der Informationstheorie in der Mitte des 20. Jahrhunderts und viele weitere Kommunikationsmodelle bauen auf dem von Shannon und Weaver entwickelten Prinzip auf. Dieses ermöglicht eine mathematische Beschreibung von Informationen und fokussiert dabei explizit technische Signale sowie deren Übertragung, Verarbeitung und

---

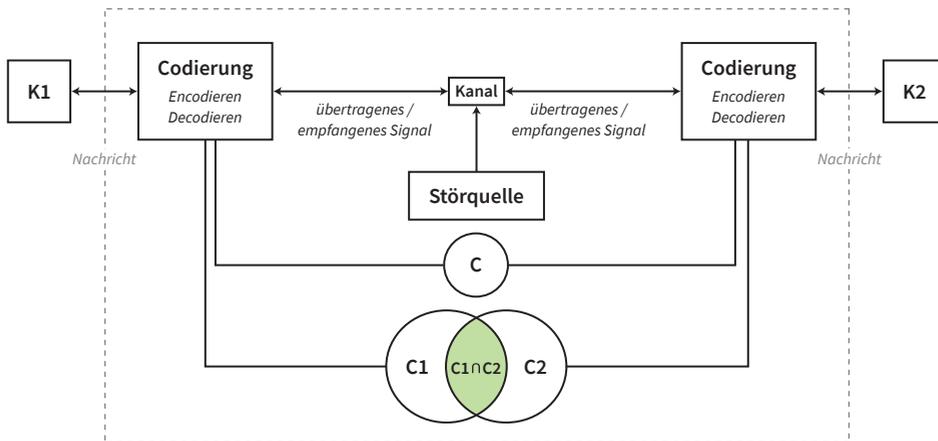
8 Bühler et al. 2017, S. 2

9 vgl. Abts et al. 2009, S. 288

10 Darstellung eines allgemeinen Kommunikationssystems nach Shannon und Weaver in: Shannon et al. 1964, S. 7

Empfang.<sup>11</sup> Ein Kommunikationsprozess nach dem Sender-Empfänger-Modell umfasst folgende wesentliche Bestandteile: An erster Stelle steht die Informationsquelle, welche eine Nachricht generiert. Diese wird anschließend an den Sender übertragen, welcher die Information in ein dem Kanal entsprechendes Signal umwandelt (encodiert). Der *Empfänger* muss das über den Kanal übermittelte Signal entschlüsseln (decodieren) und erhält somit die eigentliche Information.<sup>12</sup> Störquellen sollen dabei möglichst identifiziert und beseitigt werden, um Informationen verlustfrei zu übertragen.

Um das *Shannon-Weaver-Modell* mit seiner ausschließlichen Betrachtung der unidirektionalen Informationsübertragung weiter zu denken und darüber hinaus semantische Aspekte einzubeziehen, werden zusätzlich Modelle aus der linguistischen Perspektive herangezogen (Abb. 1). Informationsquelle und -ziel fungieren bei einer bidirektionalen Kommunikation sowohl als sendendes als auch als empfangendes System und sind somit in der Lage Informationen zu en- und zu decodieren. Im Folgenden werden diese als *Kommunikationsteilnehmer (K)* bezeichnet und verfügen über einen gemeinsamen *Code (C)* (allgemeine Symbolik). Die Kommunikationsteilnehmer K1 und K2 können dabei über den gleichen Code verfügen.



**Abb. 1** Eigene Darstellung eines Kommunikationssystems nach Shannon und Weaver<sup>13</sup>, Karl-Dieter Bunting<sup>14</sup>, Karl-Heinz Wagner<sup>15</sup>

11 vgl. Hunscha 2003, S. 6; vgl. Röhner et al. 2016, S. 21

12 vgl. ebd.; vgl. Hunscha 2003, S. 6f.

13 in Anlehnung an Shannon et al. 1964

14 in Anlehnung an Bunting 1990

15 in Anlehnung an Wagner 1997

Verfügen K1 und K2 über verschiedene Codes C1 und C2, so müssen Letztere eine Schnittmenge bilden, damit im Anschluss an die Decodierung des Signals, dessen Inhalte auch entsprechend interpretiert werden können. Eine korrekte Interpretation (Informationsgenerierung) ist nur innerhalb der Menge der Codes in der Schnittmenge möglich. Ein gemeinsamer Code bildet somit die Grundlage und notwendige Voraussetzung für das Verständnis zwischen den Kommunikationsteilnehmern.

### 2.1.2 Schnittstellen und deren Spezifik

Wie eingangs beschrieben, finden sich in jedem Kommunikationsprozess, also überall dort, wo Informationen zwischen zwei oder mehreren Systemen übertragen, ausgetauscht und verarbeitet werden, Schnittstellen. Der Medien- und Kommunikationswissenschaftler Wulf R. Halbach formulierte in *Interfaces. Medien- und kommunikationstheoretische Elemente einer Interface-Theorie*: „Um Interaktion mit einem Medium zu ermöglichen, wird eine vermittelnde Instanz benötigt, die nicht nur die Wechselseitigkeit des Nachrichtenaustausches, sondern auch des adäquaten - d. h. für die gekoppelten Systeme verstehbare - Kodierung und Dekodierung garantiert, damit „Sinn“ attribuiert werden kann. Diese „vermittelnde“ Instanz nennt man „Interface“ oder „Schnittstelle“.“<sup>16</sup> Im Folgenden soll für die Arbeitsdefinition des Begriffes *Schnittstelle* Halbach herangezogen werden:

*„Der Begriff Schnittstelle – oder englisch Interface – bezeichnet grundsätzlich den Punkt einer Begegnung oder einer Kopplung zwischen zwei oder mehr Systemen und/oder deren Grenzen zueinander.“<sup>17</sup>*

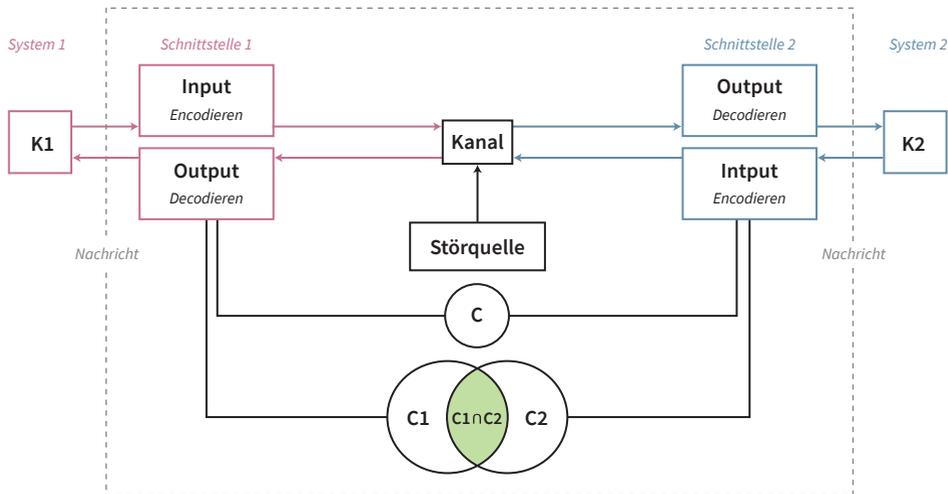
In welche Richtung die Schnittstelle dabei verläuft, wird durch die Uni- und Bidirektionalität (letztere ist beispielhaft in Abb. 2 dargestellt), die zeitliche Abfolge durch die Synchronität beschrieben. Eine weitere Charakterisierung von Schnittstellen betrifft die Verortung. Bei dieser befindet sich ein interner Begegnungspunkt innerhalb eines Systems, wohingegen eine externe Schnittstelle außerhalb des eigenen Systems zu verorten ist. Eine allgemeine Symbolik (Code) ist ebenfalls von Bedeutung (siehe Abschnitt 2.1.1). Dabei übernimmt laut Halbach eine Schnittstelle „Als technische Einrichtung [...] die Übersetzungs- und Vermittlungsfunktion zwischen gekoppelten Systemen.“<sup>18</sup> Bei dieser Kopplung soll es sich in der weiteren Betrachtung nicht zwangsweise um reine technische Systeme handeln. Auch Menschen fungieren in Publikationsprozessen als Schnittstellen.

---

16 Halbach 1994, S. 14

17 ebd., S. 168

18 ebd., S. 168



**Abb. 2** Bidirektionale Schnittstellen (in Anlehnung an Mara Kasubke<sup>19</sup>)

Mit Verweis auf die eingangs erwähnte menschliche Sinneswahrnehmung stellen demnach die Sinnesorgane den Punkt der Begegnung bzw. Koppelung zwischen dem Umwelt- und dem körpereigenen Sinnessystem dar.<sup>20</sup> Neben dieser Mensch-Außenwelt-Schnittstelle kann zwischen weiteren grundlegenden Typen natürlicher und künstlicher Schnittstellen unterschieden werden:

- Mensch-Mensch-Schnittstelle: Menschen kommunizieren miteinander
- Mensch-Maschine-Schnittstelle: Mensch und Maschine kommunizieren miteinander
- Maschine-Maschine-Schnittstelle: Maschinen kommunizieren miteinander

In Zusammenhang mit der zu entwickelnden Kommunikationsstruktur sind vor allem die Mensch-Mensch- und die Mensch-Maschine-Schnittstelle von Bedeutung. Um diese grundlegenden Typen von Schnittstellen weiter zu präzisieren, erfolgt deren Spezifizierung mit Blick auf den Publikationsprozess wissenschaftlicher Veröffentlichungen durch die Forschenden als Open-Access-Publikation. Um Schnittstellen nach allen genannten Kriterien zu prüfen, erscheint die Verwendung von vier W-Fragen praktikabel, welche in Abbildung 3 skizziert sind.

<sup>19</sup> vgl. Kasubke 2020

<sup>20</sup> vgl. ebd.

**WER** kommuniziert miteinander?

Mensch-Mensch  
Mensch-Maschine  
Maschine-Maschine



Typen/Arten von Schnittstellen

**WIE** wird kommuniziert?

asynchron  
synchron  
bidirektional  
unidirektional



Eigenschaften von Schnittstellen

**WO** wird kommuniziert?

intern  
extern



Ortsgebundenheit von Schnittstellen

**WAS** wird kommuniziert?

organisatorische Kommunikation  
technische Kommunikation



Infrastruktur von Schnittstellen

**Abb. 3** Spezifik von Schnittstellen

Die Frage „WER kommuniziert miteinander?“ gibt Aufschluss über die Typen der Schnittstellen und damit auch über alle beteiligten Kommunikationsteilnehmer. Eigenschaften von Schnittstellen und die zugrunde liegenden Kommunikationsprozesse werden über die WIE-Frage definiert. Die WO-Frage beschreibt die Ortsgebundenheit einer Schnittstelle und gibt an, ob sich diese inner- und/oder außerhalb des jeweiligen Systems befindet. Darüber hinaus findet die Infrastruktur von Schnittstellen Beachtung und vermittelt, WAS kommuniziert wird. Dabei wird zwischen organisatorischen Schnittstellen, welche den Austausch von Informationen umfassen, und technischen Schnittstellen, welche den Austausch von Daten beschreiben, unterschieden. Eine strikte Trennung ist dabei schwierig, da in vielen Prozessen parallel sowohl eine organisatorische als auch eine technische Kommunikation stattfindet, zwischen denen eindeutig zu differenzieren problematisch ist.

### 2.1.3 Interoperabilität

Ein wesentliches Kriterium für das nahtlose Zusammenwirken unterschiedlicher (technischer) Systeme ist die Interoperabilität (IOP) (lat. *opus* ‚Arbeit‘ und *inter* ‚zwischen‘).<sup>21</sup> Weltweit setzen sich verschiedenste Akteure in den unterschiedlichsten Disziplinen mit IOP auseinander, da diese auch nicht zuletzt durch Digitalisierungsprozesse von immer größerer Bedeutung ist. Unzählige Definitionen von Interoperabilität führen zu einer unüberschaubaren Anzahl an Begriffsbestimmungen. Eine Vielzahl bezieht sich ausschließlich auf die Zusammenarbeit technischer Systeme. Unter Berücksichtigung der Entwicklung einer Kommunikationsstruktur für das Publizieren wissenschaftlicher Literatur im Open-Access-Umfeld sind neben technischen Systemen auch Menschen und Organisationen in Produktionsprozesse involviert. Um sich einer Begriffsbestimmung anzunähern, entwickelten die Autoren eine projektspezifische Arbeitsdefinition des Begriffes Interoperabilität im Kontext des wissenschaftlichen Publizierens.

*Der Begriff Interoperabilität beschreibt die Fähigkeit von am wissenschaftlichen Publikationsprozess beteiligten Menschen, Institutionen und Systemen, unabhängig voneinander zu interagieren und unter Verwendung von offenen Standards so miteinander zu kommunizieren, dass Daten und Informationen effizient, fehler- und verlustfrei ausgetauscht werden können.*

Innerhalb dieser Begriffsbestimmung wird postuliert, dass die Interaktion zwischen Menschen, Organisationen, Geschäftsprozessen sowie verschiedenen Informations- und Kommunikationssystemen stattfindet. Auf diese Weise soll eine nahtlose, einheitliche und effiziente Zusammenarbeit über mehrere Systeme und Organisationen hinweg bereitgestellt werden. Die Herausgeber des Werkes *Enabling Multilingual Social Interactions and Fostering Language Learning in Virtual Worlds* formulieren ebenfalls, dass Lösungen zur IOP „[...] stets menschliche Operateure, Informationssysteme, eine Infrastruktur zum Austausch von Information und schließlich auszutauschende Information.“<sup>22</sup> beinhalten.

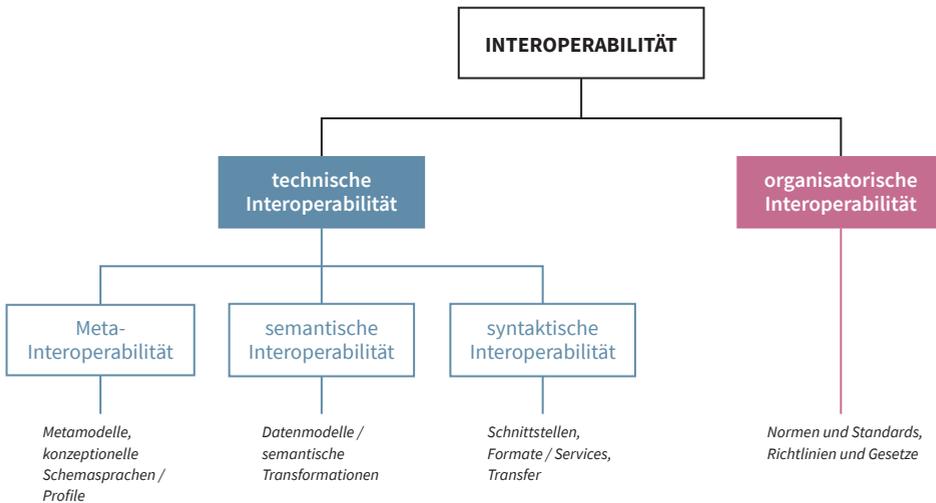
Weiterhin wird der Begriff Interoperabilität in der Literatur oft in verschiedene Ebenen aufgegliedert, jedoch findet sich hierfür keine allgemeingültige Einteilung. Daher wird die Hierarchisierung von Staub herangezogen, welcher Interoperabilität in die technische und die organisatorische IOP untergliedert (Abb. 4). Dabei teilt er die technische Interoperabilität in Meta-, semantische sowie syntaktische Interoperabilität ein und erfasst innerhalb dieser „[...] konkrete Konzepte und Werkzeuge, welche

---

21 vgl. Duden 2021

22 Bau et al.

die entsprechenden Teilaspekte der Interoperabilität realisieren [...]“<sup>23</sup>. Die syntaktische IOP beschreibt interagierende Systeme, welche in der Lage sind, Datenstrukturen zu identifizieren, zu verarbeiten und auszutauschen. Semantische IOP interpretiert Informationen und Datenstrukturen. Innerhalb der organisatorischen Interoperabilität werden Prozesse durch Workflows und Rollen beteiligter Akteure, aber auch bspw. durch Richtlinien, Gesetze, Normen (z.B. DIN, ISO) und Standards (z.B. W3C), effektiv und effizient gestaltet.<sup>24</sup>



**Abb. 4** Gliederung des Interoperabilitätsbegriffs (in Anlehnung an Peter Staub<sup>25</sup> und opengeoedu.de<sup>26</sup>)

Diese Gliederung des Interoperabilitätsbegriffes ist bei der Verwendbarkeit offener Daten von wesentlicher Bedeutung: „Diese [die offenen Daten, OA-STRUKTKOMM] sollten idealerweise mittels allgemein verfügbarer Software verarbeitet (syntaktische Interoperabilität) und interpretiert (semantische Interoperabilität) werden können, ohne dass dies etwa durch die Verwendung proprietärer, [...] nicht-standardisierter Datenformate und Schnittstellen erschwert wird (organisatorische Interoperabilität)“<sup>27</sup>.

Da die gesamte Bandbreite des Publikationsprozesses wissenschaftlicher Werke durch die Vernetzung und das Zusammenwirken verschiedenster Systeme gekenn-

23 Staub 2009, S. 22

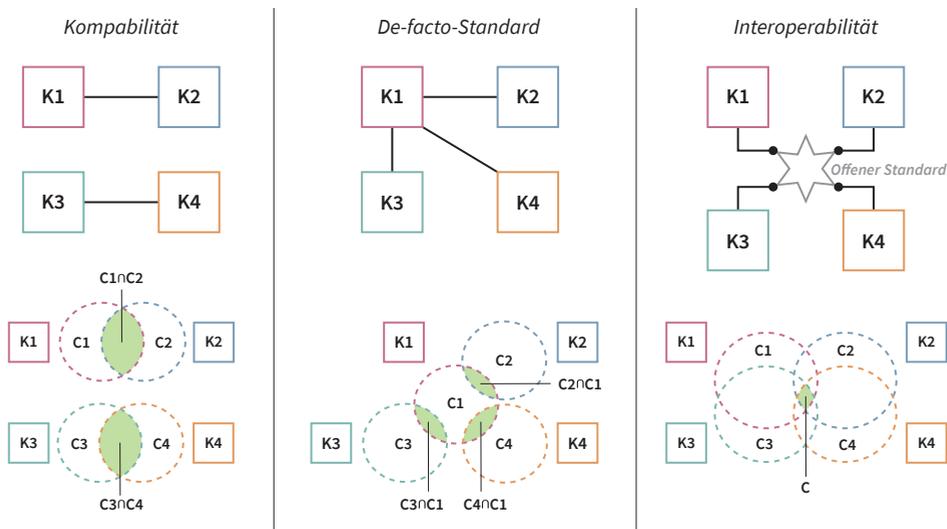
24 vgl. OpenGeoEdu 2018

25 Staub 2009, S. 22

26 OpenGeoEdu 2018

27 ebd.

zeichnet ist, existieren bereits unzählige Standards zur Überwindung der Heterogenität der beteiligten Systeme. Als Beispiel soll an dieser Stelle auf das *Portable Document Format (PDF)* verwiesen werden, welches als plattformunabhängiges Dateiformat entwickelt wurde, um Dokumente mit jedem Betriebssystem, jeder Anwendungssoftware und Hardware wirklichkeitsgetreu darstellen zu können. Wenn die verschiedenen Systeme also wirkungsvoll (interoperabel) zusammenarbeiten und somit Daten verlust- und fehlerfrei miteinander austauschen sollen, bedarf es der Verwendung von offenen Standards. Nur dadurch erlangen alle vernetzten Systeme ein einheitliches Verständnis über die miteinander auszutauschenden Dateninhalte und -strukturen und der für deren Austausch verwendeten Kommunikationsmechanismen.<sup>28</sup> Interoperabilität ist daher maßgeblich abhängig davon, ob innerhalb eines Kommunikationsprozesses offene Standards genutzt werden. Das Vorhandensein bzw. die Verwendung von offenen Standards und deren Einfluss auf die Interoperabilität kann in Form von Operabilitätsgraden beschrieben werden und wird in Abbildung 5 beispielhaft anhand der Kommunikation zwischen vier Kommunikationsteilnehmern dargestellt. Die Grafiken der oberen Reihe verdeutlichen dabei den Sachverhalt vereinfacht. Die Grafiken der unteren Reihe ergänzen diese durch eine semantische Darstellung.



**Abb. 5** Operabilitätsgrade nach vereinfachter und semantischer Betrachtung (obere Grafiken in Anlehnung an Matthias Weiss<sup>29</sup>)

<sup>28</sup> vgl. Fraunhofer-Institut für offene Kommunikationssysteme

<sup>29</sup> Weiss 2018

Jeder Kommunikationsteilnehmer (K1 bis K4) verfügt über einen eigenen Sprachraum/Code (C1 bis C4), mit dem er kommunizieren kann. Für die Kommunikation zwischen den voneinander unabhängigen Kommunikationsteilnehmern ist ein gemeinsames Verständnis und daher ein gemeinsamer Sprachraum (grün markierte Schnittmenge) notwendig. Kompatibilität als Operabilitätsgrad (Abb. 5 links) beschreibt den Zustand, bei dem zwei Kommunikationsteilnehmer in einem geschlossenen Sprachraum miteinander kommunizieren. K1 kommuniziert mit K2, braucht aber für die Kommunikation mit K3 oder K4 einen anderen Sprachraum. Da in diesem Zustand immer nur zwei Systeme zueinander kompatibel sind, kann kein effektiver Datenaustausch zwischen allen Systemen stattfinden. Die Verwendung proprietärer Formate und Standards stellt ein typisches Beispiel hierfür dar.

Ein De-facto-Standard als Operabilitätsgrad (Abb. 5 Mitte) beschreibt den Zustand, bei dem alle Kommunikationsteilnehmer zwar einen gemeinsamen Sprachraum benutzen (Sprachraum von K1 gilt als De-facto-Standard), aber nicht direkt miteinander kommunizieren können. Möchten K2 und K3 miteinander kommunizieren, geht das nur über den Sprachraum von K1. In diesem Fall sind zwar alle Systeme über K1 miteinander kompatibel, es kann jedoch kein unabhängiger Datenaustausch ohne K1 erfolgen.<sup>30</sup> Dieser Zustand ist typisch bei der Verwendung von Standards, die meistens von Industrieunternehmen oder Branchen-Akteuren entwickelt wurden, sich durch die häufige Verwendung als Quasi-Standard etabliert haben, aber (bisher) nicht durch eine Normungs- oder Zertifizierungsinstitution standardisiert sind.<sup>31</sup>

Interoperabilität als Operabilitätsgrad (Abb. 5 rechts) ist erst dann gegeben, wenn alle voneinander unabhängigen Kommunikationsteilnehmer im gleichen Sprachraum widerspruchsfrei miteinander kommunizieren und Daten unabhängig miteinander austauschen können. Dieser Zustand ist gegeben, wenn von allen Kommunikationsteilnehmern ein gemeinsamer, offener Standard verwendet wird. Für ein konkretes Beispiel ist hier das weitverbreitete Kommunikationsprotokoll Open Archives Protocol for Metadata Harvesting (OAI-PMH<sup>32</sup>) zu nennen. Das Protokoll bildet die Grundlage für einen standardisierten Datenaustausch und findet vielfach Anwendung bei Schnittstellen für die maschinelle Weitergabe von Metadaten wissenschaftlicher Publikationen.<sup>33</sup>

---

30 vgl. ebd.

31 vgl. Academic dictionaries and encyclopedias

32 vgl. Open Archives Initiative

33 vgl. Müller 2017, S. 54

## 2.2 Metadatenstandards

Für die Entwicklung eines akademischen Ansatzes einer Kommunikationsstruktur sind die Darstellung von sowie die Möglichkeiten des Austausches der in der Produktion und Distribution von OA-Werken benötigten Metadaten und Metadatenstandards von Bedeutung.

Das Anlegen, Beschreiben, Archivieren und insbesondere der Austausch von Metadaten sind grundlegende Voraussetzungen für die Sicht-, Auffind- und Nachnutzbarkeit von OA-Publikationen und dazugehörigen Forschungsdaten. Die Herausbildung von Metadatenstandards und die damit verbundene Qualitätssteigerung der Metadaten sowie die Verbesserung der Interoperabilität beim Einsatz dieser<sup>34</sup> wurden weiter oben (Kap. 1, Kap. 2.1.3) ausführlich diskutiert. Für das weitere Vorgehen wurde für eine Begriffsbestimmung das Kompetenzzentrum Interoperable Metadaten (KIM) herangezogen:

*„Ein Metadatenstandard ist eine allgemein anerkannte Sprache – mit verbindlichen Regelungen zu Vokabular, Semantik, Syntax und Schema – zur menschen- und maschinenlesbaren Beschreibung von Metadaten.“<sup>35</sup>*

Metadatenstandards ermöglichen eine einheitliche Beschreibung von Metadaten und dienen der fehlerlosen Kommunikation sowie dem effizienten Handling von Daten, hinsichtlich der Auffindbarkeit, Zugänglichkeit und Nachnutzbarkeit von OA-Inhalten. Metadatenstandards nehmen somit auch in der Kommunikationsstruktur eine bedeutende Rolle ein.

### 2.2.1 Heterogenität von Standards

Auf die Heterogenität wurde bereits mehrfach hingewiesen. Im Rahmen der Untersuchung muss diese definiert werden. Die Heterogenität von Metadatenstandards beschreibt Unterschiede und Konflikte innerhalb von Datenmodellen, Vokabular, Semantik, Syntax und Schema, die für eine fehlerlose und verlustfreie Kommunikation der Metadaten überwunden werden müssen. Im Kontext von (Meta-) Daten kann die Heterogenität in verschiedene Kategorien unterteilt werden. In der Literatur finden sich hierfür verschiedene Ansätze.<sup>36</sup> Für die nachfolgende Betrachtung wurde sich an den Arten der Heterogenität, die im Kontext der Informationsintegration verschiedener Datenbankmodelle stehen, von Leser/Naumann orientiert.<sup>37</sup>

---

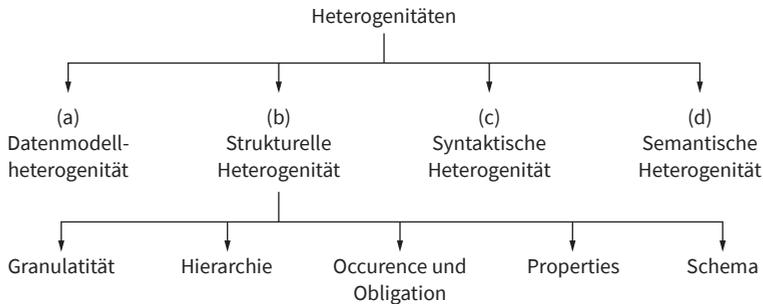
<sup>34</sup> vgl. Pedersen 2005, S.25

<sup>35</sup> vgl. Rühle 2012a, S. 8f.

<sup>36</sup> u.a. Rühle 2012a, S.5; Visser et al. 1997, S. 164f.; Busse et al. 1999, S. 4f.

<sup>37</sup> vgl. Leser/Naumann 2007, S. 60ff.

Es erfolgte eine grundsätzliche Unterscheidung in Datenmodellheterogenität (a), strukturelle (b), syntaktische (c) und semantische Heterogenität (d) (Abb. 6). Die Datenmodellheterogenität beschreibt dabei Probleme bei der Verwendung verschiedener Datenmodelle, während sich die strukturelle Heterogenität aus Unterschieden in strukturellen Darstellungsformen der Metadatenbeschreibung ergibt. Letztere kann sich in verschiedenen Ausprägungen zeigen, daher wurde innerhalb der strukturellen Heterogenität eine weitere Kategorisierung vorgenommen, um für verschiedene Konfliktsituationen entsprechende Lösungswege anbieten zu können (Abb. 6).



**Abb. 6** Eigene Darstellung der Arten von Heterogenität nach Leser/Naumann<sup>38</sup>

Die syntaktische Heterogenität ist in der Verwendung unterschiedlicher Darstellungsformen der Informationen begründet und „[...] Probleme mit der Bedeutung der verwendeten Begriffe und Konzepte“<sup>39</sup> werden von der semantischen Heterogenität zusammengefasst.

### (a) Datenmodellheterogenität

„Heterogenität im Datenmodell liegt vor, wenn das Integrationssystem und eine Datenquelle Daten in unterschiedlichen Datenmodellen verwalten.“<sup>40</sup> Leser und Naumann beziehen sich dabei im Kontext der Informationsintegration auf verschiedene Datenbankmodelle und unterscheiden objektorientierte, relationale sowie XML-Modelle. Darüber hinaus existieren weitere Datenmodelle, wie beispielsweise objektrationale, dokumentorientierte oder Graphdatenbankmodelle. Auch Metadatenbeschreibungen können im Sinne dieser Datenbankmodelle differenziert werden, was u.a. durch die Wahl der Beschreibungssprache und der Darstellungsform erfolgen kann.

<sup>38</sup> vgl. ebd.

<sup>39</sup> Leser/Naumann 2007, S. 61

<sup>40</sup> ebd., S. 65

Objektorientierte Modelle dienen prinzipiell der Beschreibung einer Ressource. Der grundsätzliche Aufbau entspricht einer Baumstruktur mit Wurzel-, Eltern- und Kindknoten, wobei der Wurzelknoten der zu beschreibenden Ressource (1) entspricht, davon abgehende Knoten wiederum entsprechen den Eigenschaften (2) mit Werten (3), die diese Ressource aufweist. Die Darstellung kann dabei grafisch, tabellarisch oder mittels einer XML-Struktur erfolgen (Abb. 7).

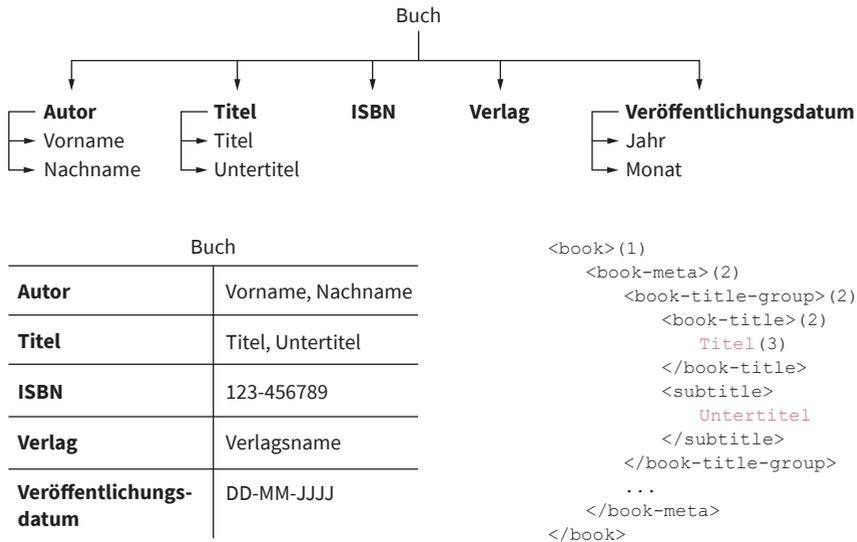


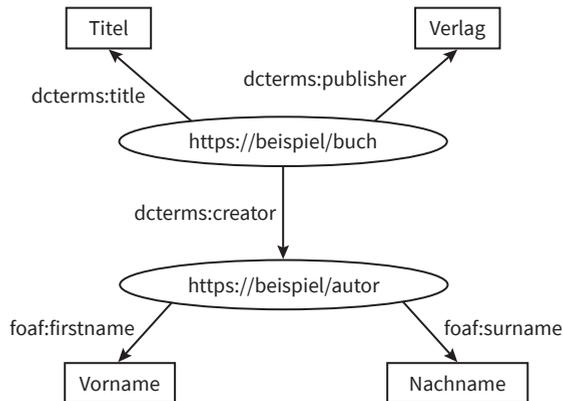
Abb. 7 Darstellungsformen eines objektorientierte Datenmodells

Relationale Modelle hingegen beruhen auf Relationen und dienen der Beschreibung der Beziehungen zwischen verschiedenen Ressourcen. Diese liegen in Tabellensammlungen vor, in der verschiedene Datensätze enthalten sind. Dabei können Tabellen miteinander verknüpft und mittels einer IDs untereinander referenziert werden.

Im Kontext der Metadatenkommunikation ist weiterhin das Resource Description Framework (RDF) zu erwähnen, welches Ressourcen beschreibt. Dieses "umfasst ein Datenmodell, das die Struktur von RDF-Metadaten beschreibt und auf speziellen Graphstrukturen basiert."<sup>41</sup> Auf diesem Modell basieren u.a. Linked Data, wobei es sich um Daten handelt, die mittels eines Uniform Resource Identifier (URI) aufgerufen werden und auf weitere verknüpfte Ressourcen verweisen können. Dies ermöglicht eine flexible Verknüpfung von Ressourcen verschiedener Standards. Die Darstellung kann neben einem RDF-Graph außerdem mittels XML erfolgen.

41 Steffens 2004, S. 56

Die Darstellung einer Metadatenbeschreibung mittels der Standards Dublin Core und Friend of a Friend (FOAF) in Form eines RDF-Graphen ist Abbildung 8 zu entnehmen.



**Abb. 8** Darstellung eines RDF-Graphen nach Enoksson et al.<sup>42</sup>

### *Überwinden der Datenmodellheterogenität*

Die Übersetzung von einem Datenmodell in ein anderes ist nicht immer problemlos möglich. Die Datenmodellheterogenität führt insbesondere dann zu Konflikten, wenn Datenmodelle verschiedene (Metadaten-) Standards enthalten und verknüpfen. Semantische Unterschiede oder nicht vorhandene Elemente können beim Datenaustausch in den Zielstandard dazu führen, dass eine fehlerhafte oder keine direkte Übertragung stattfindet. Wäre der Zielstandard des Datenaustausches in Abbildung 8 bspw. Dublin Core, könnten die FOAF-Elemente nicht direkt übertragen werden. Es bestünde die Gefahr, dass es keine entsprechenden Elemente gibt oder diese in ihrer Bezeichnung abweichen. Um diesem Problem zu begegnen, gibt es verschiedene Möglichkeiten. Zum einen kann für jeden enthaltenen Standard durch Auflösen der Uniform Resource Identifiers (URI) eine Metadatenbeschreibung abgeleitet werden, welche die entsprechenden Informationen enthält und durch die Vergabe einer Identifikation verknüpft ist. Zum anderen können die FOAF-Elemente auf Dublin Core Elemente mittels eines Mappings abgebildet werden. Für ein solches Mapping (Kap. 2.2.2) sind zusätzlich die folgenden Betrachtungen zur strukturellen, semantischen und syntaktischen Heterogenität von Bedeutung. Um von einem objektorientierten Datenmodell in ein relationales Datenmodell zu kommunizieren, ist die Vergabe eindeutiger URIs für Ressourcen, Properties und Beziehungen notwendig. Heterogenität

<sup>42</sup> Enoksson et al. 2007, S. 195

täten in der Darstellungsweise führen zu technischen Konflikten im Austausch von Metadaten. Dies gilt insbesondere für die Kommunikation zwischen grafischen und strukturierten Beschreibungen.

### (b) strukturelle Heterogenität

Wie bei der vorangegangenen Betrachtung bereits deutlich wurde, ist ein gemeinsames Datenmodell zweier Standards kein Garant für eine einheitliche Struktur oder Interoperabilität der Metadatenbeschreibungen. Es können deutliche Unterschiede auftreten, wie die Gegenüberstellung des MARC21 Records und der BITS-Metadatenbeschreibung zeigt (Abb. 9). Diese Differenzen führen zu Konflikten in der Kommunikation der Metadaten zwischen den beiden Standards, was sich u. a. durch verschiedene Properties zur Metadatenbeschreibung und durch unterschiedliche Granularität (Name vs. Vorname und Nachname) zeigt. Neben diesen wurden zusätzlich heterogene Regelungen zu Occurrence, Hierarchie und Schema betrachtet.

|  |   |
|--|---|
| <pre>020 ## \$a ISBN 041 ## \$a Sprachcode       Titel 245 10 \$a Titel       Name (Autor) 263 ## \$a Publikations-       datum (YYYYMM) 300 ## \$a Dimensionen       (Format in cm)</pre> | <pre>&lt;book-part-meta&gt;   &lt;book-title-group&gt;     &lt;book-title xml:lang="Sprachcode Titel"&gt;       Titel Original     &lt;/book-title&gt;   &lt;/book-title-group&gt;   &lt;contrib-group&gt;     &lt;contrib contrib-type="author"&gt;       &lt;name&gt;         &lt;surname&gt;Vorname(Autor)&lt;/surname&gt;         &lt;given-names&gt;Nachname(Autor)&lt;/given-names&gt;       &lt;/name&gt;     &lt;/contrib&gt;   &lt;/contrib-group&gt;   &lt;pub-date iso-8601-date="2002-11"&gt;     &lt;month&gt;Publikationsdatum (MM)&lt;/month&gt;     &lt;year&gt;Publikationsdatum (YYYY)&lt;/year&gt;   &lt;/pub-date&gt; &lt;/book-part-meta&gt;</pre> |
|--|---|

Abb. 9 MARC21 Record (links) und BITS-Metadaten (rechts) eines Buchtitels

### Granularität

Granularität ist die „Anzahl von Untergliederungen eines Elements“<sup>43</sup> und gibt den Verdichtungsgrad von Daten an. Beim Vergleich der Metadatensätze Dublin Core und BITS in Hinblick auf den Autorennamen eines Werkes werden die Unterschiede beider Standards besonders deutlich. BITS weist eine höhere Granularität gegenüber Dublin Core auf, da eine Gliederung des Autorennamens in die Elemente <surname> und <given-names> erfolgt, wohingegen Dublin Core den Vor- und Nachnamen in einem Element <dc:creator> zusammenfasst.



Abb. 10 Darstellung des Metadatum Autor-Name in Dublin Core (links) und BITS (rechts)

### Überwinden von Unterschieden der Granularität

Grundsätzlich ist es einfacher, eine Metadatenbeschreibung von einem Standard mit höherer Granularität in einen Standard mit geringerer Granularität zu transformieren. Dabei kann i.d.R. problemlos jeder benötigten Eigenschaft ein entsprechender Wert zugeordnet werden (Abb. 11). Elemente, die ursprünglich in mehrere Kindelemente untergliedert wurden (hier <name> in <surname>(1) und <given-names> (2)) müssen dafür in einem Element zusammengeführt werden (hier: <dc:creator>). Dies kann erreicht werden, indem die Inhalte der Kindelemente durch ein Trennzeichen (3)<sup>44</sup> separiert ausgegeben werden.

Nachteilig sind an dieser Stelle die unvermeidlichen Informationsverluste durch die geringere Granularität. Zudem kann es auch vorkommen, dass kein geeignetes Zielelement vorhanden ist, welches die Informationen des Quellschemas aufnehmen kann.



Abb. 11 Übersetzung eines Metadatum höherer Granularität (links) in ein Metadatum niedriger Granularität (rechts) am Beispiel des Metadatensatzes Autornamens in Quellschema BITS (links) zu Zielschema DC (rechts)

Schwieriger hingegen ist die Transformation einer Metadatenbeschreibung von einem Standard mit geringer Granularität in einen Standard mit höherer Granularität. Hier sind zunächst allgemeine Regelungen zu untersuchen, wie die Werte der jeweiligen Eigenschaften beschrieben werden, in welchem Format<sup>45</sup> sie aufgebaut sind und welche Trennzeichen vorhanden sind. Daraus können Transformationsanweisungen abge-

44 Weitere mögliche Trennzeichen sind bspw. Leerzeichen, Kommata, Schrägstriche und Minuszeichen.

45 Bei Datumsangaben können bspw. folgende Formate zur Anwendung kommen: TT/MM/JJ, JJJJ-MM-TT und TT.MM.JJJJ.

leitet und ermittelt werden, wie das jeweilige Metadatum in die Kindelemente integriert wird. Soll wie in Abbildung 12 dargestellt das Publikationsdatum von MARC21 nach BITS übertragen werden.

```

263 ## $a 2021(1)10(2)  →  <pub-date>
                             iso-8601-date="2002-11">
                             <year>2021</year>(1)
                             <month>10</month>(2)
                             </dc:creator>

```

**Abb. 12** Übersetzung eines Metadatums geringerer Granularität (links) in ein Metadatum höherer Granularität (rechts) am Beispiel des Metadatums „Publikationsdatum“ im Quellschema MARC21 (links) zum Zielschema BITS (rechts)

Zunächst muss erkannt werden, dass das Datum im Quellschema im Format „JJJJMM“ aufgebaut ist. Für die Transformation müssen demnach die ersten vier Ziffern als Jahresangabe innerhalb `<year>` und die letzten zwei Ziffern als Monatsangabe innerhalb `<month>` übertragen werden. Sind keine festen Regeln für den Aufbau von Werten erkennbar, kann dies zu Fehlern in der Informationsübertragung führen. Weitere Konflikte entstehen, wenn der Standard mit der höheren Granularität Pflichtelemente aufweist, die nicht im Standard mit geringerer Granularität vorhanden sind. Hierbei müssen geeignete Standardwerte angewendet werden, um schwerwiegende Fehler in der Validierung und in weiteren Prozessen zu vermeiden.

## Hierarchie

Die Hierarchie wird zu einem großen Teil durch die Granularität beeinflusst. Unabhängig davon kann eine allgemeine Struktur jedoch tiefer oder flacher gestaltet sein. So weist MARC21 in Abbildung 9 eine sehr flache Struktur auf, in der sich alle Informationen in einer Ebene befinden, während in BITS zahlreiche Unterebenen möglich sind. Die Granularität als auch die Hierarchieebene des Metadaterms auf Seiten von BITS ist größer. Der Vor- und Nachname des Autors befindet sich hierbei ausgehend des Wurzelements in der sechsten Ebene (`book/book-meta/contrib-group/contrib/name/*`). Bei der Dublin Core Metadatenbeschreibung hingegen befindet sich der Name in der zweiten Hierarchieebene (`metadata/dc:creator`).

### *Überwinden von Unterschieden der Hierarchie*

Heterogenitäten in der Hierarchie bereiten prinzipiell keine schwerwiegenden Konflikte, unabhängig davon, ob ein Standard mit höherer oder niedrigerer Hierarchie das Zielschema ist. Bei der Transformation muss der entsprechende XPath angegeben werden.

### Occurrence und Obligation

Ebenfalls zur strukturellen Heterogenität zählen die Unterschiede in den Regelungen der Occurrence (erlaubte maximalen und minimalen Vorkommen) sowie Obligation (verpflichtende oder optionale Angabe). Regelungen zu Occurrence und Obligation des Standards Crossref und BITS zum Metadatum Publikationsdatum sind in Abbildung 13 gegenübergestellt.



**Abb. 13** Heterogenitäten in der Occurrence und Obligation zwischen den Regelungen von BITS (links) und Crossref (rechts) zum Metadatum „Publikationsdatum“

Dabei weisen die Regelungen in BITS keinerlei Einschränkungen auf, wie am Symbol \* (1)<sup>46</sup> erkennbar ist. Die Regelungen in Crossref hingegen erfordern mindestens ein bis maximal zehn Publikationsdaten (2). Somit liegen sowohl Unterschiede im Verpflichtungsgrad als auch im Vorkommen vor.

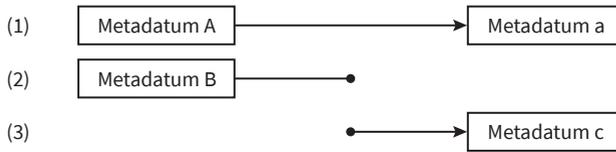
#### *Überwinden von Unterschieden bei Occurrence und Obligation*

Insbesondere Heterogenitäten hinsichtlich der Obligation können zu schwerwiegenden Konflikten führen, wenn die Informationen eines optionalen Elementes zu einem verpflichtenden Element übertragen werden sollen. Der Prozess kann zu Informationsverlust führen, wenn im Quellschema keine entsprechenden Daten enthalten sind. Um Fehler zu vermeiden, müssen geeignete Standardwerte gefunden werden, welche die Pflichtelemente ergänzen. Der umgekehrte Fall hingegen – die Transformation eines Pflichtelementes in ein optionales Element – führt i. d. R. zu keinen Konflikten. Ebenfalls unproblematisch ist die Transformation von weniger erlaubten Elementen zu mehr erlaubten Elementen. Bei der Übertragung von mehr erlaubten Elementen auf weniger erlaubten Elementen hingegen muss ggf. manuell eine geeignete Auswahl getroffen werden, um dem Zielschema zu entsprechen.

### Properties

Ein weiterer Konflikt der zwischen zwei Standards auftreten kann und der strukturellen Heterogenität zuzuordnen ist, wird dadurch beeinflusst, welche Properties (Eigenschaften) ein Standard enthält. Beim Vergleich der Properties zweier Standards können prinzipiell drei unterschiedliche Fälle auftreten (Abb. 14).

46 Das Symbol „\*“ steht für 0 bis unendlich, „+“ für 1 bis unendlich und „?“ steht für 0 oder 1.



**Abb. 14** Darstellung möglicher Konstellationen im Vergleich der Eigenschaften zweier Standards

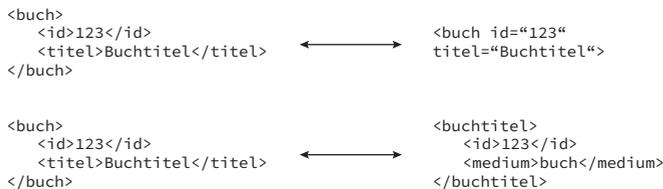
Im ersten Fall entsprechen die Eigenschaften des Quellstandards denen des Zielstandards (1). Im zweiten Fall enthält der Quellstandard eine Property, die keine Entsprechung im Zielstandard findet (2) und im dritten Fall enthält der Zielstandard Properties, die im Quellstandard nicht vorliegen (3).

### *Überwinden von Unterschieden der Properties*

Der erste Fall führt prinzipiell nicht zu Konflikten, allerdings müssen ggf. weitere Heterogenitäten in Semantik und Syntax beachtet werden. Der zweite Fall hingegen ist mit einem möglichen unvermeidbaren Informationsverlust verbunden. Im dritten Fall können Fehler in der Validierung und in nachfolgenden Prozessen auftreten, weshalb an dieser Stelle möglichst geeignete Standardwerte gefunden werden müssen.

### **Schema**

Bei der schematischen Heterogenität kann nach Leser/Naumann<sup>47</sup> der strukturellen Heterogenität zugeordnet werden. Dabei wird dasselbe Datenmodell verwendet, jedoch unterscheiden sich die Relationen oder die Darstellung der Werte als Elemente, Kindelemente und Attribute, wie in Abbildung 15 skizziert.



**Abb. 15** Beispielhafte Darstellung schematische Heterogenität hinsichtlich der Darstellung der Werte als Elemente, Kindelemente und Attribute (oben und hinsichtlich der Relation (unten))

47 vgl. Leser/Naumann 2007, S. 69ff.

*Überwinden schematischer Heterogenität*

Um Unterschiede zwischen Elementen, Kindelementen und Attributen zu überwinden, können Elemente mittels eines Transformations-Stylesheets zu Unterelementen oder Attributen oder Attribute zu Elementen oder Unterelementen transformiert werden. Ähnlich der Hierarchie entstehen hierbei in der Regel ebenfalls keine schwerwiegenden Konflikte.

**(c) semantische Heterogenität**

Während sich die strukturelle Heterogenität aus dem grundsätzlichen Aufbau und Inhalten der Standards ergibt, zeigt sich die semantische Heterogenität durch Unterschiede in der Interpretation von Namen und Symbolen.<sup>48</sup> Dabei wird sowohl die Bedeutung der Metadaterme als auch die Bedeutung der Werte an sich betrachtet. Im Fokus stehen Konflikte, die durch Homonyme und Synonyme entstehen. Bei Homonymen handelt es sich um gleiche Begriffe mit unterschiedlicher Bedeutung. So kann „Titel“ bspw. für den Titel des Autors, für den Titel und den Untertitel des Buches oder nur für den Haupttitel des Buches stehen. Weitaus häufiger sind allerdings Synonyme, unterschiedliche Begriffe mit gleicher Bedeutung, zu beachten. So ist z. B. der Zeichencode des Standards MARC21 „245 10 \$a“ ein Synonym zur `<book-title-group>` in BITS und zu `<title>` in Dublin Core. Häufig korrelieren Synonyme und Homonyme zudem mit Heterogenitäten in Granularität und Hierarchie.

*Überwinden semantischer Heterogenität*

Synonyme können die Kommunikation einer Metadatenbeschreibung zwischen verschiedenen Standards deutlich erschweren, da nicht direkt ersichtlich ist, welche Informationen des Quellschemas welchen Eigenschaften des Zielschemas zuzuordnen sind. Schwerwiegende Konflikte können auch durch Homonyme entstehen. Sind Homonyme vorhanden, so müssen die Metadaten des Quellstandards entweder unter passenden anderen Elementen ausgegeben oder durch den Zusatz von Attributen abgegrenzt werden. Liegen verschiedene Properties mit derselben Bezeichnung in unterschiedlicher Bedeutung vor, so ist eine falsche Zuordnung der Werte sehr wahrscheinlich. In der Kommunikation kann dies zu Fehlinformationen oder Informationsverlusten führen. Um solche Konflikte aufzudecken und zu überwinden, muss der jeweilige Kontext der Metadaterme und die Dokumentationen der Standards, in denen Regelungen, Definitionen und Hinweise zu jedem erlaubten Element festgehalten sind, beachtet werden. Eventuelle Differenzen in der Bedeutung oder in der Bezeichnung können auf

---

48 vgl. ebd., S. 74

diese Weise identifiziert und Metadaten entsprechend transformiert werden. Bei Synonymen können die Metadaten des Quellstandards unter dem entsprechenden Element des Zielstandards ausgegeben werden.

#### **(d) Syntaktische Heterogenität**

Syntax ist die „in einer Sprache übliche Verbindung von Wörtern zu Wortgruppen und Sätzen“ und beschreibt die „korrekte Verknüpfung sprachlicher Einheiten im Satz“<sup>49</sup>. Bezogen auf Metadaten sowie im Kontext der Heterogenität bezieht sich Syntax vor allem auf die Darstellung der Werte. Konflikte entstehen hierbei durch die Verwendungen unterschiedlicher Zahlenformate<sup>50</sup>, Zeichencodierung<sup>51</sup>, Trennzeichen sowie Datenformate<sup>52</sup>. Weitere Konflikte entstehen, wenn durch die Regelungen des Standards bereits ein konkret erlaubtes Vokabular vorgegeben ist. So kann bspw. der mögliche Inhalt eines Elements <month> auf „Januar“, „Februar“, [...], „Dezember“ oder „01“, „02“, [...], „12“ begrenzt sein. Unterscheiden sich die Regelungen verschiedener Standards hinsichtlich des Vokabulars, können an dieser Stelle Fehlermeldungen entstehen.

#### *Überwinden der syntaktischen Heterogenität*

Um Fehlinformationen oder Fehlermeldungen durch syntaktische Heterogenität zu vermeiden, müssen bei der Transformation die Regelungen zu Syntax und Vokabular beachtet werden, die in den Dokumentationen der Standards festgehalten sind. Treten Unterschiede auf, so müssen diese möglichst überwunden werden, indem geeignete Regeln für die Umwandlung der Zahlen-, Zeichen- und Datenformate formuliert werden. Unterscheidet sich das erlaubte Vokabular, so kann nach entsprechenden erlaubten Synonymen gesucht werden. Ist allerdings im Zielschema kein Vokabular mit der angestrebten Bedeutung vorhanden, so kommt es zu Informationsverlusten und geeignete Standard- oder Ausweichwerte müssen Anwendung finden.

## **2.2.2 Transformation**

Ein Hilfsmittel, welches die Kommunikation von Metadaten zwischen verschiedenen Standards unterstützt, ist das Mapping. Es bezeichnet den „Prozess des Verknüpfens verschiedener Metadatenschemata“<sup>53</sup> und stellt einen wesentlichen Schritt zur effizienten und fehlerfreien Kommunikation von Metadaten zwischen verschiedenen

---

49 Duden 2023b

50 z.B. Binärzahlen, gebrochene Zahlen, ganze Zahlen

51 z.B. UTF-8 oder UTF-16

52 z.B. Uhrzeit- und Datumsangaben

53 Kálmán et al., S. 4

Standards dar, mit dem Ziel, semantische, strukturelle und syntaktische Interoperabilität herzustellen. Für die Transformation zwischen Strukturen verschiedener Standards bzw. Quasi-Standards sind verschiedene Regeln anzuwenden. Hier sollen nun die grundsätzlichen Regeln diskutiert werden. Deren Implementierung erfolgt in den Stylesheets (siehe Kap. 3ff.), welche das jeweilige Mapping (die Transformation) durchführen.

### Regel (1:0)

Diese Regel drückt den Fall aus, dass in der sendenden Struktur ein Metadatum vorhanden und mit Daten versehen ist, aber in der empfangenden/zu erstellenden Struktur kein Pendant vorhanden ist. Daher ist für den Regelfall anzunehmen, dass kein Transformationsbedürfnis besteht, weil kein Zielelement vorhanden ist. Der Ausgangsknoten kann daher im Stylesheet entweder still übergangen werden oder es kann eine Mitteilung erzeugt werden, dass die Transformation mangels eines Zielelementes nicht durchgeführt werden kann. Im nicht regelhaften Ausnahmefall müsste ein Element in der empfangenden/zu erstellenden Struktur umgewidmet werden (1:0i, siehe Tab. 1). Hierfür kann und soll keine Standardisierung erfolgen.

### Regel (0:0)

Diese Regel wird nur der Vollständigkeit halber aufgeführt. Sie hat keinen praktischen Anwendungsfall. Wenn weder in der sendenden noch in der empfangenden/zu erstellenden Struktur das betreffende Element vorhanden ist, besteht weder ein Transformationsbedürfnis noch eine Transformationsmöglichkeit. Ein abgeleiteter Fall wäre eine beliebige andere Regel, die aber nicht angewendet werden muss, weil kein Transformationsbedürfnis besteht, d.h., in der sendenden Struktur ist ein semantisches Element vorhanden, ggf. auch mit Daten versehen, in der Zielstruktur ist ggf. dieses Element auch vorhanden, muss aber für den aktuellen Anwendungsfall nicht befüllt werden.

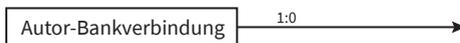
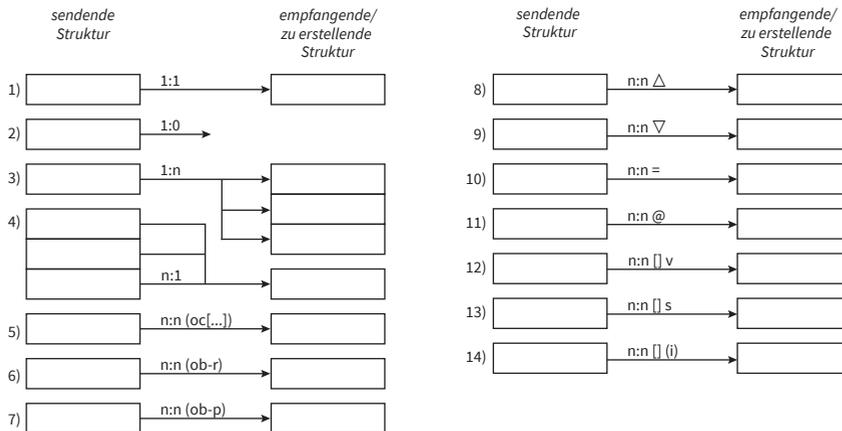


Abb. 16 Relation Autor-Bankverbindung

Weitere Regeln und mögliche Beziehungen zwischen der sendenden und der zu empfangenden/zu erstellenden Struktur werden in Abbildung 17 dargestellt:



**Abb. 17** Mögliche Beziehungen zwischen Metadatenbeschreibungen verschiedener Standards

Die Narrationen der in der Abbildung dargestellten Regeln werden in Tabelle 1 wiedergegeben.

Prinzipiell dienen diese Transformationsregeln als Grundlage für die Erstellung von Transformationsstylesheets oder Algorithmen zur Konvertierung einer Metadatenbeschreibung von einer sendenden (Quell-) Struktur in einen andere, ebenfalls standardisierte empfangende (Ziel-) Struktur. Im Normalfall sind die zu vergleichenden Strukturen bereits bekannt und jedes Element des Quellstandards muss mit dem Zielstandard in Beziehung gesetzt werden, ohne Konzentration auf bestimmte Metadaten-terme. Eine Transformation ist demnach zumeist standardorientiert bzw. bereits schon standardisiert.

Im Falle der OA-Kommunikationsstruktur hingegen sollen zunächst verschiedene Metadatenstandards kombiniert werden, um einen individuellen Umfang von Metadaten abzudecken, ohne einen neuen Standard entwickeln zu müssen. Zusätzlich muss die Kombination geeignet sein, um die Metadaten zwischen den Metadatenstandards der am Prozess beteiligten Systeme zu kommunizieren. Dabei sind die im Workflow benötigten Informationen vorgegeben. Die Hauptaufgabe der Transformation besteht in diesem Fall darin, den relevanten Metadaten entsprechend geeignete Metadaten-terme verschiedener Standards zuzuordnen und dabei Redundanzen, Informationsverlust und Fehlinformationen zu vermeiden. Die Zuordnung der Metadaten zu Metadaten-terminen sowie entsprechende Regelungen zu Syntax, Vokabular und sonstigen Eigenschaften bilden einen Teil des Lastenheftes für die Kommunikationsstruktur.

**Tab. 1** Regeln für die Transformation von Elementen zwischen zwei standardisierten Strukturen

| Nr. | Symbol   | Beziehung  |
|-----|----------|--|
| 1   | 1:1      | Element des Quellstandards entspricht genau einem Element des Zielstandards.   |
| 2   | 1:0      | Element des Quellstandards entspricht keinem Element des Zielstandards.  |
| 3   | 1:n      | Element des Quellstandards ist im Zielstandard in mehrere Elemente aufgeteilt.   |
| 4   | n:1      | Mehrere Elemente des Quellstandards sind im Zielstandard zu einem Element zusammengefasst.   |
| 5   | oc [...] | Das Element des Zielschemas weist differierende Regelungen zur Occurrence (oc) auf. Innerhalb von [...] können genaue Angaben zur Occurrence gegeben werden. |
| 6   | ob-r     | Das Element des Zielschemas weist differierende Regelungen zur Obligation (ob) auf (r-required, o-optional).   |
| 7   | ob-o     |  |
| 8   | $\Delta$ | Element des Quellstandards ist in höherer ( $\Delta$ ),<br>niedrigerer ( $\nabla$ ) oder<br>gleicher (=) Hierarchie-Ebene.                                   |
| 9   | $\nabla$ |  |
| 10  | =        |  |
| 11  | @        | Element des Quellstandards wird durch ein Attribut im Zielstandard dargestellt, spezifiziert oder von anderen Elementen differenziert.                       |
| 12  | v        | Syntax des Zielschemas erfordert differierendes Vokabular (v)<br>oder differierende Regelungen (s).  |
| 13  | s        |  |
| 14  | (i)      | Element des Quellschemas ist implizit durch Struktur, Standard oder andere Elemente des Zielschemas gegeben.   |

## 2.3 Offene Standards in der wissenschaftlichen Publikationslandschaft

Eine wesentliche Grundlage zur Erforschung des Ist-Zustandes, stellte die Identifikation und Analyse bereits vorhandener Standards und Qualitätsstandards für die Kommunikation im Problemfeld dar. Dazu wurde eine systematische Literaturrecherche durchgeführt sowie freie Gespräche ohne Interviewleitfäden mit ausgewählten Experten geführt. Die Recherche und anschließende Analyse von Standards wurden dabei unter Berücksichtigung der Vorgaben des State-of-the-Art-Publikationsworkflows zur Herstellung und Verbreitung von OA-Monografien<sup>54</sup> durchgeführt.

In einer ersten Grobrecherche wurden 41 Informationsquellen<sup>55</sup> systematisch nach 16 ausgewählten Suchbegriffen in jeweils deutscher und englischer Sprache durchsucht<sup>56</sup> und entsprechende Literatur gesichtet und geprüft, ob diese Relevanz eine tiefere Analyse darstellt. Im Anschluss an die Literaturanalyse wurden die Erkenntnisse der Recherche in einem umfangreichen Glossar mit verlinkten Hintergrundinformationen zusammengetragen. Das Glossar umfasst neben Standards auch Datenformate, Lizenzen, Developer, Institutionen sowie Akteure im Bereich Standards im wissenschaftlichen Publizieren (Stand 17.05.2021: 234 Begriffe). Zur Eingrenzung des Glossars auf die wesentlichen Begriffe, d.h. offene Standards im wissenschaftlichen Publizieren, wurden entsprechende Kriterien zur Identifizierung einzelner Spezifikationen entwickelt. Mittels der erarbeiteten Kriterien wurde zunächst als Grundlage für die Analyse eine spezifische Arbeitsdefinition des Begriffes Standard im Kontext von Open-Access-Publikationsprozessen entwickelt:

*Ein offener Standard für wissenschaftliches Publizieren beschreibt eine einheitlich dokumentierte, weithin anerkannte, vielfach angewandte, offen zugängliche und erweiterbare Spezifikation, die bei der Erstellung, Beschreibung, Herstellung und Verbreitung wissenschaftlicher Publikationen angewandt wird.*

Die Verwendung dieser spezifischen Arbeitsdefinition und der darin aufgeführten Kriterien ermöglichte eine Filterung der wesentlichsten Standards aus der großen Masse an Standards.

---

54 Böhm et al. 2020

55 u.a. ResearchGate, ScienceOpen, Karlsruher Virtueller Katalog (KVK), SLUB, OAPEN, BASE, OpenAIRE, Zenodo, ROAR, DOAJ, Google

56 Die insgesamt 32 Suchbegriffe begrenzten sich durch die Doppelung von Begriffen während der Übersetzung auf 29 Suchbegriffe.



## 2.4 Anwendung und Analyse OA-Workflow

Für die konkrete Betrachtung und Beschreibung aller Schnittstellen im gesamten OA-Publikationsprozess dient das allgemeingültige Workflowmodell<sup>58</sup> als Grundlage. Es lässt sich feststellen, dass zwischen einzelnen Prozessen, bei denen zwei oder mehrere Akteure bzw. Systeme miteinander kommunizieren, grundlegend ein Austausch von Informationen (organisatorische Kommunikation) und/oder Daten (technische Kommunikation) stattfindet. Für die Entwicklung einer Kommunikationsstruktur ist vor allem die technische Kommunikation und deren Schnittstellen von Bedeutung, da diese organisiert werden müssen. Die organisatorische Kommunikation an Schnittstellen soll an dieser Stelle als gegeben betrachtet werden. Es erfolgte eine Analyse des OA-Workflowmodells unter Anwendung der theoretischen Grundlagen, welche nachfolgend dargelegt wird.

### 2.4.1 Schnittstellen

In Hinblick auf die Verortung von Schnittstellen liefert das Workflowmodell mit seinen Prozessbeschreibungen bereits einen detaillierten Überblick aller internen Prozessschnittstellen, bei denen die Kommunikation ausschließlich innerhalb des Systems der publizierenden Institution stattfindet. Dies betrifft überwiegend Schnittstellen zwischen einzelnen Prozessen, bei denen Daten und/oder Informationen von vorherigen Prozessen konsumiert und (in der Regel durch den Prozess veränderte) an nachfolgende weitergegeben werden. Ein Beispiel dafür ist in Abbildung 19 dargestellt. Der Prozess M-Pv-310<sup>59</sup> konsumiert Informationen (I1<sup>60</sup> = die zu erwartenden Produktionskosten) vom vorgelagerten Prozess M-Pv-300 und gibt veränderte (produzierte) Informationen (O1 = Soll-Kalkulation) an den nachfolgenden Prozess M-Pv-320 weiter.



**Abb. 19** Beispiel einer Prozessschnittstelle im allgemeingültigen Workflowmodell

Die Beschreibung von externen Schnittstellen, also die technische Kommunikation der publizierenden Institution mit den Systemen weiterer Akteure, ist demzufolge für die Entwicklung eines akademischen Ansatzes einer Kommunikationsstruktur von großer

<sup>58</sup> vgl. Böhm et al. 2020, S. 55ff.

<sup>59</sup> vgl. ebd., S. 192

<sup>60</sup> I = Input, O = Output

Bedeutung. Die Anforderungen dieser Systeme hinsichtlich der Datenübertragung und -beschaffenheit müssen dabei für einen effizienten Austausch innerhalb einer strukturierten Kommunikation Berücksichtigung finden. Für die Betrachtung externer Schnittstellen wird davon ausgegangen, dass alle im Workflowmodell erfassten Prozesse, die ausgelagert werden können, von externen Akteuren und deren Systemen ausgeführt werden. Diese Annahme stellt das Maximum externer Schnittstellen dar. Ausgehend von dieser Annahme wurden alle möglichen externen Kommunikationsteilnehmer, die an OA-Publikationsprozessen beteiligt sein können, herausgearbeitet (Anhang B). Dabei wurden die Kommunikationsteilnehmer insofern möglich und sinnvoll institutionalisiert. Damit soll hervorgehoben werden, dass die aufgeführten Akteure in der Regel sowohl ein organisatorisches System (Mensch) als auch ein technisches System (Maschine) darstellen können.

Anschließend wurde untersucht, an welchen Prozessschnittstellen im allgemeingültigen Workflowmodell, welche externen Kommunikationsteilnehmer mit der publizierenden Institution in einen Austausch gehen. Dabei konnten 47 Prozesse<sup>61</sup> identifiziert werden, bei denen eine Kommunikation mit externen Beteiligten und somit eine externe Schnittstelle im Publikationsprozess auftreten kann. Davon findet in 29 Prozessen eine technische Kommunikation statt, bei der eine Kommunikationsstruktur potenziell zur Anwendung kommen kann. Eine spezifische Beschreibung all dieser Prozesse mit technischer Kommunikation ist daher erforderlich. Im Folgenden sollen anhand von drei konkreten Beispielen Schnittstellen innerhalb des OA-Publikationsprozesses nach den eingangs vorgestellten Faktoren beschrieben werden. Die Beschreibung erfolgt dabei nur beispielhaft, da die Prozesse aufgrund ihrer Allgemeingültigkeit in der Praxis auf unterschiedlichste Art ausgeführt werden können.

### **Beispiel 1: Prozess P-EP-100 – Produkt auf primärer Veröffentlichungsplattform veröffentlichen<sup>62</sup>**

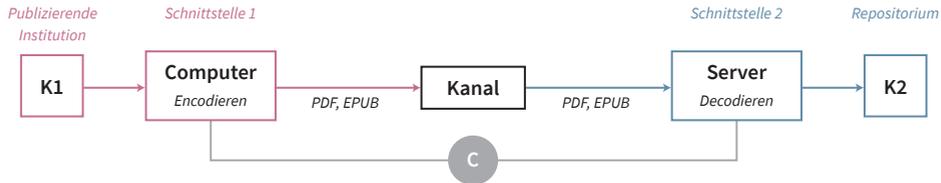
Bei diesem Prozess wird die elektronische Ausgabe der finalen Publikation auf die primäre Veröffentlichungsplattform (in der Regel das institutionelle Repositorium) hochgeladen. Die Spezifik der bei diesem Prozess vorhandenen Schnittstelle wird durch folgende Kriterien beschrieben:

---

61 Von insgesamt 78 Prozessen: 30 M-Prozesse (managementorientierte Prozesse), 24 I-Prozesse (inhaltsorientierte Prozesse), 24 P-Prozesse (produktionsorientierte Prozesse). Null-Prozesse wurden nicht untersucht.

62 vgl. Böhm et al. 2020, S. 168

- K = 2: Publizierende Institution (sendende Infrastruktur),  
Repositorium(empfangende Infrastruktur)
- Eigenschaften: synchron, unidirektional
- Infrastruktur: technische Kommunikation
- Typ: Mensch-Maschine-Schnittstelle



**Abb. 20** Darstellung der Schnittstellen-Spezifik beim Prozess P-EP-100: Produkt auf primärer Veröffentlichungsplattform veröffentlichen

Die Kommunikationsteilnehmer bei diesem Prozess sind zum einen die publizierende Institution als sendendes und zum anderen das Repositorium als empfangendes System. Eine der publizierenden Institution angehörige Person (Mensch) lädt die elektronische Ausgabe eines Werkes (z.B. in den Formaten PDF und EPUB) auf den Server des primären Repositoriums hoch. Hierbei empfängt ein Server (Maschine) die gesendeten Daten. Der Kanal für die stattfindende Datenübertragung kann beispielsweise ein Web-Formular sein, welches den Datentransfer zwischen dem sendenden und dem empfangenden System regelt. Zur Anwendung kommen hierfür beispielsweise Austausch-Standards<sup>63</sup> wie HTTPS. Die stattfindende Kommunikation erfolgt synchron, da der Server die gesendeten Daten ohne Zeitverzögerung annimmt und decodiert. Da eine Übertragung von Daten erfolgt, findet eine technische Kommunikation statt.

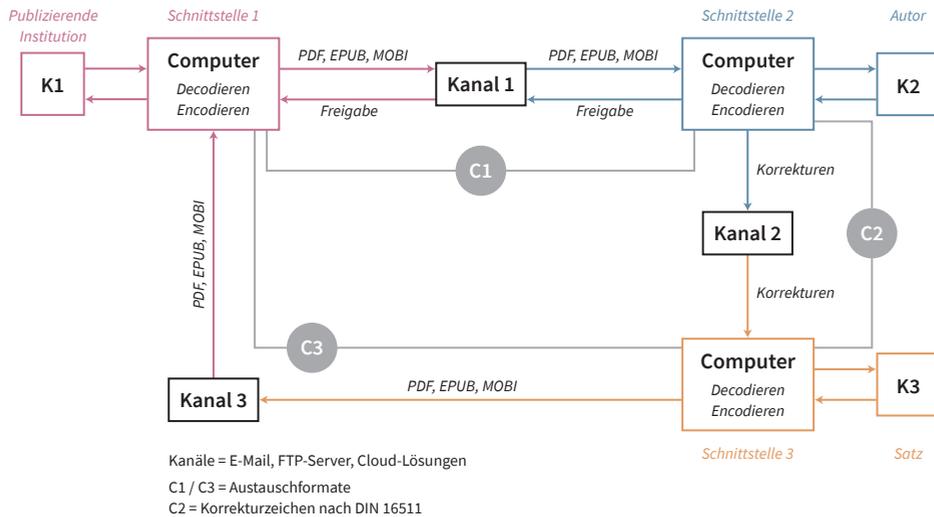
### Beispiel 2: Prozess P-CiP-160 – Freigabe vom Content-Urheber einholen<sup>64</sup>

Innerhalb dieses Prozesses erfolgt die Übermittlung der Publikation in allen gewünschten Ausgabeformaten zur Freigabe. Die Daten werden kontrolliert und ggf. Korrekturwünsche erteilt. Folgende Kriterien beschreiben die Spezifik des Prozesses P-CiP-160:

- K = 3: Lektorat, Autor, Satz
- Eigenschaften: bi-/unidirektional, a-/synchron
- Infrastruktur: organisatorische und technische Kommunikation
- Typ: Mensch-Maschine-Schnittstelle

<sup>63</sup> vgl. Böhm et al. 2021

<sup>64</sup> vgl. Böhm et al. 2020, S. 154



**Abb. 21** Darstellung der Schnittstellen-Spezifika beim Prozess P-CiP-160: Freigabe vom Content-Urheber einholen

An diesem Prozess sind drei Kommunikationsteilnehmer beteiligt, welche als sendendes und empfangendes System agieren: die publizierende Institution, Autoren sowie der Satz. Die publizierende Institution übermittelt die Publikation in allen gewünschten Ausgabeformaten (z.B. in den Formaten PDF, EPUB, MOBI) via E-Mail, über einen FTP-Server oder Cloud-Lösungen an die Autoren. Über die Autoren erfolgt die Kontrolle der gesendeten Formate. Gibt es keine weiteren Korrekturen, so senden die Autoren eine Freigabe (Information) per E-Mail an die zu publizierende Institution. Im Falle von Korrekturwünschen senden die Autoren diese an den Satz. Nach Einarbeitung der Korrekturen übermittelt der Satz diese wieder in den gewünschten Ausgabeformaten an die publizierende Institution. Die Kommunikation zwischen K1 und K2 erfolgt bidirektional, K2 und K3, sowie K3 und K1 kommunizieren unidirektional. Standards, welche innerhalb dieses Prozesses zur Anwendung kommen können neben Korrekturzeichen nach DIN 16511, verschiedene Austausch-Standards, wie beispielsweise HTTPS sein.

### Beispiel 3: Prozess I-Qe-130 – Peer-Review beauftragen<sup>65</sup>

Bei diesem Prozess beauftragt die publizierende Institution ein Gutachten mit einem Peer-Review am wissenschaftlichen Werk und übermittelt den Content. Die Spezifik der bei diesem Prozess vorhandenen Schnittstelle wird durch folgende Kriterien beschrieben:

- K = 2: Publizierende Institution, Gutachter
- Eigenschaften: asynchron, bidirektional
- Infrastruktur: organisatorische und technische Kommunikation
- Typ: Mensch-Maschine-Schnittstelle

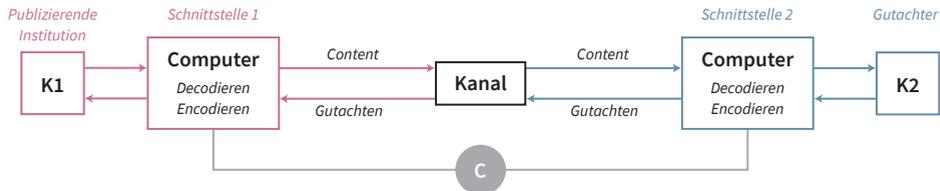


Abb. 22 Darstellung der Schnittstellen-Spezifik beim Prozess I-Qe-130: Peer-Review beauftragen

Die Kommunikationsteilnehmer bei diesem Prozess, publizierende Institution und Gutachter, agieren bei diesem Kommunikationsprozess als sendendes und empfangendes System. Mitarbeitende der publizierenden Institution senden den Content des wissenschaftlichen Werks (z.B. im Format Word oder PDF) und die für das durchzuführende Peer-Review-Verfahren benötigten Informationen an die Person, die das Gutachten durchführt. Der Übertragungskanal für diese technische und organisatorische Kommunikation ist zumeist E-Mail. Nach der Durchführung des Peer-Reviews wird das Gutachten an Mitarbeitende der publizierenden Institution übermittelt. Die Kommunikation findet dabei asynchron statt. Zur Anwendung kommen können qualitätssichernde Richtlinien und Vorgaben der publizierenden Institution bezüglich der Verfahren und Methoden des Peer-Reviews.

Die beschriebenen Ansätze und theoretischen Grundlagen zu Kommunikation, Schnittstellen und Interoperabilität leisten ein grundlegendes Verständnis und bieten hilfreiche Werkzeuge, um sich der Schnittstellen-Spezifika im wissenschaftlichen Publikationsprozess zu nähern. Der Schlüssel für einen effektiven, fehler- und verlustfreien Datenaustausch zwischen interoperablen Systemen ist dabei die Verwendung offener Standards.

<sup>65</sup> vgl. ebd., S. 106

### 2.4.2 Metadatenmapping

Um sich der technischen Konstruktion weiter anzunähern, wurde außerdem untersucht, welche Daten einer OA-Publikation von welchen Standards abgedeckt werden können. Dazu wurden in einem ersten Schritt alle Daten, insbesondere Metadaten, innerhalb des Workflowmodells gesammelt, beschrieben und in sechs Datengruppen gegliedert (Anhang C). Auf Grundlage der beschriebenen Metadaten wurde die technische Dokumentation einer Stichprobe von zwölf offenen Standards analysiert, mit den Daten des OA-HVerlages abgeglichen und in einer Datenmatrix passende Elemente identifiziert und zusammengetragen. Die Analyse erhebt keinen Anspruch auf Vollständigkeit und Richtigkeit, da sich die technischen Dokumentationen teilweise als sehr komplex darstellen, dennoch verdeutlicht die Matrix die Verschachtelung einzelner Standards und wie diese einander durchdringen (siehe auch Kap. 3.2).<sup>66</sup> Diese Erkenntnis ist wesentlich für die Entwicklung eines Ansatzes einer Kommunikationsstruktur.

Wie in Abschnitt 2.2.2 erläutert, unterstützt Mapping die Kommunikation von Metadaten zwischen verschiedenen Standards. Daher wurden ausgehend von der Metadatenstruktur des OA-Workflowmodells (Anhang C) und unter Beteiligung der Quasi-Standards BITS, ONIX, Crossref sowie Schema.org Transformationsregeln für das Mapping erarbeitet (Anhang D<sup>67</sup>). Die Datengruppe Peer-Review wurde um die Metadaten ergänzt, welche sich durch den Informationsworkflow, den Peer-Review-Prozess sowie der erforderlichen Metadaten in BITS und Crossref ergeben. Die Daten zur Barrierefreiheit wurden entsprechend der von ONIX, Crossref und Schema.org geforderten Metadaten präzisiert. Dabei wurde versucht, den gesamten Umfang der im Produktions-, Publikations- und Distributionsprozess benötigten Metadaten abzudecken und somit nicht auf nur ein Anwendungsszenario zu begrenzen.

Der Metadatensatz aus dem Workflowmodell wurde in funktionsbezogene Kategorien gegliedert, welche sich aus der hauptsächlichen Funktion der jeweiligen Metadaten ergeben: Deskriptive Metadaten, die der näheren Beschreibung des Werks dienen, Administrative Metadaten, die bspw. den Produktions- und Distributionsprozess unterstützen, strukturelle Metadaten, die ein Dokument strukturieren und Linkage-Daten, die Beziehungen zu anderen Ressourcen beschreiben. Die Datengruppen einer Open-Access-Publikation im OA-HVerlag aus Anhang C wurden wie folgt in die Metadatenkategorien einsortiert:

---

<sup>66</sup> Reiche et al. 2023, S.4ff.

<sup>67</sup> Die Darstellung des Mapping in Anhang D basiert auf der Arbeit von Sannemann (2021). Nach 2021 wurden die Datengruppen aus Anhang C, auf der das Mapping basiert, überarbeitet. Die Datengruppen in Anhang D wurden entsprechend aktualisiert.

- **Deskriptive Metadaten**
  1. Daten zu Werk und Lizenzen
  2. Daten zu Mitwirkenden
- **Administrative Metadaten**
  1. Daten zu Werk und Lizenzen
  3. Daten zur Produktion
  4. Daten zu Klassifikation
  5. Daten zu Peer-Review
  6. Daten zu Barrierefreiheit
- **Strukturelle Metadaten:** werden bereits durch die Strukturierung der OA-Inhalte mittels BITS beschrieben, weshalb diese nicht weiter beachtet werden müssen.
- **Linkage/Relationship-Daten:** werden zum jetzigen Zeitpunkt bei den bisherigen Auftragsdaten nicht gesondert betrachtet. Beziehungen zwischen Werk und dessen Komponenten sind größtenteils durch strukturelle Metadaten beschrieben, während Beziehungen zwischen Werken zum Teil in deskriptiven und administrativen Metadaten enthalten sind.

Um das Ziel einer möglichst simplen und leicht verständlichen Kommunikationsstruktur zu unterstützen, wurde von einem Hauptstandard ausgegangen. Während des Mappings wurde demnach in einem ersten Schritt eine Zuordnung eines Metadaterms zum gewählten Hauptstandard angestrebt. Erst wenn dieser keinen geeigneten Term aufwies, wurden alternative Standards hinzugezogen. Da der BITS-Standard in der Strukturierung der OA-Inhalte des OA-HVerlag bereits Anwendung findet, ist dieser am ehesten als Hauptstandard geeignet. Als alternative Standards wurden ONIX, Schema.org und Crossref u. a. wegen ihrer detaillierten Ausprägung hinsichtlich Peer-Review und Barrierefreiheit in Betracht gezogen. Eine detaillierte Auswertung wird in der Arbeit von Sannemann diskutiert.<sup>68</sup>

---

<sup>68</sup> vgl. Sannemann 2021, S. 80ff.



## 3 Grundlagen zur Entwicklung einer Kommunikationsstruktur

---

Der akademische Ansatz einer Kommunikationsstruktur soll es Hochschulen, Universitäten und Bibliotheken ermöglichen, im Betrieb eines OA-Verlages Arbeits- und Kostenaufwände erheblich zu senken. Auch ohne herstellerische Kenntnisse, soll ein übersichtliches und einfaches Hinterlegen und Kommunizieren von Metadaten im gesamten Publikationsprozess ermöglicht werden. Auf diese Weise können Prozessineffizienzen eliminiert und der Automatisierungsgrad erhöht werden.

### 3.1 Rahmenbedingungen für die Diskussion

Die Untersuchungen der im Feld vorhandenen Standards bzw. Quasi-Standards, die für die Gesamtheit der denkbaren Workflows für Open-Access-Publikationen herangezogen werden können bzw. müssen, wurde im Artikel „Zur Situation des strukturierten Austauschs von Metainformationen über die Herstellungs- und Distributionsprozesse von Open-Access-Publikationen“<sup>69</sup> dargelegt. Zusammenfassend kann festgestellt werden, dass sich, ähnlich wie sich die natürlichen Sprachen entwickelt haben, artifizielle Sprachen, zu denen Metadaten-Sets unabhängig von ihrer technischen Ausprägung zu zählen sind, unabhängig und domainspezifisch entwickelt haben und derzeit auch rasant weiterentwickeln. Da sich die Domänen aber überlappen, im Zentrum stehen buchartige Dokumentationen, lassen sich substantielle Schnittmengen zwischen den Standards feststellen. Untersucht man im ersten Schritt die Elemente der Standards auf semantische Übereinstimmung, stellt man, aufgeschlüsselt auf einzelne Datengruppen, starke Überschneidungen bis teilweise 100% fest<sup>70</sup>. Dies ist auch nachvollziehbar, da die Foki der Entwicklungen zwar auf domainspezifischen Fragestellungen beruhen, aber grundlegende Metadatenelemente in allen Standards adressiert werden. Beispielfhaft kann dies am semantischen Element „Titel“ nachgewiesen werden, das in allen untersuchten Standards auftrat, was nicht verwundert, werden doch Dokumente in der Regel mit einem Titel gekennzeichnet. Insgesamt wurden in 12 untersuchten Standards lediglich sieben Elemente identifiziert, deren semantische Entsprechungen in allen Standards auftraten. Für die Gesamtheit aller untersuchten Elemente der betrachteten Standards waren dies 7%. Weitere 10 Elemente (10%) waren in 11 der betrach-

---

<sup>69</sup> Reiche et al. 2023

<sup>70</sup> ebd., S. 7, Tabelle 4

teten Standards zu finden<sup>71</sup>. Die Anzahl der gefundenen Übereinstimmungen ist ein Indikator für die semantische Übertragbarkeit von Daten zwischen den einzelnen Standards. Da diese für eine ausreichende Anzahl von Metadatenelementen als gegeben angenommen wurde, wurde eine Untersuchung der technischen Übertragbarkeit durchgeführt. Da die als semantisch gleich festgestellten Metadatenelemente in den einzelnen Standards technisch unterschiedlich codiert werden, kann bei einer Übertragung, trotzdem „das Gleiche gemeint ist“, ein technischer Aufwand an den Schnittstellen anfallen. Diese Untersuchung wurde aus Aufwandsgründen für sieben Standards aus der ursprünglich 12 Standards umfassenden Stichprobe durchgeführt. Es wurde festgestellt, dass die Übertragung von Metadatenelementen mit einer semantischen Übereinstimmung zu 90,85% mit geringem Aufwand übertragen werden können. Werden Transferanforderungen zwischen Elementen einbezogen, in denen keine semantische Übereinstimmung festgestellt wurde, sinkt der Anteil der Transfers mit geringem Aufwand auf 39,32%. Auch das ist nachvollziehbar, weil der Definitionsaufwand für Regeln, welche diese Fälle beschreiben, erheblich höher ist. Gleiches gilt für die anschließende technische Umsetzung. Damit sei kurz zusammenfassend die aktuelle Situation in der Domain der Metadatenstandards beschrieben, auf deren Basis nachfolgend Überlegungen angestellt werden sollen, wie diese mit dem Ziel der Verringerung des Transferaufwandes an den Schnittstellen zwischen den Standards, oder besser zwischen den anwendenden Systemen zu verbessern ist.

### 3.2 Diskussion möglicher Strategien

Grundsätzlich sind hier zwei Strategien denkbar. Strategie 1 (S1) zielt auf die Schaffung einer einheitlichen Kommunikationsstruktur nach dem Muster von JDF ab. Das Ziel wäre die Schaffung einer einzigen und vollständigen Kommunikationsstruktur, welche alle im Feld befindlichen Standards und Quasi-Standards integriert und damit ablöst. Es ergäbe sich ein quasi-geschlossener Standard<sup>72</sup>. Eine Alternativstrategie (S2) kann darin bestehen, dass alle relevanten Metadatenstrukturen bestehen bleiben und projekt- bzw. domainspezifisch die relevanten Metadaten dynamisch zusammengefasst werden, wobei Regelwerke für die jeweiligen Transformationen zur Anwendung kommen. Das Ergebnis wäre ein offener Verbund von Standards und Spezifikationen.

---

71 ebd., S. 7, Tabelle 3

72 Durch die Verwendung offener Technologien, wie XML, würde dieser Standard nicht hermetisch abgeschlossen sein. Die Weiterentwicklung würde aber von einem Konsortium für alle Aspekte vorgenommen werden, was sicher einerseits zu höherer Konsistenz, aber andererseits auch zu behäbigeren Prozessen führen würde.

### 3.2.1 Diskussion von S1

Der naheliegendste Lösungsansatz für das Problem der heterogenen Metadatenstandards liegt darin, die vorhandenen zu einem einzigen zusammenzufassen und damit einen *lingua franca* für die Metadatenkommunikation im Zusammenhang mit Open-Access-Publikationen zu schaffen. Ziel wäre die Schaffung eines vollständigen Sets an Metadaten, das den Informationsbedarf für alle Anwendungsfälle abdeckt. Ein hoher Anspruch, bei dem neben den angeführten Vorteilen auch einige Nachteile zu beachten sind.

Eine Metadatenstruktur mit Vollständigkeitsanspruch wird notwendigerweise komplex. Der Quasi-Standard JDF, welcher dieser angedachten Lösung am Nächsten kommt, hat für einen als etwas weniger umfangreich als die zu betrachtenden Publikationsprozesse einzuschätzenden Definitionsrahmen mehrere tausend semantische Elemente definiert, um alle Produkte, Prozesse, Materialien und Maschinen, die potentiell jeden Herstellungsworkflow beschreiben zu können. Zudem wurden Mechanismen implementiert, um dynamische Elemente, wie Mengen und Zähler, die als endliche Liste nicht darstellbar sind, abbilden zu können<sup>73</sup>. Das Konzept nutzt zur Beschreibung von Druckprodukten einerseits die XML inhärente Baumstruktur aus<sup>74</sup> und hat andererseits in Form von Linkmechanismen implementierte Regeln, um einen Metadatenraum zur Produktbeschreibung aufzuspannen. Wohlgeformtheit und Validität werden durch klassische XML-Parser, die JDF-Richtigkeit durch weitere Passes durch den JDF-Parser gesichert<sup>75</sup>. Um dem Anspruch gerecht zu werden, alle denkbaren Produkte mit ihren Produktionsworkflows über ihre Metadaten abbilden zu können, ohne dass in einer Kommunikationsinfrastruktur alle Metadatenfelder immer betrachtet (d.h. geparsed) werden müssen, wurden Interoperability Conformance Specifications entwickelt, welche die minimalen Anforderungen an eine bestimmte Klassen an Produktionsmaschinen festlegen, so dass Metadaten, welche diese Maschinenklasse nicht betreffen, nicht verarbeitet werden müssen. Dies senkt den Aufwand und die Fehlerquote an den Prozessschnittstellen. Die Pflege dieser ICS übernahmen und übernehmen Arbeitsgruppen<sup>76</sup>.

Zur Entstehungsgeschichte wurde weiter oben bereits referiert. Es ist festzuhalten, dass es starke Marktplayer geben muss, die einerseits die Notwendigkeit und die Synergieeffekte einer solchen vereinheitlichten Kommunikationsstruktur erkennen und andererseits den Willen und die Marktmacht haben, diese durchzusetzen. Zudem muss über lange Zeiträume gedacht werden. Zwischen der Veröffentlichung der Version JDF 1.1 und der aktuellen Version 1.7 im August 2020 liegen fast 20 Jahre.

---

73 vgl. CIP4 Organization, S. 83ff.

74 ebd. S. 22ff

75 ebd. S. 11ff

76 ebd. S. 724

Nicht zu unterschätzen ist weiterhin die Notwendigkeit, eine lebendige Community zu etablieren, welche den (Quasi-) Standard pflegt und weiterentwickelt. Dazu gehört nicht nur die DTD bzw. das XML-Schema, sondern auch die Entwicklung von Tools (im Falle von JDF den JDF-Parser), Entwicklungsumgebungen, Best-Practice-Beispielen usw. Für diese Community sind auch dauerhafte, gemeinschaftliche finanzielle Aufwendungen notwendig. Das Beispiel zeigt aber, dass die Strategie theoretisch umsetzbar wäre.

Abschließend sollen die Situationen verglichen werden, in denen ein Entwicklungsstart vollzogen wurde bzw. werden würde. JDF startete in den 2000er Jahren. Die Entscheidung, auf XML-Technologien zu setzen, war damals sehr innovativ und aus heutiger Perspektive weitsichtig. Es wurde auch von Beginn an eine zentrale Infrastruktur mit Software-Repository und Möglichkeiten der Communitybildung aufgebaut. Da die Entwickler- und später auch die Anwendercommunity weltweit operierte, wurden die Face-to-Face-Treffen durch Telefonkonferenzen erweitert. An die Möglichkeiten, welche heute die Kommunikationssysteme bieten, war noch nicht zu denken. Die aktuellen technischen Infrastrukturen würden sogar bessere Möglichkeiten für eine kooperative Entwicklung bieten.

Während zum Start der Entwicklung von JDF die meisten Workflows in den Druckereien bereits etabliert und beschrieben und damit die semantische Bildung als weitestgehend abgeschlossen betrachtet werden konnte und damit die Voraussetzung einer Schaffung einer vollständigen Metadatenstruktur gegeben war, ist das aktuelle Definitionsfeld, bestehend aus mindestens Verlags- und Bibliotheks-Workflows ungleich heterogener und volatiler. Eine als vollständig zu betrachtende Metadatenstruktur würde wahrscheinlich eine größere Dimension annehmen als die vorliegende JDF-Version 1.7 und würde eine tiefere Struktur aufweisen müssen.

Eine Kommunikationsstruktur sollte, wie auch JDF, ausschließlich dem Transfer der für die Produktion relevanten Metadaten zwischen den einzelnen Workflowteilnehmern realisieren. D.h., die standardisierte Metadatenkommunikation definiert lediglich, dass die empfangenden Systeme einen korrekten Metadatensatz empfangen und verarbeiten bzw. die sendenden diesen generieren und versenden können. Die interne Verarbeitung der Daten in den beteiligten Systemen ist nicht Gegenstand der Standardisierung. Die Daten können sowohl in der nativen Struktur des Standardmetadatenformats gespeichert werden oder aber auch in anderen, proprietären, Datenformaten. Damit relativiert sich das Problem der Altdatenübernahme bzw. -konvertierung, welche ein ernstzunehmendes Hemmnis für die Entscheidung zu Entwicklung einer einheitlichen Kommunikationsstruktur darstellen kann.

### 3.2.2 Diskussion von S2

Als ein Anwendungsfall für die unabhängige Entwicklung von Standards bzw. Quasi-Standards und die anwendungsspezifische Kombination soll der EPUB<sup>77</sup>-Mechanismus für digitale Bücher unter Einbeziehung der Recommendations von HTML<sup>5</sup><sup>78</sup> und CSS<sup>79</sup> betrachtet werden. Betrachtet man den EPUB-Standard genauer, werden zudem weitere Standards, wie XML<sup>80</sup>, DAISY<sup>81</sup>, SVG<sup>82</sup>, Dublin Core<sup>83</sup> und ZIP<sup>84</sup> und anderen implementiert. Zusätzlich werden Pixelgrafikformate und in den neueren Versionen auch Audio- und Video- sowie verschiedene Schriftformate integriert. Alle diese Standards und Quasi-Standards werden unabhängig von der EPUB-Spezifikation gepflegt und von dieser in kontrollierten Versionen übernommen. EPUB bietet damit als ein Rahmenwerk, das verarbeitenden Werkzeugen, hier den Lesealgorithmen für elektronische Bücher, die Möglichkeit gibt, präzise und konsistent auf die Daten zuzugreifen und diese richtig zu interpretieren und im Sinne der Anwendung zu produzieren.

Auch hier findet sich der Ansatz wieder, ein reines Transportformat für Informationen zur Verfügung zu stellen. Die Erzeugung bzw. die interne Verarbeitung der Daten ist, wie bei JDF, nicht Gegenstand der Spezifikation. Damit wird in diesem speziellen Anwendungsfall gesichert, dass der Standard einerseits betriebssystemunabhängig verwendet werden kann, denn eBook-Lesealgorithmen werden auf allen Plattformen angeboten und genutzt, und die große am Markt befindliche Zahl an physischen Lesegeräten mit den unterschiedlichsten Eigenschaften unterstützt wird.

Strategie 2 erfordert also eine Möglichkeit, Produktions- und Metadaten, die verschiedenen Standards und Spezifikationen zugehörig sind, in einer informationstechnischen Struktur zu speichern und zu verwalten. Diese sollte von ihrer Grundanlage her konsistent und robust sein, da auf Anwenderseite beträchtliche Aufwände für deren Integration generiert werden, die nachhaltig sein müssen, andererseits aber flexibel, um auf Änderungen in der hochperformativen Medienwelt reagieren zu können. D.h., es müssen einerseits Speicherstrukturen für nicht-normative Dokumente geschaffen werden, wobei nicht-normativ bedeutet, dass diese Daten nicht dem Standard des Containerformats angehören, sondern anderen und andererseits dem Containerformat zugehörige, standardisierte Dokumente, die den Zugriff auf die nicht-normativen Daten regeln.

---

77 W3C 2023a

78 W3C 2011

79 W3C 2023b

80 W3C 2008

81 DAISY Consortium 2023

82 W3C 2018

83 Dublin Core 2023b

84 ISO 2021

In elektronischen Büchern ist dieser Mechanismus dergestalt implementiert, dass über den normierten Zugriff auf das Dokument `META-INF/container.xml` auf das Verzeichnis mit den nicht-normativen Dokumenten (in der Regel ist dies das Verzeichnis EPBS) verwiesen wird. Der Zugriff auf die nicht-normativen Dokumente wird im Dokument `content.opf` hinterlegt, in dem einerseits die Metadaten der Publikation, im *manifest plane* die Definition aller Ressourcen und im *spine plane* deren buchartige Anordnung enthalten sind<sup>85</sup>. Ähnlich wie bei Strategie 1 müssen, ein XML-basiertes Dokumentenformat vorausgesetzt, zuerst die Anforderungen an Wohlgeformtheit und Gültigkeit geprüft werden. Dafür können die vorhandenen XML-Parser verwendet werden. Soll darüber hinaus auch die strukturelle Richtigkeit geprüft werden, müssen spezielle Validatoren nach dem Vorbild des JDF-Parsers angewendet werden. Für elektronische Bücher sind diverse am Markt vorhanden<sup>86</sup>. Für eine strukturiertes Kommunikationsformat für OA-Publikationen müsste ein solches Validierungswerkzeug entwickelt werden. Dafür böte sich zum Beispiel Schematron<sup>87</sup> an.

### 3.2.3 Konklusion und Extraktion der Anforderungen

Beide Strategien weisen eindeutige Vor- und Nachteile aus. Aus der Diskussion im Stakeholder-Workshop kristallisierte sich aber als Mehrheitsmeinung als realistisch umsetzbare Variante Strategie 2 heraus. Insofern soll die Extraktion der Anforderungen an eine Kommunikationsstruktur auf der Grundlage dieser Strategie erfolgen, wobei sinnvolle Elemente der Strategie 1 integriert werden sollen. Nachfolgend sollen die Aspekte diskutiert werden, die als Hauptanforderung an eine Kommunikationsstruktur nach Strategie 2 umgesetzt werden sollten.

#### Technologischer Unterbau

Es kann vorausgesetzt werden, dass eine Kommunikationsstruktur, die sich in die bestehenden informationstechnischen Infrastrukturen einbetten soll, die aktuellen Technologien für Datenaustauschformate, wie XML oder der ECMA<sup>88</sup>-Standard JSON<sup>89</sup>, verwenden sollte. Beide Technologien sichern die Forderungen, menschen- und maschinenlesbar sowie betriebssystem- und programmiersprachenunabhängig zu sein und ohne Zusatzaufwand erstellbar zu sein, sprich im Textformat vorzuziehen. Für JSON spricht die Schlankeheit des Datenaustauschformats, für XML der höhere Verbreitungsgrad. Letzteres gibt für den zu entwickelnden Ansatz den Ausschlag, denn

---

85 siehe §3.1 in W3C 2023a

86 z.B. EPUB-Check in W3C 2023c

87 Schematron 2021

88 ECMA International ist eine private, internationale Standardisierungsorganisation.

89 JSON 1999; ECMA International

die meisten der in einem OA-Workflow zu beachtenden Standards<sup>90</sup> basieren auf XML-Technologien, was bei einer späteren Implementierung der Transformationsregeln als vorteilhaft angesehen wird.

### **Metadaten und Produktionsdaten**

Grundsätzlich sind für eine strukturierte Kommunikation zwei Vorgehensweisen bzgl. der Produktionsdaten zu beachten. Diese können einerseits integraler Bestandteil der Kommunikationsstruktur sein und mit dieser übertragen werden. Es läge dann ein Containerformat nach dem Vorbild von EPUB vor. Alternativ wäre es möglich, lediglich die Metadaten in einer Kommunikationsstruktur zu übertragen und darin nur die Möglichkeit der Verlinkung auf die Produktionsdaten (Ressourcen) vorzusehen. Diese Vorgehensweise liegt zum Beispiel JDF zugrunde, dass eine reine Metadatenstruktur zur Verfügung stellt und Produktionsdaten ausnahmslos extern verlinkt. Die Diskussionen während des Stakeholder-Workshops haben ergeben, dass die Schaffung eines Containerformats die bevorzugte Lösung darstellt. Eine Begründung liegt darin, dass derzeit (noch) nicht standardisierte Inhalte, wie Peer-Review-Dokumente oder Messdaten, welche im Zusammenhang mit OA-Publikationen eine wichtige Rolle spielen, ebenfalls in dem Container gespeichert und transportiert werden sollen.

### **Datenablagestruktur**

Die Datenablagestruktur muss in ihrer Ablage robust und zukunftssicher sein. Komplexität spielt eine untergeordnete Rolle. Ähnlich wie bei EPUB und DCF<sup>91</sup> soll eine standardisierte Ordnerstruktur mit ebensolchen bezeichnungsregelnden bearbeitenden Systemen konsistenten Zugriff auf die Meta- und die Produktionsdaten gewährleisten. Dabei können die Zugriffsregeln externalisiert oder über ein Dokument, das die Regeln des Zugriffs auf die Meta- bzw. Produktionsdaten enthält, geregelt werden. Für die Kommunikationsstruktur wird der letzte Fall, nach dem Vorbild von EPUB, vorgeschlagen.

### **Validierungswerkzeug**

Unter der Voraussetzung, dass eine Metamarkupsprache zur Konstruktion der Kommunikationsstruktur verwendet wird, kann für die Sicherstellung von Wohlgeformtheit und Gültigkeit ein vorhandener Parser, im Falle von XML zum Beispiel der Xerces<sup>92</sup>, verwendet werden. Dieser kann auch eingesetzt werden, um XML-basierte Daten, die anderen DTD oder Schemata entsprechen, zu prüfen. Da die Kommunikationsstruktur aber lediglich einen Container zum Transport der Daten vom Sender zum Empfänger darstellen soll, ist deren Prüfung nicht zwingend notwendig, sondern obliegt

---

90 vgl. Böhm 2021

91 Design rule for Camera File system, vgl. JEITA

92 vgl. Xerces apache 2022

dem schreibenden Werkzeug beim Sender, das dafür sorgen muss, dass ausschließlich korrekte Daten in den Container eingebracht werden. Die empfangende Seite kann die sie betreffende Daten ebenfalls kontrollieren und im Falle einer fehlgeschlagenen Prüfung zurückweisen. Das heißt, ein Validierungswerkzeug muss lediglich Wohlgeformtheit, Gültigkeit und der inneren strukturellen Integrität der im Definitionsbereich der Kommunikationsstruktur befindlichen Metadaten sicherstellen.

### **Operabilität**

Wie in Abschnitt 2.1.3 ausgeführt, wird der höchste Operabilitätsgrad mit der Interoperabilität erreicht. Die Kompatibilität zwischen zwei Systemen kann mit Hilfe von Regeln an der Schnittstelle zwischen zwei beteiligten Systemen gesichert werden. Dies stellt den aktuellen Zustand in der Kommunikation von Systemen in OA-Publikationsworkflows dar. Sind mehr als zwei Systeme beteiligt, und dies ist häufig der Fall, erhöht sich die Anzahl der Schnittstellen und damit der Implementierungsaufwand an diesen. Abhilfe kann hier ein de-facto-Standard schaffen. Diesen Operabilitätsgrad zu erreichen, sollte Minimalanforderung sein. Fernziel sollte aber die Erreichung von Interoperabilität sein.

### **Schnittstellen**

Die Kommunikationsstruktur muss mannigfaltige Schnittstellen bedienen. Als Ausgangspunkt sollen die in Kapitel 2.3.1 diskutierten Schnittstellen sein, die sich aus dem allgemeinen Workflowmodell für OA-Publikationen ableiten lassen. Der akademische Ansatz wird diese auf Grund des Umfangs nur exemplarisch respektieren können.

### **Abgrenzung**

Die Kommunikationsstruktur bildet lediglich einen Container, der von der sendenden Infrastruktur angelegt und mit Meta- und Produktionsdaten befüllt wird. Die Vorgehensweise dafür unterliegt nicht der Standardisierung. Sicherzustellen ist lediglich die Validität der Daten im Container. Gleiches gilt für die empfangende Infrastruktur. Sie setzt einen validen Datenbestand voraus und erwartet einen regelgerechten Zugriff auf die Daten. Art und Weise der Datenverarbeitung in der empfangenden Infrastruktur unterliegen ebenfalls nicht der Standardisierung und werden von dieser individuell geregelt.

### **Community**

Die stellvertretend für die Strategie diskutierte Entwicklung von JDF hat gezeigt, dass für die Erstellung eines vollständigen und quasi-hermetischen Standards eine starke Community notwendig ist. Diese war durch das Engagement von vier Global Playern gegeben. Zudem erscheint die Druckbranche weniger differenziert als die Verlags- bzw. Bibliotheksbranche. Eine hermetische Community nach dem Vorbild von CIP4 wird

also nicht zu schaffen sein. Mit dem W3C global und der IG Produktmetadaten<sup>93</sup> und diversen Initiativen im Bibliotheksbereich, wie der Arbeitsgemeinschaft der Verbundsysteme<sup>94</sup>, die AG Universitätsverlage<sup>95</sup> und anderer im deutschsprachigen Bereich gibt es bereits verschiedene Einzelinitiativen, die, wie bereits oben diskutiert, heterogen aufgestellt sind, so dass ein Vorgehen nach dem Vorbild von CIP4 nicht zielführend erscheint. Jedoch wäre es angebracht, dass neben den selbstorganisierenden Kräften in der Branche vor allem integrierende Impulse von Fördermittelgebern durch zielgerichtete Steuerung der Fördermittel eingebracht werden.

Basierend auf den beschriebenen Strategien und Anforderungen an eine Kommunikationsstruktur formulierte das Forschungsteam folgende Arbeitsdefinition:

*Eine Kommunikationsstruktur ist eine allgemein gültige, erweiterbare Struktur, die der bidirektionalen Informationsübertragung von publikations- und produktionsbezogenen Daten zwischen prozessbeteiligten Systemen dient und einen effizienten sowie automatisierten Produktions- und Distributionsworkflow von OA-Veröffentlichungen sicherstellt.*

Neben den dargestellten Anforderungen sind weitere, allgemeine Kriterien für die Entwicklung einer Kommunikationsstruktur von Bedeutung. Eines der wesentlichsten Kriterien ist der unbeschränkte, kostenfreie Zugang zur Struktur, um Möglichkeiten der Weiterentwicklung projektunabhängig zu gewährleisten und um die damit einhergehende Herstellerunabhängigkeit, welche für die Integration in verschiedene Prozesse und Workflows notwendig ist, sicherzustellen. Die Struktur wurde daher als Open-Source-Lösung gedacht. Da die Struktur unabhängig von individuellen Workflows und im Prozess beteiligter Systeme Einsatz finden soll, gilt als weiteres Kriterium die Allgemeingültigkeit. Darüber hinaus soll die Struktur als Werkzeug einfach in der Anwendung sein, das heißt, auch ohne herstellerische Kenntnisse nutzbar sein. Die Struktur soll problemlos und ohne Installationsaufwand in Publikationsworkflows integriert werden können. Die Kommunikationsstruktur hat den Anspruch, alle für den OA-Publikationsprozess relevanten Metadaten zu integrieren.

---

93 vgl. Börsenverein

94 vgl. DNB 2022

95 vgl. ag universitätsverlage



## 4 Akademischer Ansatz

---

Im Rahmen des Forschungsprojektes soll keine vollständige Implementierung des Ansatzes erfolgen. Darzustellen ist ein aus den oben diskutierten theoretischen Grundlagen abgeleiteter Ansatz, der als Denkmodell für weitere Entwicklungen dienen kann. Die Entwicklung von JDF hat gezeigt, dass eine leistungsfähige Lösung nur durch eine lebendige Community mit einer Entwicklungsperspektive von Dekaden erreicht werden kann. Neben den Unternehmen müssen auch Forschungseinrichtungen integriert werden und ihren Beitrag leisten. Ergänzend zu den oben ausgeführten theoretischen Grundlagen soll der nachfolgende Ansatz als Nukleus weitergehender Entwicklungen betrachtet werden.

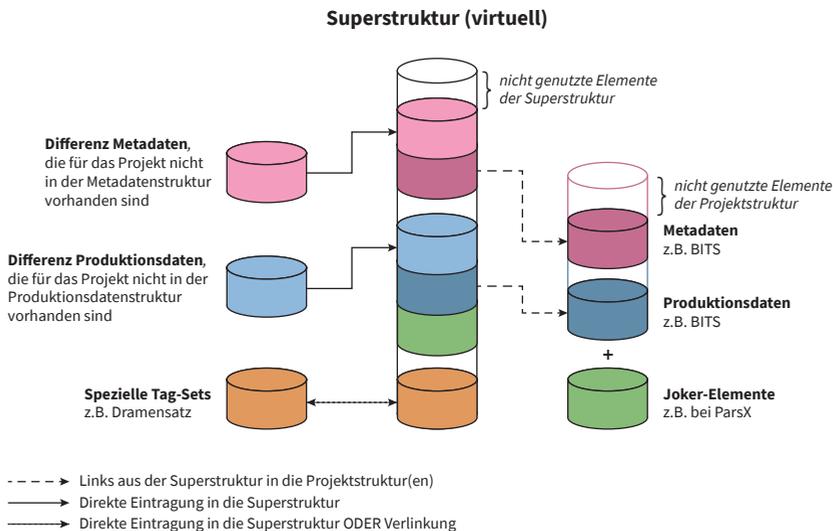
Wie in Kapitel 3 dargelegt, wurde für die Entwicklung eines akademischen Ansatzes einer Kommunikationsstruktur als Grundlage das Denkmodell eines Containerformats herangezogen. So kann sichergestellt werden, dass, gleich einer Auftragstasche, alle für den Herstellungs- und Distributionsprozess notwendigen Dokumente, Daten, Metadaten und Standards an einem Ort hinterlegt und Informationen möglichst einfach ausgelesen werden können. Gleichzeitig bietet ein solches Format Möglichkeiten der freien Erweiterbarkeit und individueller Anpassung.

### 4.1 Aufriss der Idee

Die Grundidee besteht darin, dass zuerst entsprechende der oben formulierten Abgrenzungsbedingungen (Kap. 3.2.3) eine sendende Infrastruktur, zum Beispiel ein Workflowmanagementsystem (WFMS) in einer Bibliothek und eine empfangende Infrastruktur (zum Beispiel ein Repository) anzunehmen sind. Beide Instanzen haben unterschiedliche Sichten auf Meta- und Produktionsdaten in verschiedenen Phasen eines Workflows für OA-Publikationen. Für beide Infrastrukturen wird vorausgesetzt, dass sie die Daten in strukturierter Form halten, wobei die technische Umsetzung dieser Struktur unerheblich ist. Da wie oben ausgeführt (Kap. 3.1) die semantischen Übereinstimmungen in den untersuchten Standards und Quasi-Standards sehr hoch ist, ist es möglich, eine virtuelle Super-Struktur zu denken, die in erster Näherung alle möglichen semantischen Elemente in Form einer domänenspezifischen Ontologie oder, im einfacheren Fall, einer (möglichst abgeschlossenen) Liste enthält. D.h., jedes Element eines beteiligten Standards bzw. Quasi-Standards referenziert auf einen Platzhalter in der Superstruktur, der dessen semantische Bedeutung repräsentiert. Elemente aus verschiedenen Standards, welche die gleiche semantische Bedeutung haben, referenzieren also auf den selben Platzhalter und sind wiederum über diesen in den einzelnen

Strukturen erreichbar. Da die Kommunikationsstruktur projektbezogen aufgebaut wird, ist anzunehmen, dass in Abhängigkeit der teilnehmenden (sendenden und empfangenden) Infrastrukturen verschiedene Standards und Quasi-Standards verwendet werden. Diese können über den Mechanismus der gedachten unendliche Liste in theoretisch jeder Kombination integriert werden.

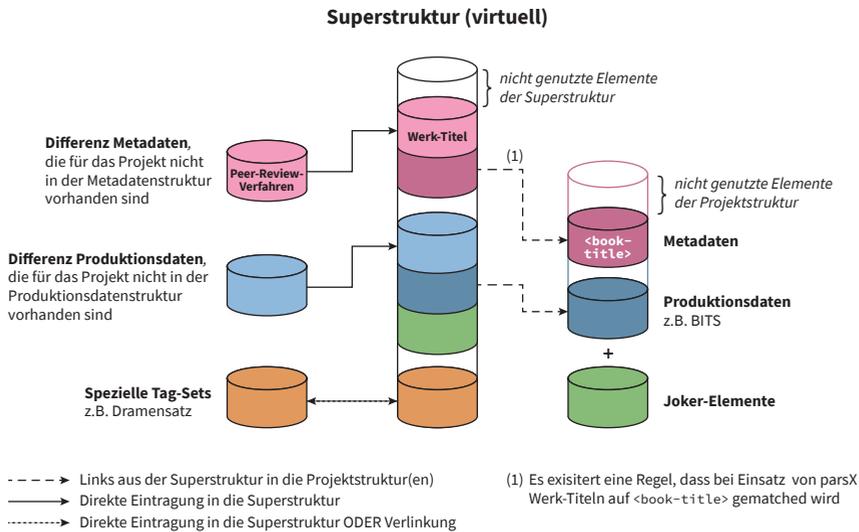
Abbildung 23 illustriert diese Idee. Es wird davon ausgegangen, dass in einer projektspezifischen Kommunikationsstruktur ein „Leitstandard“ verwendet wird, sinnvollerweise jeweils jener, den die sendende Infrastruktur verwendet. Für das Beispiel in der Abbildung wird BITS angenommen. Es wird sich folgender Zustand einstellen: Es werden a) Elemente der Struktur mit meta- bzw. Produktionsdaten befüllt sein, es werden b) in der Struktur vorgesehene Elemente für das spezifische Projekt nicht verwendet und bleiben daher leer (nicht genutzte Elemente der Projektstruktur) und c) wird es Meta- bzw. Produktionsdaten geben, die projektspezifisch zu transportieren sind, aber im Leitstandard (hier BITS) nicht definiert sind (Differenz Meta- bzw. Produktionsdaten). Zusätzlich können sehr seltene Elemente (zum Beispiel für den Dramensatz) auftreten. Einige Standards bzw. Quasi-Standards benutzen für die Abdeckung des Falls c) sogenannte Joker-Elemente, die ebenfalls berücksichtigt werden müssen.



**Abb. 23** Grundsätzlicher Ansatz einer auf einer „Superstruktur“ basierenden Kommunikationsstruktur

Der Mechanismus für die Bildung der projektspezifischen Containerstruktur unterscheidet also folgende Fälle:

1. Das in der virtuellen Superstruktur definierte semantische Element ist in einer oder mehreren beteiligten Standards oder Quasistandards befüllt. Die Informationen verbleiben in den jeweiligen Strukturen und werden aus der Superstruktur verlinkt.
2. Es müssen Informationen transportiert werden, welche keine semantische Entsprechung im Leitstandard haben. Diese werden direkt in die Superstruktur „eingetragen“<sup>96</sup>.
3. Projektspezifisch nicht genutzte Elemente des Leitstandards werden nicht verlinkt. Damit wird ein einfacher Mechanismus etabliert, der es ermöglicht, bestehende, strukturiert abgelegte Datenbestände nicht „umschlüsseln“ zu müssen, sondern direkt aus der sendenden Infrastruktur zu übernehmen.



**Abb. 24** Einsatz von Regeln bei der Verknüpfung semantischer Elemente

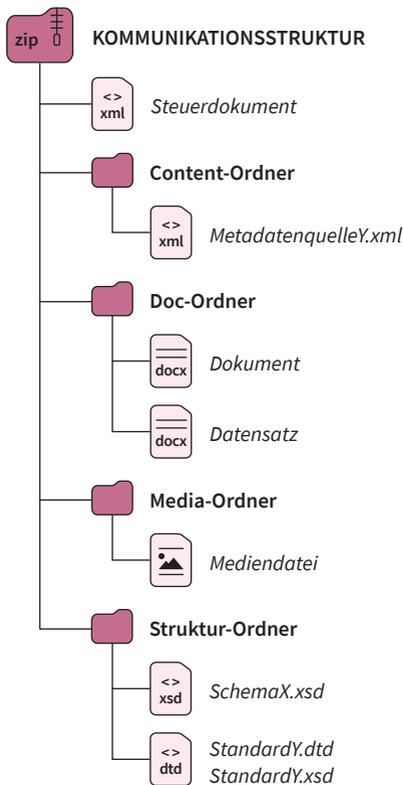
Abbildung 24 zeigt ein Beispiel für das Referenzieren des semantischen Oberbegriffs (hier Werk-Titel) auf das konkrete Datenfeld des in der Kommunikationsstruktur verwendeten Leitstandards (hier BITS) über eine Regel. Die Anforderungen an die Transformation wurden in Kapitel 2.2 beschrieben. Die einzelnen Regeln müssen für die jeweiligen beteiligten Strukturen implementiert werden. Im Beispiel ist die Regel einfach, denn es liegt sowohl technische, wie auch strukturelle ( $\Delta$ ), syntaktische

96 siehe S. 78

sowie semantische Integrität vor (siehe auch Kap. 2.2.1). Der Eintrag in die Liste an die Position, die dem semantischen Begriff des Werkstitels entspricht, enthält also lediglich den URI auf das Element `<book-title>` in der Instanz, die der BITS-DTD entspricht. (101: Content/instanz.xml/book/book-meta/book-title-group/book-title).

## 4.2 Aufbau und Funktionsweise

Wie sich das Containerformat der Kommunikationsstruktur zusammensetzt, ist in Abbildung 25 skizziert: Der Container, der alle Informationen enthalten soll, wird durch den ZIP-Ordner „Kommunikationsstruktur“ dargestellt. In diesem befinden sich alle Informationen, Dokumente und Daten, die zu einem Projekt vorliegen. Zu jedem Projekt wird demnach ein eigener Ordner erstellt. Der übersichtliche und einfache Aufbau soll das Anlegen eines Projektes auch ohne tiefgehende technische und herstellerische Kenntnisse ermöglichen.



**Abb. 25** Aufbau der Kommunikationsstruktur als Containerformat

Jegliche Inhalte und Informationen eines OA-Projekts werden den Ordnern „Content“, „Doc“, „Media“ und „Struktur“ zugeteilt. Die Originaldokumente, welche mit einer OA-Publikation verbunden sind, werden im Ordner „Doc“ hinterlegt. Dabei handelt es sich um die Ausgangsversionen, bspw. im .docx- oder .xlsx-Format. Medien, welche in die Dokumente eingebettet wurden, befinden sich im „Medien“-Ordner. Der Ordner „Content“ wiederum dient der Sammlung aller Produktions-Metadatenquellen. Da die Struktur die Kombination verschiedener Metadatenstandards ermöglichen soll, wird je Standard eine Metadatenquelle erstellt. Die zu den Standards gehörenden Document-Type-Definitions (DTD) und XML-Schema-Definitions (XSD), welche die Struktur und Inhalte dieser definieren, befinden sich im Ordner „Struktur“. Auch individuell erstellte Schemata können dort hinterlegt werden. Um die einzelnen Informationen und Daten aus den jeweiligen Ordnern zu verbinden, ist ein „Steuerdokument“ notwendig. Dieses befindet sich wie die beschriebenen Ordner in der zweiten Ebene. Hierbei handelt es sich um eine XML-Datei, die sich ähnlich zur OPF<sup>97</sup>-Datei eines EPUBs verhält. Darin sind alle Dokumente, Medieninhalte, Standards und Metadaten referenziert, welche für die OA-Publikation relevant sind.

Ein Ziel der Kommunikationsstruktur ist die Kombination verschiedener Standards, um alle notwendigen Metadaten abbilden zu können. Für jedes Metadatum muss demnach ein entsprechender Metadaterm eines Standards zugeordnet werden. Das Steuerdokument fungiert hierbei als eine Art „Matching-Dokument“, welches zum einen die Metadaten mit den Metadatermen aus der entsprechenden Metadatenquelle verbindet und zum anderen Metadatenquellen den dazugehörigen Standard-Schemata zuordnet. Das Steuerdokument besteht grundsätzlich aus sieben Bestandteilen, welche sich ebenfalls an der OPF-Datei des EPUBs orientieren:

Zu Beginn steht die XML-Deklaration mit der Angabe des Schemas. Anschließend folgen in Abschnitt zwei die Projekt-Metadaten, die den aktuellen Produktionsstand (1) sowie einen Zeitstempel (2) enthalten:

```
<projekt-metadaten>
  <historie>
    <event type="Prozess_A"(1) datum="YYYY-MM-DD"(2)/>
  </historie>
</projekt-metadaten>
```

Der dritte Abschnitt enthält Referenzen zu den verschiedenen Metadatenquellen (4) im Ordner „Content“ und ordnet diesen das entsprechende Schema (5) aus dem Ordner „Struktur“ zu. Dabei wird je eine numerische ID (3) vergeben.

---

97 Open Packaging Format

```
<content>
  <instanz id="B"(3) href="Pfad_Metadatenquelle_B"(4)
    schema="Pfad_Schema_B"(5)/>
</content>
```

Der größte Abschnitt folgt an vierter Stelle und besteht aus einem Matching der Metadaterme mit den entsprechenden Metadatenknoten aus den Metadatenquellen. Für jedes Metadatum (7) wird hierbei der XPath des entsprechenden Metadatenknotens (9) im jeweiligen Quelldokument (8) zugeordnet. Das jeweilige Quelldokument wird über die Angabe der zuvor vergebenen ID (3) referenziert. Zudem wird jedem Metadatum eine Nummer (6) zugeordnet, um beim Auslesen der Metadaten möglichst einfach und fehlerfrei darauf zugreifen zu können.

```
<meta var="C"(6) name="Metadatum"(7)
id_ref="ID_Metadatenquelle_B"(8)>
XPath_des_Metadatenknotens(9)
</meta>
```

In Abschnitt fünf werden wiederum alle Dokumente, welche dem Werk zugehörig sind, referenziert (11) und mit IDs (10) versehen:

```
<dokument id="D"(10) href="Pfad_Dokument_D"(11)/>
```

Alle Medieninhalte, wie bspw. Bilder, werden im darauffolgenden Abschnitt sechs hinterlegt, unter Angabe von ID (12), Typ (13) und Referenzlink (14):

```
<medium id="E"(12) type="img"(13) href="Pfad_Medium_E"(14)/>
```

In den Abschnitten sieben und acht wiederum werden Dokumente zu Peer-Review (15) und Forschungsdaten (17) zugeordnet, indem die ID (3) der entsprechenden Instanzen aus Abschnitt drei referenziert wird (16):

```
<peer-review>(15)
  <gutachten id_ref="ID_Dokument"(16)/>
</peer-review>
<forschungsdaten>(17)
  <messung id_ref="ID_Dokument"(18)/>
</forschungsdaten>
```

## 4.3 Anwendungsszenarien

Der oben beschriebene Ansatz soll nun auf seine Anwendbarkeit überprüft werden. Dazu werden nachfolgend einige Anwendungsfälle beschrieben, welche typischerweise über die Kommunikationsstruktur ausgeführt werden. Für diese müssen die jeweiligen Regelwerke (Stylesheets) entwickelt werden. Die Anwendungsfälle werden entlang des Lebenszyklus' eines Meta- bzw. Produktionsdatensatzes einer OA-Publikation entwickelt.

### 4.3.1 Produktlebenszyklus

Der Produktlebenszyklus einer OA-Publikation ist in Abbildung 26 skizziert und wird in den nachfolgenden Kapiteln beschrieben.

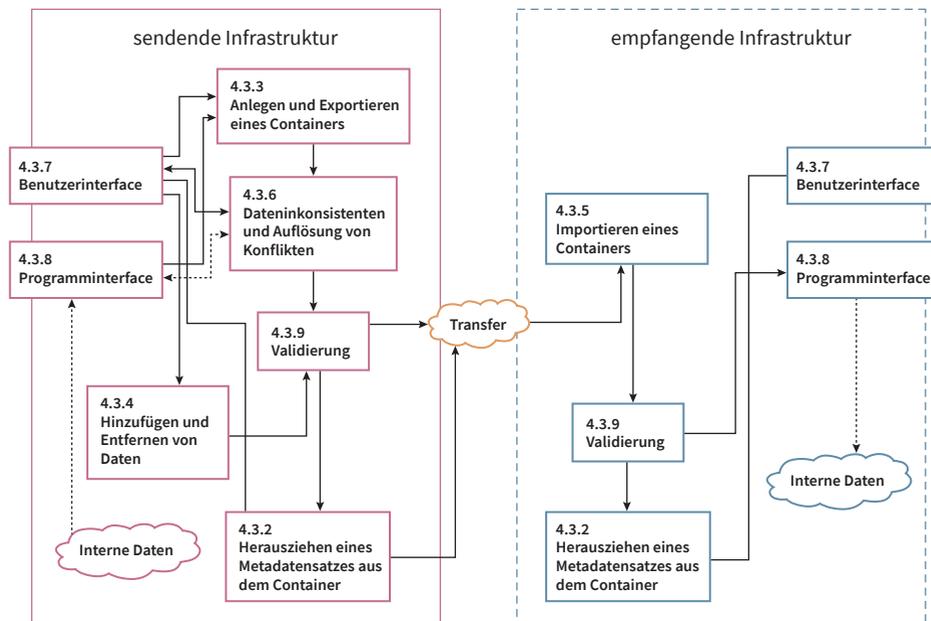


Abb. 26 Produktlebenszyklus einer OA-Publikation

### 4.3.2 Herausziehen eines Metadatensatzes aus dem Container

Der wichtigste Anwendungsfall ist offenbar die Extraktion von speziellen Metadatensätzen aus den gesammelten Metadaten des Containers. Es muss also gewährleistet werden, dass jegliche Anfrage an den Metadatenbestand (und konsequenterweise an den gesamten Datenbestand) gestellt werden kann. Ein einfacher konkreter Anwendungsfall wäre die Extraktion von Autor und Titel. Ein spezieller, aber häufiger Anwendungsfall wäre die Extraktion von Metadaten und die Generierung eines (standardisierten) Metadatensatzes für die Übermittlung der Daten an ein Repository. Eine Anfrage könnte daher lauten: Extrahiere alle Daten, die für die Übermittlung des Werkes an ein Repository notwendig sind und generiere dafür einen Metadatensatz, z.B. im ONIX-Format. Für diesen Use Case wird von einer erfolgreichen Transaktion ausgegangen. Im Fehlerfall soll, wie in Abschnitt 4.3.6 beschrieben, vorgegangen werden.

### 4.3.3 Anlegen und Exportieren eines Containers

Ein Container kann interaktiv (siehe Abschnitt 4.3.7) oder durch ein erstellendes Programm angelegt werden. Es müssen die Verzeichnisstruktur und mindestens die Steuerungsdatei (`projekt.xml`) angelegt werden. Das Exportieren besteht aus dem Erstellen des Containerformats (`*.zip`). Der Transfer von der erstellenden Stelle zur empfangenden ist nicht Bestandteil der Spezifikation.

### 4.3.4 Hinzufügen und Entfernen von Daten

Voraussetzung für diesen Anwendungsfall ist ein in Use Case 4.3.2 angelegtes Containerformat. Das Hinzufügen und Entfernen von Daten kann interaktiv (Abschnitt 4.3.7) oder automatisiert (4.3.8) erfolgen. Die Daten können entweder in einen geschlossenen Container eingefügt oder extrahiert werden oder der Container wird geöffnet und die Daten mittels Dateisystemoperationen manipuliert.

### 4.3.5 Importieren eines Containers

Das Importieren eines Containers obliegt vollständig dem importierenden System. D.h., das Öffnen des komprimierten Containers, der Zugriff auf die Daten und die Überführung in die Datenstrukturen des Systems sind nicht Bestandteil der Spezifikation.

### 4.3.6 Dateninkonsistenzen, Auflösung von Konflikten

Beim Zusammenführen verschiedenen Datenquellen mit Schnittmengen in den Metadatensätzen sind mögliche Dateninkonsistenzen zu berücksichtigen. Mögen die einzelnen Datensätze auch in sich wohlgeformt und gültig sein, können sich zwangsläufig in deren Kombination strukturelle Inkonsistenzen ergeben.<sup>98</sup> Diese müssen zum einen festgestellt werden. Die Feststellung kann beim Hinzufügen von Daten in den Container (Kap. 4.3.2) erfolgen und nach Änderungen von Containerinhalten durchgeführt werden (4.3.9).

### 4.3.7 Benutzerinterface

Für die Generierung der Anfragen soll ein plattformunabhängiges Benutzerinterface benutzt werden können. Mit dessen Hilfe soll es Nutzern möglich sein, interaktiv, durch Auswahlmöglichkeiten die in 4.3.1 beschriebenen Anfragen an den Datencontainer zu stellen. Die Ergebnisse sollen in verschiedenen Formaten, wie zum Beispiel als Excel-Liste, CSV und den wichtigsten Metadatensätzen für Repositorien exportiert werden können. Weiterhin sollen Nutzer mit Hilfe des Interfaces in der Lage sein, neue Metadaten und Nutzdaten in den Container zu importieren. Das Nutzerinterface soll NICHT dazu dienen, einzelnen Metadaten im Container anzulegen, zu ändern oder zu löschen (Kap. 4.3.3).

### 4.3.8 Programminterface

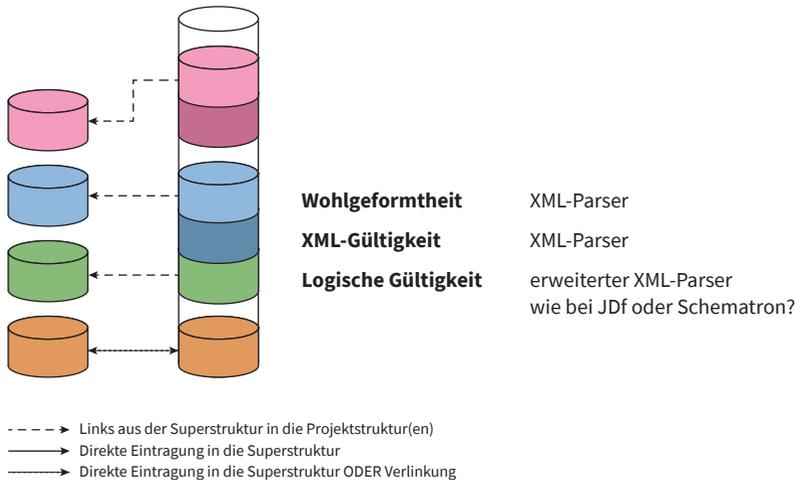
Externe Anwendungen benötigen eine Programmierschnittstelle, welche die gleichen Funktionalitäten anbietet, wie das in 4.3.7 beschriebene Benutzerinterface.

### 4.3.9 Validierung

Es soll möglich sein, eine Konsistenzprüfung des Containers vorzunehmen. Konkret bedeutet das, dass die zentrale Steuerungsdatei (`projekt.xml`) traversiert wird und alle Einträge auf innere Kohärenz und auf Korrektheit der Links in den Container überprüft wird. Werden Fehler festgestellt, wird wie in Abschnitt 4.3.6 vorgegangen.

---

98 Sannemann S. 45ff.



**Abb. 27** Einsatz von Validierungswerkzeugen zur Qualitätssicherung der Kommunikationsstruktur

#### 4.3.10 Optionale Anwendungsfälle

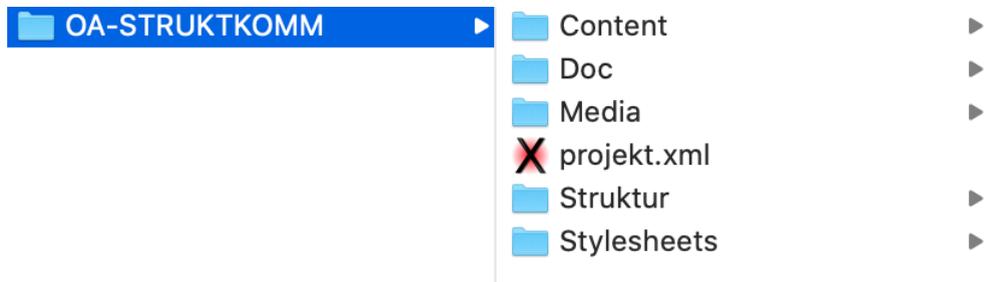
Es wäre denkbar, das Containerformat durch einen zusätzlichen Ordner zu erweitern, in den bei jedem Importvorgang von Daten bereits die standardisierten Metadatensätze angelegt/upgedated werden. Der Vorteil bestünde darin, dass im Bedarfsfalle nicht erst auf die Anfrage eine Generierung (Abschnitt 4.3.1) angestoßen werden müsste, sondern direkt auf die fertigen Datensätze zugegriffen werden kann. Diese Vorab-Generierung müsste bei jeder Manipulation (Hinzufügen, Entnehmen, Ändern) der Daten, sowohl durch das Benutzerinterface oder das Programminterface vorgenommen werden.

### 4.4 Beispielhafte Umsetzung

Nachdem die Idee erörtert wurde, soll nun nachgewiesen werden, dass diese, zumindest im begrenzten theoretischen Rahmen, umsetzbar ist. Dazu wurde eine Entwicklungsumgebung aufgesetzt, die es ermöglicht, einzelne Anwendungsfälle des Dokumentenlebenszyklus für eine projektspezifische Publikation zu testen.

#### 4.4.1 Vorschlag für die Struktur der Containerinstanz für die strukturierte Kommunikation

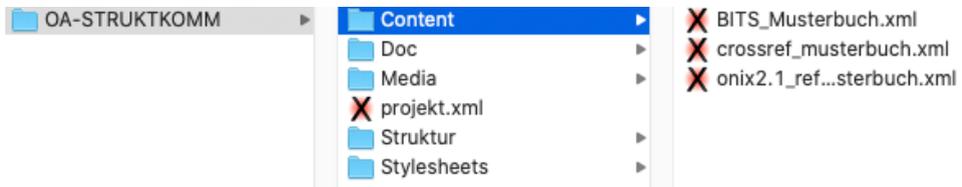
Die Funktionsweise des Containers für die strukturierte Kommunikation wird nachfolgend beispielhaft vorgestellt. Der oben ausgeführte Implementierungsvorschlag kann folgendermaßen umgesetzt werden. In einem Verzeichnis (im Beispiel OA\_SRUKTKOMM), werden die Dateien und Verzeichnisse, welche später in einen Container verpackt werden, angelegt. Dies entspricht den Anwendungsfällen 4.3.7, wenn die Struktur interaktiv über ein Benutzerinterface angelegt wird bzw. 4.3.8, wenn die sendende Infrastruktur unter Verwendung von Funktionen, welche durch ein Programminterface zur Verfügung gestellt werden, angelegt werden. In dem Verzeichnis müssen mindestens die Projektsteuerdatei (`projekt.xml`) sowie die Unterverzeichnisse `Content` für strukturierte Instanzen, `Doc` für alle nicht standardisierten Dokumente wie Peer-Reviews, Messdaten etc., `Media` für alle Mediendateien, wie Bilder, Audio- und Videocontent etc., `Struktur` für alle verwendeten DTDs und Schemata, die zur Validierung der Daten notwendig sind, sowie `Stylesheets`, in dem alle für das Projekt relevanten Transformationsstylesheets abgespeichert werden. Eine Übersicht über die grundlegende Verzeichnisstruktur bietet Abbildung 28.



**Abb. 28** Grundsätzliche Struktur eines Containers für die strukturierte Kommunikation von Daten und Metadaten für OA-Publikationen

Zentrales Dokument ist die Projektsteuerdatei `projekt.xml`. Sie gleicht in ihrer Funktionalität der Instanz `content.opf` in einem EPUB-Container. Die Projektsteuerdatei ist direkt im Wurzelverzeichnis des Projektes angeordnet. Auf den Verlinkungsmechanismus über `META-INF/container.xml` wird verzichtet.

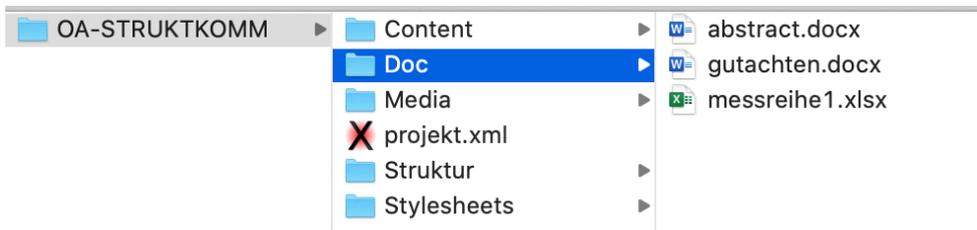
Als Inhaltsdaten werden in dem beschriebenen Beispielprojekt folgende verwendet: Im Verzeichnis „Content“ werden drei Musterinstanzen gespeichert, welche strukturierte Projektdaten gemäß verschiedenen Standards bzw. Quasistandards enthalten. Konkret handelt es sich um eine BITS-Instanz (`BITS-Musterbuch.xml`), eine Crossref-Instanz (`crossref-Musterbuch.xml`) und eine ONIX-Instanz (`ONIX-Musterbuch.xml`).



**Abb. 29** Musterinstanzen im Content-Verzeichnis des Containers

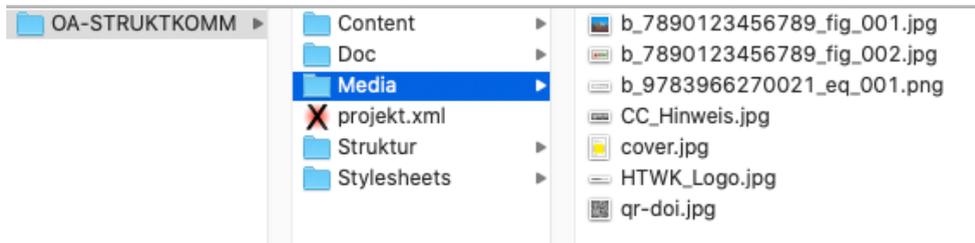
Das BITS-Musterbuch entspricht dabei der strukturierten Instanz der Monografie „Open-Access-Publikationsworkflow für akademische Bücher“, das als Ergebnisdokumentation des Projekts OA-HVerlag verfasst wurde.

Im Verzeichnis „Doc“ können alle projektspezifischen Daten, die nicht standardisierten strukturierten Datenformaten entsprechen, abgespeichert werden. Diese Daten werden im Container nur transportiert, sie unterliegen aber keinen Qualitätssicherungsprozessen im Sinne einer standardisierten Kommunikation. Beispielfhaft werden ein Abstract (`abstract.doc`), ein Peer-Review-Gutachten (`gutachten.doc`) und die Dokumentation einer Messreihe (`messreihe1.xlsx`) dort abgespeichert.



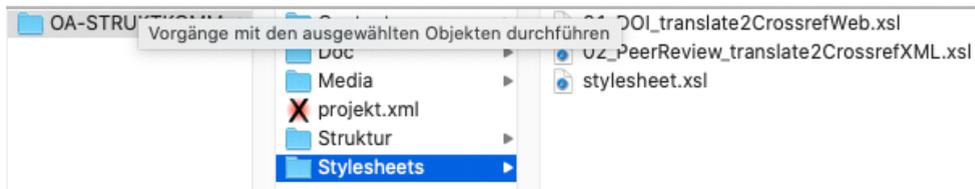
**Abb. 30** Verzeichnis für nicht standardisierte Daten

Alle Mediendateien, welche aus den strukturierten Dokumenten im Content-Verzeichnis referenziert werden, können im Verzeichnis „Media“ hinterlegt werden. Dieses Verzeichnis stellt das Standard-Verzeichnis dar. Referenzieren die Mediendaten aus den strukturierten Daten mit URI in Form von relativen Pfadangaben auf gleicher Ebene oder als Unterverzeichnis vom Content-Verzeichnis, kann diese Struktur beibehalten werden. Damit soll es ermöglicht werden, ein Publikationsprojekt direkt in den Container zu kopieren. Eine Referenzierung der Mediendaten innerhalb der Projektdatei muss dann nicht erfolgen (siehe S. 76).



**Abb. 31** Verzeichnis für Mediendaten

Obwohl die Generierung und Manipulation von Metadaten im Container nicht standardisiert ist und den jeweiligen Infrastrukturen obliegt, besteht die Möglichkeit, Stylesheets, welche projektspezifisch zwischen diesen ausgetauscht werden sollen oder müssen, im Verzeichnis „Stylesheets“ abzulegen. Ein Mechanismus zur Erklärung der Verwendung der Stylesheets ist in diesem ersten Entwurf nicht vorgesehen. Hier wäre die Etablierung eines organisatorischen Workflows denkbar, der nebenläufig zum Versand des Containers zu etablieren wäre.



**Abb. 32** Verzeichnis für Stylesheets

Wie oben beschrieben (siehe 4.1.9) kann vorgesehen werden, dass die im Container enthaltenen Metadaten auf Wohlgeformtheit und Gültigkeit geprüft werden. Für die Projektsteuerinstanz `projekt.xml` sind diese obligatorisch, weshalb `projekt.xsd` im Verzeichnis enthalten sein muss, für alle anderen Metadaten optional. Es wird davon ausgegangen, dass die Prüfung bereits durch die sendende Infrastruktur durchgeführt wurde. Wenn, was beim offenen Standard XML möglich ist, Standards-Strukturen durch private Elemente erweitert wurden, würde bei der Eingangsprüfung beim empfangenden System die Validierung misslingen. Für den letzteren Fall müssen daher die entsprechenden geänderten Document Type Definitions bzw. Schemata den XML-Parsern bekanntgemacht werden. Die betreffenden Instanzen werden im Ordner Struktur hinterlegt.

Das Verzeichnis kann auch für die Ablage von Schematron Rules verwendet werden, welche für den dritten Pass der Validierung verwendet werden können.

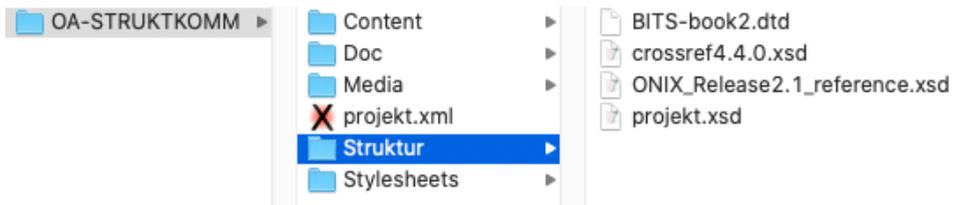


Abb. 33 Verzeichnis für DTDs und Schemata

Damit ist die Verzeichnisstruktur des Containers eingeführt. Dem Gebot der Offenheit folgend, soll es möglich sein, die Verzeichnisstruktur zu erweitern. Diese Offenheit ist in dem Konzept dahingehend angelegt, dass die Instanzen und ihre Elemente aus der Projektdatei referenziert werden. Damit ist es möglich, einerseits „private“, d.h., projektspezifische, Unterverzeichnisse anzulegen und andererseits diese gegebenenfalls zu lokalisieren. Möglich ist dies durch die konzeptionelle Anlage der Projektsteuerinstanz. Deren Verwendung soll nachfolgend vorgestellt werden.

Das Schema für die XML-Instanz der Projektsteuerdatei befindet sich vollständig im Anhang E. Dessen Anwendung soll nachfolgend am oben eingeführten Beispiel dokumentiert werden.

Wurzelelement ist das Element `<projekt>`.

```
<projekt xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:noNamespaceSchemaLocation="Struktur/projekt.xsd">
  ...
</projekt>
```

Im ersten Kindelement-Knoten können Informationen zur Erstellungshistorie der Projektinstanz hinterlegt werden:

```
<projekt-metadaten>
  <historie>
    <event type="Erstellung" datum="2021-07-29"/>
  </historie>
</projekt-metadaten>
```

Im `<content>`-Knoten werden alle strukturierten Instanzen referenziert. Für jede wird ein Kindknoten `<instanz>` mit dem Inhaltsmodell `EMPTY` angelegt. Die notwendigen Metadaten werden in drei Attributen hinterlegt. Für jeden `<instanz>`-Knoten wird eine für die Instanz unikat Idenfikationsnummer generiert (`id="1"`), die physische Instanz (die sich in der Regel innerhalb des Containers befindet, es können aber auch externe Ressourcen referenziert werden), wird mit dem Attribut `href` verbunden (`href="Content/BITS_Musterbuch_Monografie.xml"`). Für die Qualitätssicherung der Metadaten kann im Container das Schema hinterlegt werden, gegen welches die referenzierte Instanz geprüft werden kann. Es sind drei Fälle zu unterscheiden:

1. die Qualitätssicherung obliegt der schreibenden Infrastruktur, dann ist das Attribut nicht anzugeben;
2. die Instanz ist gültig gegen ein Standard-Schema, dann kann
  - a) ein URI angegeben werden, der auf die Ressource des Standard-Schemas zeigt oder
  - b) die Schema-Instanz kann physisch im Verzeichnis Struktur abgelegt werden oder
3. eine Standard-Instanz wurde durch private Elemente erweitert. Dann ist diese projektspezifische Instanz im genannten Verzeichnis zu hinterlegen (`schema="Struktur/BITS-book2.dtd"`).

```
<content>
  <instanz id="1"
    href="Content/BITS_Musterbuch_Monografie.xml"
    schema="Struktur/BITS-book2.dtd" />
  ...
</content>
```

Der nachfolgende Geschwisterknoten `<metadaten>` enthält die Referenzen auf die einzelnen Metadaten, die in den eingebundenen Metadaten-Instanzen enthalten sind. Der oben aufgerissenen Idee folgend, wird beispielhaft eine endliche Liste verwendet. Diese Liste wurde aus den weiter oben diskutierten angepassten Metadaten des Workflowmodells aus dem Projekt OA-HVerlag generiert. Diese kann nur den Anspruch erheben, innerhalb dieses Workflowmodells vollständig zu sein. Die Untersuchungen der Schnittmengen zu weiteren Standards und Quasi-Standards<sup>99</sup> haben eine ausreichende Anzahl ergeben, so dass die (offene) Liste für die Demonstration der Idee als geeignet angesehen werden kann. Für eine künftige Superstruktur muss durch eine Community eine vollständige Liste bzw. ein Thesaurus oder eine Ontologie nach dem Beispiel von JDF entwickelt werden. Die hier verwendete Liste befindet sich im Anhang

---

<sup>99</sup> vgl. Reiche et al. 2023

C. Aus dieser Liste werden die Attributwerte der Attribute `var` und `name` der Kindknoten `<meta>` des Knotens `<metadaten>` entnommen. Das Attribut `var` entspricht dabei der Nummer innerhalb der Liste und `name` enthält die semantische Bezeichnung des zugehörigen Metadatums. Das weitere Attribut `id-ref` referenziert auf die Instanz im Ordner `Content`, in welcher ein semantisch übereinstimmendes Element vorhanden und in dessen Knoten der Metadaten-Wert gespeichert ist. Hier ist zu beachten, dass es für den Fall, dass im `Content`-Verzeichnis zwei (oder mehr) strukturierte Instanzen vorhanden sind, ein semantisches Metadatum in beiden (oder mehr) Standards vorkommen kann. In diesem Fall wären dann auch die oben diskutierten Heterogenitäten zu beachten. Es werden zwei Vorgehensweisen für diesen Fall vorgeschlagen:

1. die sendende Infrastruktur entscheidet, auf welche strukturierte Instanz verwiesen wird und trägt nur eine `id-ref` in die Liste ein. Eine Referenz auf die andere(n) Instanz(en), welche die selbe Metadateninformation enthält, erfolgt nicht. Die Entscheidung, auf welche Instanz referenziert wird, liegt bei der sendenden Infrastruktur.
2. Die sendende Infrastruktur trägt beide (oder mehrere) Referenzen auf Metdatennoten in den verschiedenen Instanzen ein. Dann entstehen mehrere `<meta>`-Knoten mit gleichen Attributwerten für `var` und `name` und unterschiedlichen für `ref_id`.

```
<metadaten>
  <meta var="101" name="Werk-Titel-Hauptsprache"
  id_ref="1">book/book-meta/book-title-group/book-title</meta>
  <meta var="102" name="Werk-Untertitel-Hauptsprache"
  id_ref="1">book/book-meta/book-title-group/subtitle</meta>
  ...
  <meta var="105" name="Abstrakt-Hauptsprache"
  id_ref="2">ONIXMessage/Product/OtherText[@language='ger']
  </meta>
  ...
  <meta var="110" name="Abbildung1-DOI"
  id_ref="3">doi_batch/body/content_item[@component_type=
  'chapter']/component_list/component[@parent_relation=
  'isPartOf']/doi_data/doi</meta>
  <meta var="999" name="Rechnungsnummer">0815/23</meta>
</metadaten>
```

Auf Anregung des Stakeholder-Meetings wurde, neben der Empfehlung, die Idee eines Containerformats zu entwickeln, die Möglichkeit diskutiert, neben den strukturierten Meta- und Produktionsdaten auch Dokumente, die in einem wissenschaftlichen Publikationsworkflow anfallen, wie Mess- und Peer-Review-Daten, transportieren zu können. Der akademische Vorschlag folgt dieser Idee, indem in der Projektsteuer-Ins-

tanz ein Knoten <dokumente> vorgesehen wird, der als Sammlung der Referenzen aller nicht strukturierten Dokumente angesehen werden soll. Die Referenzen werden in Kindknoten mit dem Elementnamen <dokument> hinterlegt. Diese folgen dem Inhaltsmodell EMPTY, enthalten also nur Attribute. Das Attribut id enthält als Attributswert eine projekt-unikate Identifikationsnummer, die automatisch generiert werden kann. Das Attribut href enthält eine Referenz auf ein Dokument im Doc-Pfad (siehe Abb. 29).

```
<dokumente>
  <dokument id="3" href="doc/abstract.doc"/>
  <dokument id="4" href="doc/gutachten.doc"/>
  <dokument id="5" href="doc/messreihe1.xls"/>
</dokumente>
```

Die Dokumente im Doc-Verzeichnis können auch durchaus strukturierte Daten enthalten (zum Beispiel komma-separierte Messreihen etc.). Nichtstrukturierte Daten ist daher in dem Sinne zu verstehen, dass diese Daten nicht qualitätsgesichert (z.B. geparsed) werden und auf deren Strukturen keine Referenzen ausgebildet werden, sondern dass diese lediglich im Container transportiert werden.

Es sollen auch Referenzen auf externe Dokumente zugelassen werden, so dass die betreffenden Dokumente nicht physisch im Container transportiert werden, aber über den Mechanismus zum Projekt zugehörig gekennzeichnet werden. Die Auswertung (siehe Kap. 4.1.5) obliegt der lesenden Infrastruktur und soll nicht geregelt werden.

Wie oben beschrieben, können Mediendaten als Produktionsdaten zu einem Publikationsprojekt separat im Verzeichnis Media für den Transfer abgelegt werden. In der Projektsteuerinstanz können diese Mediendaten im Knoten <media> referenziert werden. Für jede Medieninstanz wird ein Kindknoten <medium> angelegt. Dieser ist vom Inhaltsmodell EMPTY. Im Attribut id wird eine projekt-unikate Referenznummer hinterlegt, die einer ID in der Struktur einer der Content-Instanzen enthalten ist.

```
<media>
  <medium id="cover" type="img" href="Media/cover.jpg"/>
  <medium id="logo" type="img" href="Media/HTWK_Logo.jpg"/>
  <medium id="qr" type="img" href="Media/qr-doi.jpg"/>
  <medium id="cc" type="img" href="Media/CC_Hinweis.jpg"/>
  <medium id="b_7890123456789_part_002_chapter_003_fig_001"
    type="img" href="Media/b_7890123456789_fig_001.jpg"/>
  <medium id="b_7890123456789_part_002_chapter_003_fig_002"
    type="img" href="Media/b_7890123456789_fig_002.jpg"/>
  <medium id="b_9783966270021_eq_001"
    type="img" href="Media/b_9783966270021_eq_001.png"/>
</media>
```

Dieser Mechanismus ist nur optional anzuwenden, denn er generiert Aufwand beim Anlegen des Containers. Die Separation der Mediendaten kann aber dann von Vorteil sein, wenn die empfangende Infrastruktur die Mediendaten in eine Mediendatenbank oder ein MAMS<sup>100</sup> importiert. Sie können dann direkt mit dem Verweis in die Projektdaten-Instanz importiert werden. Im Regelfall wird aber, wie oben beschrieben, ein Medienverzeichnis parallel zum oder unterhalb der Instanz im Content-Verzeichnis existieren. Für diesen Fall muss der <media>-Knoten nicht angelegt werden.

Im Lastenheft ist vermerkt, dass im Publikationsprozess von (wissenschaftlichen) OA-Veröffentlichungen Informationen anfallen, die nicht oder nur unvollständig in den strukturierten Daten abgebildet werden. Sie werden in der Regel in einem nebenläufigen (meist manuellen) Workflow behandelt. So werden Messdaten in einem Repository oft nicht direkt mit der Publikation, aber für die Verifikation der Ergebnisse und für die weiterführende Forschung verfügbar gemacht. Mit der Kommunikationsstruktur sollen einerseits die Daten transportiert werden können, andererseits sollen die Verarbeitungswflows der empfangenden Infrastruktur in die Lage versetzt werden, direkt auf die zu verarbeitenden Daten im Container zugreifen zu können.

```
<peer-review>
  <gutachten id_ref="4"/>
</peer-review>
```

```
<forschungsdaten>
  <messung id_ref="5"/>
</forschungsdaten>
```

Die beiden Knoten stehen exemplarisch für weitere Dokumentenarten, welche mit dem Container transportiert werden können. Hier muss einerseits diskutiert werden, welche weiteren Dokumentenarten dies betrifft und wie deren Beschreibung über die Metadaten ggf. erweitert werden kann, d.h., welche Attribute noch verwendet werden können. Denkbar wäre zum Beispiel ein Attribut, hinter dem die Version eines Dokumentes hinterlegt werden kann, da die o.a. Dokumente typischerweise Veränderungen über die Zeit unterliegen.

Damit wurden die Anwendungsfälle 4.3.3 und 4.3.5 des Dokumentenlebenszyklus<sup>1</sup> hinreichend beschrieben. Implizit trifft dies auch auf den Anwendungsfall 4.3.4 zu, bei dem in einem bestehenden Container einzelne Metadaten manipuliert, d.h., hinzugefügt, geändert oder eliminiert werden. Für alle drei Anwendungsfälle besteht die Notwendigkeit, die Datenkonsistenz des Containers, für 4.3.3 und 4.3.4 nach Abschluss der Transaktion, für 4.3.5 ggf. vor dem Beginn des Imports zu validieren.

---

100 Media Asset Management System

#### 4.4.2 Validieren der Struktur(en) innerhalb eines Containers

Die Projektsteuerinstanz `projekt.xml` ist durch die sendende Infrastruktur wohlgeformt und gültig in den Container zu integrieren. Dies kann durch einen einfachen XML-Parser, zum Beispiel den Open-Source-Parser Xerces aus dem Apache-Projekt<sup>101</sup> erfolgen, der die Projektinstanz gegen `Struktur/Projekt.xsd` parsed.

Wie im Kapitel 3.2.3 bereits diskutiert, obliegt die Qualitätskontrolle der Daten im Container der schreibenden/sendenden Infrastruktur. Es wird also vorausgesetzt, dass die Daten im Container wohlgeformt und gültig sind. Ebenso kann die empfangende Infrastruktur die Daten im Container vor der Verarbeitung prüfen und im Fehlerfall den Container ggf. ablehnen.

Es wurde aber auch der Fall besprochen, dass eine der zu sendenden strukturierten Instanzen einer DTD oder einem Schema entspricht, das durch private Elemente erweitert wurde. In diesem Fall muss für die Qualitätssicherung bei der empfangenden Infrastruktur die geänderte Regelinstanz mit dem Container transportiert werden.

Die beschriebenen Prüfungen sind regelhafte Prüfungen, die in einem jeden XML-basierten Workflow durchgeführt werden können bzw. müssen. Zusätzlich gehört zu den Aufgaben der sendenden Infrastruktur die Sicherung der Integrität der Daten im Container. Dabei sind folgende Fälle zu unterscheiden:

1. Es ist zu prüfen, ob alle für das Publikationsprojekt notwendigen Metadaten mit einem Eintrag oder einer Referenz im Knoten `<metadata>` als `<meta>`-Knoten in der Projektsteuerinstanz enthalten sind.
2. Es ist zu prüfen, ob alle aus der Projektsteuerinstanz verlinkten Ressourcen auch im Container (oder einer externen Speicherstelle) vorhanden sind, d.h., dass keiner der Links invalid ist.
3. Es ist zu prüfen, dass sich im Container keine nicht in der Projektsteuerinstanz verlinkten Ressourcen befinden.
4. Es kann geprüft werden, ob alle in den enthaltenen Produktions- bzw. Metadateninstanzen Verweise auf Ressourcen gültig sind, d.h., dass alle Ressourcen zuverlässig mit dem Container transportiert werden.

---

101 vgl. Xerces apache 2022

Die beschriebenen Prüfungen können durch drei Technologien umgesetzt werden:

1. Durch einen erweiterten Parser, welcher die Prüfungen in einem dritten Pass durchführt. Dafür müsste ein bestehender XML-Parser erweitert werden. Dies generiert einerseits erheblichen Aufwand und andererseits das Problem, dass bei künftigen Erweiterungen jeweils eine neue Version des Parsers zur Verfügung gestellt werden muss. Dass dieser Ansatz umsetzbar ist, beweisen der JDF-Parser und diverse Validierungswerkzeuge für EPUB-Instanzen.
2. Die Regeln für die Validierungen können in XSLT-Stylesheets hinterlegt werden.
3. Für die Validierung können Schematron-Business-Rules verwendet werden.

Die Technologien können auch in Kombination eingesetzt werden.

#### 4.4.3 Anwendungsfall: Herausziehen eines Metadatensatzes aus dem Container

Im Kapitel 4.3.2 wurde der Anwendungsfall Herausziehen von Metadatensätzen aus dem Container vorgestellt. Nachfolgend sollen zwei Varianten des Anwendungsfalls dargestellt werden. Sie sollen beispielhaft die Vorgehensweise bei der fallspezifischen Extraktion von Datensätzen dokumentieren.

Die vorgestellten Formatierung und Transformation werden in der Entwicklungsumgebung ausgeführt, welche bereits für das im Projekt OA-HVerlag entwickelten und im Projekt OA-STRUKTKOMM validierte Pop-Up-Publikationssystem<sup>102</sup> verwendet wurde.

##### 4.4.3.1 Generierung eines Formulars zur Unterstützung der DOI-Aktivierung

Folgendes Szenario soll beschrieben werden. Das Publikationsprojekt ist im Projektcontainer zusammengestellt, die Projektsteuerdatei `projekt.xml` ist wohlgeformt und gültig. Für das Publikationsprojekt soll eine DOI via `Web Deposit Form`<sup>103</sup> angemeldet werden. Dies erfolgt derzeit über ein Formular. Mittels eines Stylesheets (`01_DOI_translate2CrossrefWb.xsl`) können aus den Publikationsinstanzen die jeweiligen Informationen herausgezogen und in einem PDF-Formular für die Eingabe zusammengestellt werden. Beispielhaft wird in Abb. 34 der Aufruf des Stylesheets auf der Kommandozeile gezeigt.

---

<sup>102</sup> Böhm et al. 2023

<sup>103</sup> Web-Formular zur manuellen Eingabe von Metadaten

```

MA_Sannemann_Entwicklungsumgebung — zsh — 89x5
michaelreiche@VdM-GU002-MacBook-Pro MA_Sannemann_Entwicklungsumgebung % ./formatiere.sh K
ommunikationsstruktur_Musterbuch/projekt.xml Stylesheets/01_DOI_translate2CrossrefWeb.xsl
Outputs/DOI-Formular.pdf

```

**Abb. 34** Aufruf des Formatierungsstylesheets zur Generierung des DOI-Formulars

Nachfolgend soll exemplarisch eine Ausschnitt aus dem Stylesheet gezeigt werden, welches dokumentiert, wie der Eintrag des Werkstitel auf dem Formular erfolgt:

```

<fo:block font-size="12pt" space-after="9mm" font-weight="normal">
  <xsl:text>Werktitel: </xsl:text>
  <xsl:variable name="Translate-Meta" select="projekt/metadaten/meta
    [@var='101']" />
  <xsl:for-each select="projekt/content/instanz">
    <xsl:if test="@id=$Translate-Meta/@id_ref">
      <xsl:variable name="Metadatenquelle_A">
        <xsl:value-of select="$Pfad" />
        <xsl:value-of select="@href" />
      </xsl:variable>
      <xsl:choose>
        <xsl:when
test="document($Metadatenquelle_A)/book/book-meta/book-title-group/
book-title[normalize-space(.)]">
          <xsl:apply-templates
select="document($Metadatenquelle_A)/book/book-meta/book-title-group/
book-title"/>
        </xsl:when>
        <xsl:otherwise>
          <xsl:value-of select="$Warning"/>
          <xsl:message>
            <xsl:text>Titel: keine Informationen
              vorhanden</xsl:text>
          </xsl:message>
        </xsl:otherwise>
      </xsl:choose>
    </xsl:if>
  </xsl:for-each>
</fo:block>

```

Das vollständige Stylesheet ist im Anhang F dokumentiert. Abbildung 35 zeigt einen Ausschnitt aus dem automatisch generierten Formular zur Unterstützung der Eintragung der DOI für das Publikationsprojekt.

**Vorlage: Webformular zu DOI Aktivierung auf Crossref**

Werktitel: Titel der Open-Access-Publikation

---

Data Type Selection

---

**Select Data Type:** Book

---

Book-Information

---

**Select Book Deposit Type:** Monograph

**Book Type:** Monograph (Warning: Information could not be found)

Abb. 35 Ausschnitt aus dem DOI-Formular

Die Ausgabe des Werktitel oben ist das Ergebnis des oben vorgestellten Ausschnitts aus dem Formatierungsstylesheet. Das vollständige Formular ist im Anhang F.2 hinterlegt.

#### 4.4.3.2 Erzeugung einer Upload-Instanz zur Registrierung von Peer-Review-Daten

Sollen für eine Publikation, die bereits eine DOI besitzt, Peer-Review-Daten übertragen werden, kann dies durch einen Upload (Direct Deposit of XML<sup>104</sup>) einer XML-Instanz vorgenommen werden. Es soll als zweiter Anwendungsfall gezeigt werden, wie dieser Batch als Anfrage an den Container generiert werden kann. Dazu wird im vorgestellten Beispiel ein Transformationsstylesheet (`@2_DOI_translate2CrossRefWeb.xsl`) auf die Projektsteuerinstanz `projekt.xml` angewendet.

---

<sup>104</sup> Metadatenupload mittels einer XML-Instanz



```

michaelreiche@VdM-GU002-MacBook-Pro MA_Sannemann_Entwicklungs Umgebung % ./transformiere.s
h Kommunikationsstruktur_Musterbuch/projekt.xml Stylesheets/02_PeerReview_translate2Cross
refXML.xsl Outputs/Peer-Review.xml

Transformiere XML: Kommunikationsstruktur_Musterbuch/projekt.xml unter Benutzung des Sty
lesheets Stylesheets/02_PeerReview_translate2CrossrefXML.xsl nach Outputs/Peer-Review.xml

Metadatenquellen included:
../Kommunikationsstruktur_Musterbuch//Content/BITS_Musterbuch.xml
../Kommunikationsstruktur_Musterbuch//Content/Onix2.1_Musterbuch_Monografie.xml
../Kommunikationsstruktur_Musterbuch//Content/Crossref_Musterbuch_Monografie.xml
../Kommunikationsstruktur_Musterbuch//Content/Crossref_Peer-Review_Musterbuch_Monografie.
xml

Batch-ID 30082023-01
Zeitstempel 30.08.2023, 10:18
Hinweis: Als depositor wird der Verlag eingesetzt
Fehlerhinweis zu depositor_eMail: Keine Informationen vorhanden
Hinweis: Als registrant wird der Verlag eingesetzt

```

**Abb. 36** Transformation zur Hinterlegung von Peer-Review-Daten

Das Stylesheet und die Ergebnisinstanz werden im Anhang G dokumentiert.

Die beiden vorgestellten Anwendungen sollen beispielhaft für die vielen denkbaren Anfragen an die Metadaten im Container stehen. Das Grundprinzip des Zugriffs über die virtuelle Superstruktur, auf welche über die Linkliste der semantischen Metadaten in der Projektsteuerinstanz zugegriffen werden kann, macht die Anfragen sehr flexibel. Es können, wie vorgestellt, standardisierte Zugriffsmöglichkeiten der XML-Familie, wie Transformation und Formatierung, verwendet werden, genauso können aber auch Zugriffe über Programmierschnittstellen implementiert werden.

Damit ist der akademische Ansatz des Austauschs strukturierter und nicht strukturierter Meta- und Produktionsdaten mittels eines Containers und integriert über eine virtuelle Metadaten-Superstruktur beschrieben. Die Limitierungen und Weiterentwicklungsmöglichkeiten werden im nachfolgenden Kapitel diskutiert.



## 5 Schlussbetrachtung

---

Der vorgestellte akademische Ansatz ist explizit als Denkmodell zu verstehen. Es soll und kann als initiales Grundmodell dienen, welches kooperativ weiterentwickelt und implementiert werden muss. Eine vollständige Umsetzung ist im Rahmen eines Forschungsprojektes nicht möglich und war auch nicht intendiert. Vielmehr war es möglich, gemeinsam mit den Stakeholdern, welche die wichtigen Player im Entwicklungsprozess einer solchen Kommunikationsstruktur darstellen, die Rahmenbedingungen und die Entwicklungsansätze zu diskutieren und festzulegen. Nun liegt ein Konzept vor, welches die Möglichkeit bietet, heraus aus dem akademischen Elfenbeinturm in einer gut aufgestellten Community eine praxisverwendbare Kommunikationsstruktur zu entwickeln. Der vorgestellte Ansatz unterliegt natürlicherweise einigen Limitierungen, welche hier diskutiert werden sollen.

Die Diskussion zum Überblick über offene Standards im wissenschaftlichen Publizieren hat gezeigt, dass diese eine hohe Heterogenität aufweisen. Dies war der Ausgangspunkt für die Überlegungen zum vorgelegten akademischen Ansatz. Der Schwierigkeit, welche diese Situation mit sich bringt, nämlich dass eine Vollständigkeit der virtuellen Superstruktur in einem ersten Ansatz nicht zu erreichen ist, wurde dahingehend abgeholfen, dass der Metadatenansatz, welcher aus dem im Projekt OA-HVerlag entwickelten allgemeingültige Workflowmodell abgeleitet wurde, als Leit-Metadatenansatz verwendet wird. Im Artikel „Zur Situation des strukturierten Austauschs von Metainformationen über die Herstellungs- und Distributionsprozesse von Open-Access-Publikationen“<sup>105</sup> wird gezeigt, dass diese Vorgehensweise sinnvoll ist. Es ist aber die nicht zu vermeidende Limitierung zu konstatieren, dass die Liste der Metadaten, die dem akademischen Ansatz zugrunde liegt, nicht vollständig ist. Sie wird aber als ausreichend für die Beschreibung des Konzeptes einer virtuellen Superstruktur angesehen. Dieser Metadatenansatz kann im Rahmen weiterer Forschungen problemlos erweitert werden.

Eine zweite, damit zusammenhängende, Limitierung besteht darin, dass der grundlegende Metadatenansatz in erster Näherung als Liste zu verstehen ist. Strukturierte Daten, das impliziert schon die Bezeichnung, unterliegen einer Ordnung und sind damit höherwertiger als Thesauri abzubilden. Einen Ansatz hin zu dieser Art der semantischen Strukturierung wird in diesem Projekt bereits vorgelegt und im eben angeführten Artikel dargelegt. Die Liste der Metadatenelemente wird dabei in sechs Gruppen (Daten zu Werk und Lizenzen, zu Mitwirkenden, zur Produktion, zur Klassifikation, zu Peer-Reviews und zur Barrierefreiheit) eingeteilt. Hier können weitere

---

105 vgl. Reiche et al. 2023

Gruppen, wie zum Beispiel zu Metainformationen zur Förderung der Publikation oder zu betriebswirtschaftlichen bzw. Auftragsdaten zugefügt werden. Denkbar wäre auch, weitergehende Verhältnisse zwischen den Metadaten und deren Kategorien abzubilden, was die Struktur in Richtung einer Ontologie entwickeln würde.

Drittens konnten wegen der limitierten Ressourcen in einem Forschungsprojekt die Anwendungsfälle nur beispielhaft und in einem eingeschränkten Umfang implementiert werden. Vorteilhaft dabei war, dass das für das im Projekt OA-HVerlag verwendete Entwicklungssystem, das auch dem Publikationssystem zugrunde liegt und vollständig aus Open-Source-Softwarewerkzeugen besteht, dafür weiterverwendet werden konnte. Damit konnten die Funktionalitäten exemplarisch implementiert werden. Diese Limitierung ist aber insofern akzeptabel, weil der Ansatz der Kommunikationsstruktur auf den reinen Transport der Daten zwischen den Infrastrukturen abstellt und die Implementierung der beschriebenen Anwendungsfälle (und auch weiterer) in der Verantwortung der schreibenden bzw. lesenden Systeme liegt und mit deren Technologien durchgeführt wird. Daher muss die Entwicklungsumgebung jene lediglich simulieren.

Das Rückgrad der Kommunikation von Daten und Metadaten in der Medienindustrie wird durch XML-Technologien gebildet. Nichtsdestoweniger werden alternative Ansätze, wie zum Beispiel die Verwendung von JSON, diskutiert. Obwohl der akademische Ansatz im XML-Denkschema vorgestellt wird, wäre eine Übertragung der Idee auf eine andere technologische Basis problemlos möglich.

Trotz der angesprochenen Limitierungen wird der akademische Ansatz sowohl vom Forschungsteam wie auch von den beteiligten Stakeholdern, denen er auf der Transfertagung „Road to Open Meta“<sup>106</sup> am 12. September 2022 vorgestellt wurde. Die Ergebnisse der Forschungsprojekte OA-HVerlag und OA-STRUKTKOMM sollen in Teilen im anlaufenden Forschungsprojekt OA-WFMS<sup>107</sup> praktisch umgesetzt werden.

Das Beispiel der Entwicklung von JDF hat gezeigt, dass eine leistungsfähige Kommunikationsstruktur nicht im akademischen Elfenbeinturm entstehen kann. Auch wenn die zahlreichen Stakeholder des Projektes im Workshop und in unzähligen bilateralen Gesprächen in hohem Maße zur Ideenentwicklung beigetragen haben, obliegt es doch diesen, die nicht zu bestreitende Notwendigkeit einer allgemeinen Kommunikationsstruktur, ob nun eingeschränkt auf Open-Access-Publikationen oder den gesamten Verlagsbereich, praktisch umzusetzen. Der Vorteil des akademischen Ansatzes besteht in seinem Abstraktionsgrad, der es ermöglicht, Denkansätze zu generieren, die sicherstellen, die bereits von den Unternehmen geleistete Arbeit, die zu den bestehenden

---

106 OA-STRUKTKOMM 2022

107 Das Projekt Workflow-Management-Systeme für Open-Access-Hochschulverlage (OA-WFMS) wird von September 2023 bis August 2025 vom Bundesministerium für Bildung und Forschung (BMBF) gefördert (vgl. Universität Potsdam 2023).

standardisierten und quasi-standardisierten Strukturen geführt hat, zu respektieren und diese projektbezogen in den jeweils sinnvollen Kombinationen einzusetzen. Damit sollen zumindest Hemmschwellen abgebaut werden, die eine ergebnisoffene Diskussion zu einem offenen Kommunikationsstandard verhindern könnten. Dies soll als wichtigstes Outcome des Projektes OA-STRUKTKOMM verstanden werden.

Diese Community muss sich finden. Weiter oben wurden bereits Beispiele für derartige Ansätze diskutiert (Kap. 3.2). Die Präsentationen der Ergebnisse der TOA-Tagung am 12. Juli 2023 hat weiterhin bezüglich der Interessensbereiche der teilnehmenden Forschergruppen eine große Schnittmenge in den Themen aufgezeigt. Hier muss sich nun eine Interessengemeinschaft, wie CIP4, finden und die gemeinsamen Ziele definieren. Dass dies möglich ist, zeigt die Entwicklung von JDF.

Die erste naheliegende praktische Aufgabe dieser Interessengemeinschaft sollte dann die Entscheidung über die Art der Leitstruktur sein. Der vorliegende akademische Ansatz stellt auf eine offene Liste ab, die aus den Metadaten des Projektes OA-HVerlag abgeleitet ist. Damit kann diese Liste, wie bereits diskutiert, nur vollständig gegenüber der Kommunikation von Prozessen, die innerhalb dieses Modells definiert wurden, sein. Die Untersuchung der Standards und Quasi-Standards in der Domain der Workflows von OA-Publikationen (und darüber hinaus) zeigt aber eine weitaus größere Vielfalt an Metadaten. Zudem ist die Entwicklung nicht abgeschlossen, neue, bisher noch nicht bearbeitete Themenfelder werden bearbeitet. So erschienen während der Projektlaufzeit mindestens zwei neue Standards, die in den akademischen Ansatz zu integrieren wären. So wurden im Jahr 2022 mit CRediT durch die NISO vierzehn neue Contributor Roles<sup>108</sup> angeboten. Die Verwendbarkeit dieser Anwenderrollen für das Forschungsthema wurde in einer Bachelorarbeit mit dem Titel „Einordnung der NISO Contributor Roles in die aktuellen Herstellerworkflows“ untersucht<sup>109</sup>. Weiterhin wurde 2023 mit der neuen Version von FundsXML ein Vorschlag zur strukturierten Erfassung von Fördermittelinformationen<sup>110</sup> veröffentlicht, der zu untersuchen wäre, inwieweit er in das Workflowmodell und nachfolgend auch in den akademischen Ansatz zu integrieren wäre. Es ist anzunehmen, dass im Forschungszeitraum weitere Standardisierungsprojekte, vor allem im Bibliotheksbereich, erfolgreich abgeschlossen wurden. Insofern wäre es eine weitere Aufgabe der Community, eine Meta-Studie durchzuführen, welche den im Rahmen des Projektes erforschten und veröffentlichten *Überblick über offene Standards im wissenschaftlichen Publizieren* (siehe Kapitel 2.3) aktualisiert.

Die bereits abgeschlossenen, vom BMBF initiierten Projekte zur Verbesserung der Open-Access-Infrastruktur in Deutschland haben einen wertvollen Beitrag zur Etablie-

---

108 CRT Contributor Roles Taxonomy (2023)

109 Die Arbeit wurde im Juni 2023 verteidigt war bis zum 28. August noch nicht im Qucosa-Publikationsserver der HTWK erschienen.

110 FundsXML (2017)

rung einer OA-Community geleistet. Diesen Weg gilt es mit der gleichen Konsequenz und mit demselben Enthusiasmus weiterzugehen. Die Ergebnisse dieses wie auch der anderen Projekte der abgeschlossenen Aufrufe sind Puzzlesteine für eine leistungsfähige OA-Kultur in den deutschsprachigen Ländern und darüber hinaus. Nur so kann der aus der Wissenschaft heraus vorangetriebene, tiefgreifende Kulturwandel, der notwendig ist, um die Herausforderungen des 21. Jahrhunderts zu meistern, gelingen.





# Literaturverzeichnis

---

- Abts, Dietmar; Müller, Wilhelm (Hrsg.) (2009): Masterkurs Wirtschaftsinformatik: Kompakt, praxisnah, verständlich – 12 Lern- und Arbeitsmodule. Vieweg+Teubner Verlag, Wiesbaden. ISBN 9-7838-3480-0022
- Academic dictionaries and encyclopedias (n. d.): De-facto-Standard. URL: <https://de-academic.com/dic.nsf/dewiki/308799>, Letzter Zugriff: 31.05.2021
- ag universitätsverlage (2023): Institutionelle Publikationsdienste für eine offene Wissenschaft. <https://ag-univerlage.de/>, Letzter Zugriff: 09.07.2023
- Bau, Nico; Gerz, Michael; Schmitz, Hans-Christian; Schwarze, Arne (n. d.): Interoperabilität in sicherheitskritischen Kreisen. Fraunhofer-Institut für Kommunikation, Informationsverarbeitung und Ergonomie FKIE. URL: [https://www.sifo.de/sifo/shareddocs/Downloads/files/poster-10\\_interoperabilitaet\\_schmitz\\_if18.pdf?\\_\\_blob=publicationFile&v=1](https://www.sifo.de/sifo/shareddocs/Downloads/files/poster-10_interoperabilitaet_schmitz_if18.pdf?__blob=publicationFile&v=1)
- Böhm, David; Grossmann, Alexander; Reiche, Michael; Schrader, Antonia (2020): Open-Access-Publikationsworkflow für akademische Bücher. Ein Handbuch für Hochschulen und Universitäten. Open-Access-Hochschulverlag, HTWK Leipzig. DOI: 10.33968/9783966270175-00
- Böhm, David; Grossmann, Alexander; Reiche, Michael; Tillmann, Diana (2021): Überblick über offene Standards im wissenschaftlichen Publizieren. Poster. HTWK-Leipzig. DOI: 10.14293/S2199-1006.1.SOR-PPNKUIH.v1
- Böhm, David; Grossmann, Alexander; Reiche, Michael; Tillmann, Diana (2023): OA-Produktionssystem. Open Source Software. GitHub. GNU 3.0 Lizenz. <https://github.com/OA-STRUKTKOMM/OA-Produktionssystem/tree/main>
- Börsenverein (n. d.): IG Produktmetadaten. <https://www.boersenverein.de/interessengruppen/ig-produktmetadaten/>, Letzter Zugriff: 09.08.2023
- Busse, Susanne; Kutsche, Ralf-Detlef; Leser, Ulf; Weber, Herbert (1999): Federated Information Systems: Concepts, Terminology and Architectures. In: Forschungsberichte des Fachbereichs Informatik. Bericht Nr. 99-9. Technische Universität Berlin. [https://www.researchgate.net/publication/2277599\\_Federated\\_Information\\_Systems\\_Concepts\\_Terminology\\_and\\_Architectures](https://www.researchgate.net/publication/2277599_Federated_Information_Systems_Concepts_Terminology_and_Architectures)

Bühler, Peter; Schlaich, Patrick; Sinner, Dominik (2017): Visuelle Kommunikation. Bibliothek der Mediengestaltung. Springer-Verlag GmbH Deutschland. DOI 10.1007/978-3-662-53770-1\_1

Bünting, Karl-Dieter (1990): Modell einer wechselseitigen Kommunikation. In: Einführung in die Linguistik. Frankfurt/ am Main: Hain (Athenäum Taschenbücher; 2011; Sprachwissenschaft)

CRT Contributor Roles Taxonomy (2023): CRediT. 14 Contributor Roles. <https://credit.niso.org/>, Letzter Zugriff: 27.08.2023

DAISY Consortium (2023): DAISY Format. In: The DAISY Consortium. <https://daisy.org/activities/standards/daisy/>, Letzter Zugriff: 07.08.2023

DNB (23.05.2018): Lieferung von Metadaten für monografische Netzpublikationen an die Deutsche Nationalbibliothek. Metadaten-Kernset Monografien, Hörbücher und Hochschulprüfungsarbeiten. Version 2.1 URL: <https://d-nb.info/1160064784/34> Letzter Zugriff: 19.09.2023

DNB (2022): Arbeitsgemeinschaft der Verbundsysteme (AGV). In: Deutsche Nationalbibliothek. [https://www.dnb.de/DE/Professionell/Standardisierung/AGV/agv\\_node.html](https://www.dnb.de/DE/Professionell/Standardisierung/AGV/agv_node.html), Letzter Zugriff: 09.08.2023

Dublin Core (2023a): DCMI History. <https://www.dublincore.org/about/history/>, Letzter Zugriff: 05.06.2023

Dublin Core (2023b): Dublin Core™. <https://www.dublincore.org/specifications/dublin-core/>, Letzter Zugriff: 07.08.2023

Duden (2021): Interoperabilität. URL: <https://www.duden.de/rechtschreibung/Interoperabilitaet>, Letzter Zugriff: 14.05.2021

Duden (2023a): Granularität. URL: <https://www.duden.de/rechtschreibung/Granularitaet>, Letzter Zugriff: 20.08.2023

Duden (2023b): Syntax. URL: <https://www.duden.de/rechtschreibung/Syntax>, Letzter Zugriff: 20.08.2023

Duden (2023c): Mapping. URL: <https://www.duden.de/rechtschreibung/Mapping>, Letzter Zugriff: 30.08.2023

Ecma International (n. d.): ECMA-404. The JSON data interchange syntax. 2nd edition, December 2017. <https://www.ecma-international.org/publications-and-standards/standards/ecma-404/>, Letzter Zugriff: 09.08.2023

- Enoksson, Frederik; Plamér, Matthias; Naeve, Ambjörn (2007): An RDF modification protocol, based on the needs of editing Tools. In: Metadata and Semantics. Conference Paper. Springer, Boston, MA. S.191–199. DOI: 10.1007/978-0-387-77745-0\_18
- Fraunhofer-Institut für offene Kommunikationssysteme (n. d.): Interoperabilität geht nur gemeinsam. URL: <https://www.fokus.fraunhofer.de/de/dps/themen/interoperabilitaet>, Letzter Zugriff: 31.05.2021
- FundsXML (2017): FundsXML. The pan-European format. URL: <http://www.fundxml.org/what-is-fundxml/>, Letzter Zugriff: 28.08.2023
- Halbach, Wulf R. (1994): Interfaces: Medien- und kommunikationstheoretische Elemente einer Interface-Theorie. Wilhelm Fink Verlag, München. ISBN 3-7705-2934-0
- Hunscha, Sonja (2003): Kommunikations- und Interaktionsmodelle. eine Ausarbeitung von Sonja Hunscha zum Seminar Multimodale Mensch-Maschine-Kommunikation. bei Bernhard Jung und Alf Kranstedt Sommersemester 2003. Universität Bielefeld, Technische Fakultät.
- ISO (2021): ISO/IEC 21320-1:2015. Information technology – Document Container File – Part 1: Core. In: ISO. <https://www.iso.org/standard/60101.html>, Letzter Zugriff: 07.08.2023
- CIP4 Organization (16.08.2020): JDF Specification. Release 1.7. <https://www.cip4.org/files/cip4-2022/Documents/Specifications/JDF%20Specification%201.7.pdf> Letzter Zugriff: 12.09.2023
- JEITA (n. d.): AV&IT Technology Standardization. Digital Cameras. In: Japan Electronics and Information Industries Association. [https://www.jeita.or.jp/cgi-bin/standard\\_e/list.cgi?cateid=1&subcateid=4](https://www.jeita.or.jp/cgi-bin/standard_e/list.cgi?cateid=1&subcateid=4), Letzter Zugriff: 08.08.2023
- Jeude, Kirsten (12.04.2013): Auf Du und Du mit Dublin Core & Co. – Berufsfeld Metdaten-Management im Kontext der Bibliothek der Zukunft. Leibnitz-Informationszentrum Wirtschaft. In: slideshare. <https://de.slideshare.net/suvanni/auf-du-und-du-mit-dublin-core-co-berufsfeld-metadatenmanagement-im-kontext-der-bibliothek-der-zukunft>, Letzter Zugriff: 30.08.2023
- JSON (1999): Einführung in JSON. <https://www.json.org/json-de.html>, Letzter Zugriff: 09.08.2023
- Kálmán, Tibor; Kurzawe, Daniel; Lordick, Harald; Mache, Beata; Schmunk, Stefan; Walkowski, Niels-Oliver (2012): Metadaten Crosswalk Umsetzung exemplarischer Crosswalks in der in AP 1.2 entwickelten Metadaten-Registry (M 3.4.1). Projekt DARIAH-DE Aufbau von Forschungsinfrastrukturen für die e-Humanities.

- Kasubke, Mara (2020): Was ist eine Schnittstelle? In: 1CRM, visual4 GmbH. URL: <https://1crm-system.de/prozessoptimierung/was-ist-eine-schnittstelle/>, Letzter Zugriff: 27.05.2021
- Leser/Naumann (2007): Informationsintegration. Architekturen und Methoden zur Integration verteilter und heterogener Datenquellen, Heidelberg: dpunkt.verlag GmbH, ISBN 3-8986-4400-6
- Library of Congress (26.04.2023): Standard Generalized Markup Language (SGML). ISO 8879:1986, In: Sustainability of Digital Formats: Planning for Library of Congress Collections, <https://status.iso.org/incidents/dldzn680scx2>, Letzter Zugriff: 05.06.2023
- Müller, Uwe (2017): Standards und Best Practices im Kontext von Open Access. In: Praxishandbuch Open Access. De Gruyter Saur. URL: <https://www.degruyter.com/document/doi/10.1515/9783110494068-007/html>
- HTWK FIM (2023): OA-HVerlag. Beschreibung des Projektes. In: HTWK Leipzig. URL: <https://fim.htwk-leipzig.de/forschung/forschungsprojekte-der-fakultaet/open-access-hochschulverlag/projektbeschreibung>, Letzter Zugriff: 31.08.2023
- OA-STRUKTKOMM (2022): Transfertagung Road to Open Meta. In: OA-STRUKTKOMM. URL: <https://oa-strukt Komm.htwk-leipzig.de/transfertagung/dokumentation>, Letzter Zugriff: 30.08.2023
- OA-STRUKTKOMM (2023): OA-STRUKTKOMM. URL: Open-Access-Strukturierte-kommunikation <https://oa-strukt Komm.htwk-leipzig.de/start>, Letzter Zugriff: 30.08.2023
- Open Archives Initiative (n. d.): Open Archives Initiative. Protocol for Metadata Harvesting. Interoperability through Metadata Exchange. <https://www.openarchives.org/pmh/> Letzter Zugriff: 12.09.2023
- OpenGeoEdu (2018): Open Data. Vorlesung. Merkmale offener Daten. Interoperabilität. URL: <https://learn.opengeoedu.de/opendata/vorlesung/merkmale/interoperabilitaet>, Letzter Zugriff: 14.05.2021
- Pedersen, Susanne (2005): Interoperabilität heterogener Informationsquellen im Gesundheitswesen auf Grundlage von Standards für die medizinische Kommunikation und Dokumentation. Dissertation. Carl von Ossietzky Universität Oldenburg. <http://oops.uni-oldenburg.de/145/1/pedint05.pdf>
- Reiche, Michael; Tillmann, Diana; Grossmann, Alexander; Böhm, David (2023): Zur Situation des strukturierten Austauschs von Metainformationen über die Herstellungs- und Distributionsprozesse von Open-Access-Publikationen. Open-Access-Hochschulverlag, HTWK Leipzig. DOI: 10.33968/9783966270526-00

- Röhner, Jessica; Schütz, Astrid (2016): Psychologie der Kommunikation. Basiswissen Psychologie. Springer Fachmedien, Wiesbaden. DOI 10.1007/978-3-658-10024-7\_2
- Rühle, Stefanie (2012a): Kleines Handbuch Metadaten. Metadaten. KIM Kompetenzzentrum Interoperable Metadaten. [https://wiki.dnb.de/download/attachments/43523047/201209\\_metadaten.pdf](https://wiki.dnb.de/download/attachments/43523047/201209_metadaten.pdf)
- Rühle, Stefanie (2012b): „Kleines Handbuch Metadaten. Metadatenprofile. KIM Kompetenzzentrum Interoperable Metadaten. [https://wiki.dnb.de/download/attachments/43523047/201209\\_metadatenprofile.pdf](https://wiki.dnb.de/download/attachments/43523047/201209_metadatenprofile.pdf)
- Sannemann, Selina (2021): Analyse von Metadaten in der Produktion und Distribution von Open-Access-Publikationen. Masterarbeit. HTWK Leipzig. Zum Zeitpunkt der Veröffentlichung des Forschungsberichtes stand diese noch nicht auf den Qucosa-Dokumentenserver zur Verfügung.
- Schematron (2021): Schematron: A Language for Validating XML. In. Schematron. <https://schematron.com/>, Letzter Zugriff: 08.08.2023
- Shannon, Claude E.; Weaver, Warren (1964): The Mathematical Theory of Communication. <https://www.mathes.in.tum.de/file/10i99mf5b34so/Publications/2004/Stef04/Stef04.pdf>
- Staub, Peter (2009): Über das Potenzial und die Grenzen der semantischen Interoperabilität von Geodaten. Ein operationelles Verfahren zur Nutzung verteilter Systeme in Geodaten-Infrastrukturen. Dissertation. Eth Zürich. <https://www.research-collection.ethz.ch/handle/20.500.11850/16385>
- Steffens, Ulrike (2004): Metadaten-Management für kooperative Anwendungen. Dissertation. Technischen Universität Hamburg-Harburg.
- Universität Potsdam (2023): Workflow-Management-Systeme für Open-Access-Hochschulverlage (OA-WFMS). In: Projekte. URL: <https://www.uni-potsdam.de/de/cio/projekte>, Letzter Zugriff: 29.08.2023
- Visser, Pepijn R.S.; Jones, Dean M.; Bench-Capon, T.J.M.; Shave, M.J.R. (1997): An Analysis of Ontology Mismatches; Heterogeneity versus Interoperability. University of Liverpool. <https://aaai.org/papers/0021-an-analysis-of-ontology-mismatches-heterogeneity-versus-interoperability/>
- VLB (2023a): Auflage, Publikationsort & Erscheinungsjahr. In: VLB. Eine Marke von MVB. VLB-ONIX-Empfehlungen. URL: <https://vlb.de/hilfe/vlb-onix-empfehlungen/auflage-publikationsort-und-erscheinungsjahr> Letzter Zugriff: 19.09.2023

VLB (2023b): Lieferbarkeitsangabe & Publikationsstatus. In: VLB. Eine Marke von MVB. VLB-ONIX-Empfehlungen. URL: <https://vlb.de/hilfe/vlb-onix-empfehlungen/lieferbarkeitsangabe-publikationsstatus> Letzter Zugriff: 19.09.2023

VLB (2023c): Produktformangaben. In: VLB. Eine Marke von MVB. VLB-ONIX-Empfehlungen. URL: <https://vlb.de/hilfe/vlb-onix-empfehlungen/produktformangaben> Letzter Zugriff: 19.09.2023

Wagner, Karl Heinz (1997): Einführung in die Sprachwissenschaft. Kapitel 3: Semiotik. In: Universität Bremen. URL: <http://www.fb10.uni-bremen.de/khwagner/grundkurs1/kapitel3.aspx>, Letzter Zugriff: 14.05.2021

W3C (2008): Extensible Markup Language (XML) 1.0 (Fifth Edition). W3C Recommendation. 26 November 2008. W3C (MIT, ERCIM, Keio). <http://www.w3.org/TR/2008/REC-xml-20081126/>, Letzter Zugriff: 07.08.2023

W3C (2011): HTML5. A vocabulary and associated APIs for HTML and XHTML. W3C Working Draft 25 May 2011. W3C (MIT, ERCIM, Keio). <http://dev.w3.org/html5/spec-LC/>, Letzter Zugriff: 07.08.2023

W3C (2018): Scalable Vector Graphics (SVG) 2. W3C Candidate Recommendation 04 October 2018. W3C (MIT, ERCIM, Keio, Beihang). <https://www.w3.org/TR/2018/CR-SVG2-20181004/>, Letzter Zugriff: 07.08.2023

W3C (2023a): EPUB 3.3. W3C Recommendation 25 May 2023. International Digital Publishing Forum and World Wide Web Consortium. <https://www.w3.org/TR/2023/REC-epub-33-20230525/>, Letzter Zugriff: 07.08.2023

W3C (2023b): Cascading Style Sheets. home page. <https://www.w3.org/Style/CSS/>, Letzter Zugriff: 07.08.2023

W3C (2023c): epubcheck. In: GitHub. <https://github.com/w3c/epubcheck>, Letzter Zugriff: 08.08.2023

Weiss, Matthias (2018): Die Bedeutung der Interoperabilität für das Internet der Dinge. URL: <https://www.ias.uni-stuttgart.de/service/begriffslexikon/bedeutung-der-interoperabilitaet-fuer-das-internet-der-dinge/>, Letzter Zugriff: 31.05.2021

Xerces apache (2022): The Apache Xerces™ Project – xerces.apache.org. <https://xerces.apache.org/>, Letzter Zugriff: 08.08.2023





# Anhang

---

# Anhang A

---

Offene Standards im wissenschaftlichen Publizieren



# Anhang B

---

Externe Kommunikationsteilnehmer

**Externe Kommunikationsteilnehmer der publizierenden Institution**

|  |   |
|--|---|
| Autor  | Verfasser des wissenschaftlichen Werks und Besitzer der Urheberrechte.  |
| Gutachter  | Begutachtet das Werk nach fachlichen Gesichtspunkten (z.B. im Peer-Review).   |
| Lektorat   | Nimmt fachliche, sprachliche oder technische Korrekturen am Content vor.  |
| Grafikdienstleister  | Ist für die Bildbearbeitung und Erstellung von Grafiken sowie die Umsetzung von Cover und Umschlag verantwortlich.  |
| Druckdienstleister   | Ist für den Druck und Vertrieb der physischen Ausgaben der Publikation verantwortlich   |
| Satzdienstleister  | Ist für die medienneutrale Umsetzung des Werks verantwortlich.  |
| Marketing  | Ist zuständig für das Marketing-Konzept und dessen Umsetzung zur Bewerbung des finalen Produktes verantwortlich.  |
| Aggregator (z.B. VLB, NewBooks, Crossref)                        | Bietet Dienstleistungen für die Streuung und Distribution von Titelmetadaten an und sorgt so für eine höhere Sicht- und Auffindbarkeit der Publikation.           |
| Bibliothek (z.B. DNB, SLUB)                                      | Ist als Institution verantwortlich für die Archivierung und Zugänglichkeit von Publikationen.   |
| Buchhändler bzw. Barsortiment (z.B. KNV, Libri)                  | Sind verantwortlich für die Zugänglichkeit und den Vertrieb von physischen Werken.  |
| Datenbanken (z.B. OAPEN, DOAB)                                   | Stellen das digitale Werk Open Access zur Verfügung und erhöhen die Sicht- und Auffindbarkeit der Publikation.  |
| Repositorium (z.B. Qucosa)                                       | Dokumenten- und Publikationsserver von wissenschaftlichen Institutionen und Verlagen, sichert die Zugänglichkeit und langfristige Archivierung von Publikationen. |
| Rezipienten (oder auch Kunden, z.B. Besteller einer Publikation) | Bezeichnet im Allgemeinen den Leser der wissenschaftlichen Publikation.   |

# Anhang C

---

Metadaten im OA-HVerlag

## C.1 Metadaten einer Open-Access-Publikation im OA-HVerlag

Der vorliegende Datensatz des OA-HVerlag-Workflowmodells wird in sechs Kategorien unterteilt. Diese enthalten jeweils eine verschiedene Menge an Datenelementen und sind grundsätzlich erweiterbar.

### 1. Daten zu Werk und Lizenzen

|      |  |
|------|--|
| 1.01 | Werk-Titel (Hauptsprache)  |
| 1.02 | Werk-Untertitel (Hauptsprache)   |
| 1.03 | Werk-Titel (Zweitsprache)  |
| 1.04 | Werk-Untertitel (Zweitsprache)   |
| 1.05 | Hierarchische bibliografische Relation (Reihenzugehörigkeit)             |
| 1.06 | Abstrakt (Hauptsprache)  |
| 1.07 | Abstract (Zweitsprache)  |
| 1.08 | Werk-DOI   |
| 1.09 | Kapitel-DOI  |
| 1.10 | Abbildung-DOI  |
| 1.11 | Tabelle-DOI  |
| 1.12 | Hardcover-ISBN   |
| 1.13 | Softcover-ISBN   |
| 1.14 | EPUB/MOBI-ISBN   |
| 1.15 | ePDF-ISBN  |
| 1.16 | Produktsprache (nach ISO 639-1)  |
| 1.17 | Art der Hochschulschrift (Bachelor / Master / Diplom / Forschungsarbeit) |
| 1.18 | Ladenpreis-Hardcover   |
| 1.19 | Ladenpreis-Softcover   |
| 1.20 | Ladenpreis elektronische Formate   |
| 1.21 | Auflage  |
| 1.22 | Art der Auflage (Neuausgabe / Nachdruck)                                 |
| 1.23 | Publikationsort  |
| 1.24 | Erscheinungsland   |
| 1.25 | Publikationsstatus (Noch nicht erschienen / erschienen)                  |
| 1.26 | Förderer   |
| 1.27 | Cover-Bild   |
| 1.28 | Copyright-Information  |
| 1.29 | CC-Lizenz  |
| 1.30 | Access-Status  |
| 1.31 | Art des Urhebers   |
| 1.32 | Rechteinhaber  |
| 1.33 | Auftragsnummer   |

### 2. Daten zu Mitwirkenden

|      |                         |
|------|-------------------------|
| 2.01 | Autor-Name              |
| 2.02 | Autor-Vorname           |
| 2.03 | Autor-Titel             |
| 2.04 | Autor-Wohnadresse       |
| 2.05 | Autor-PLZ               |
| 2.06 | Autor-Wohnort           |
| 2.07 | Autor-Institution       |
| 2.08 | Autor-Telefon-Nr.       |
| 2.09 | Autor-E-Mail-Adresse    |
| 2.10 | Autor-Kurzvita          |
| 2.11 | Autor-ORCID-ID          |
| 2.12 | Autor-ISNI              |
| 2.13 | Autor-Autorenbild       |
| 2.14 | Autor-Bankverbindung    |
| 2.15 | Beteiligte-Name         |
| 2.16 | Beteiligte-Vorname      |
| 2.17 | Beteiligte-Titel        |
| 2.18 | Beteiligte-E-Mail       |
| 2.19 | Beteiligte-Institution  |
| 2.20 | Herausgeber-Name        |
| 2.21 | Herausgeber-Adresse     |
| 2.22 | Herausgeber-PLZ         |
| 2.23 | Herausgeber-Land        |
| 2.24 | Herausgeber-E-Mail      |
| 2.25 | Herausgeber-Website-URL |

### 3. Daten zur Produktion

|      |  |
|------|--|
| 3.01 | Produktformen                              |
| 3.02 | Datei der elektronischen Ressource (Abruf) |
| 3.03 | EPUB-Version                               |
| 3.04 | Umfang Print Insgesamt                     |
| 3.05 | Umfang Print Römisch                       |
| 3.06 | Umfang Print Arabisch                      |
| 3.07 | Abbildungszahl-S/W                         |
| 3.08 | Abbildungszahl-farbig                      |
| 3.09 | Print-Format-Breite                        |
| 3.10 | Print-Format-Höhe                          |
| 3.11 | Print-Farbigkeit                           |
| 3.12 | Papierqualität Innenteil                   |
| 3.13 | Papierqualität Umschlag                    |
| 3.14 | Angabe zur Veredelung                      |

### 3. Daten zur Produktion

|      |                           |
|------|---------------------------|
| 3.15 | Bindungsart               |
| 3.16 | Erstauflage Hardcover     |
| 3.17 | Erstauflage Softcover     |
| 3.18 | Erscheinungstermin-Print  |
| 3.19 | Erscheinungstermin-Online |
| 3.20 | Lieferbarkeitsstatus      |
| 3.21 | Tabellenanzahl S/W        |
| 3.22 | Tabellenanzahl farbig     |
| 3.23 | Rückenstärke              |

### 4. Daten zu Klassifikation

|      |                             |
|------|-----------------------------|
| 4.01 | Warengruppe-Index           |
| 4.02 | Warengruppe                 |
| 4.03 | thema-Klassifikation        |
| 4.04 | BIC                         |
| 4.05 | DDC-Sachgruppe              |
| 4.06 | Kundengruppe / Zielgruppe   |
| 4.07 | URL primäres Repositorium   |
| 4.08 | Schlagwörter (Hauptsprache) |
| 4.09 | Keywords (Zweitsprache)     |

### 5. Daten zu Peer-Review

|      |  |
|------|--|
| 5.01 | Peer-Review-Titel                        |
| 5.02 | Peer-Review-Nr.                          |
| 5.03 | Peer-Review-Verfahren                    |
| 5.04 | Peer-Review-Sprache                      |
| 5.05 | Peer-Review-Überprüfungsdatum            |
| 5.06 | Gutachter-Name                           |
| 5.07 | Gutachter-Rolle                          |
| 5.08 | Institution (die Begutachtung einreicht) |

### 6. Daten zu Barrierefreiheit

|      |  |
|------|--|
| 6.01 | Zugänglichkeitsbeschreibung  |
| 6.02 | Angabe zu menschliche Sinneswahrnehmungen (text, visuell)  |
| 6.03 | Angabe zu physiologischen Gefahren   |
| 6.04 | Konformität zu WCAG 2.0  |
| 6.05 | Konformität zu ISO 14289-1:2014  |
| 6.06 | Anpassungen zur Optimierung der Zugänglichkeit (z.B. vollständig alternative Beschreibungen, Zugang zu math. Inhalten) |

## C.2 Glossar der Metadaten einer Open-Access-Publikation im OA-HVerlag

| Nr.                                  | Datenelement   | Beschreibung   |
|--------------------------------------|--|--|
| <b>1. Daten zu Werk und Lizenzen</b> |  |  |
| 1.01                                 | Werk-Titel (Hauptsprache)  | Titel des Werkes in der Hauptsprache (z.B. deutsch)  |
| 1.02                                 | Werk-Untertitel (Hauptsprache)                                     | Untertitel des Werkes in der Hauptsprache (z.B. deutsch)   |
| 1.03                                 | Werk-Titel (Zweitsprache)  | Titel des Werkes in der Zweitsprache (z.B. englisch)   |
| 1.04                                 | Werk-Untertitel (Zweitsprache)                                     | Untertitel des Werkes in der Zweitsprache (z.B. englisch)  |
| 1.05                                 | Hierarchische bibliografische Relation (Reihenzugehörigkeit)       | stellt die Beziehungen zu einem übergeordneten Gesamtwerk dar (Titelmeldung DNB: fakultativ)                             |
| 1.06                                 | Abstrakt (Hauptsprache)  | Kurzbeschreibung des Inhalts in der Hauptsprache (z.B. deutsch), z.B. bei Titelmeldung KNV                               |
| 1.07                                 | Abstract (Zweitsprache)  | Kurzbeschreibung des Inhalts in der Zweitsprache (z.B. englisch), z.B. bei Titelmeldung KNV                              |
| 1.08                                 | Werk-DOI   | DOI des gesamten Werks   |
| 1.09                                 | Kapitel-DOI  | kapitelweise Zuordnung von DOIs für spezifischere Zitierung  |
| 1.10                                 | Abbildung-DOI  | Zuordnung DOI für jede Abbildung für spezifischere Zitierung, z.B. Crossref-Metadaten                                    |
| 1.11                                 | Tabelle-DOI  | Zuordnung DOI für jede Tabelle für spezifischere Zitierung, z.B. Crossref-Metadaten                                      |
| 1.12                                 | Hardcover-ISBN   | ISBN der gedruckten Hardcover-Ausgabe, z.B. bei Titelmeldung KNV   |
| 1.13                                 | Softcover-ISBN   | ISBN der gedruckten Softcover-Ausgabe, z.B. bei Titelmeldung KNV   |
| 1.14                                 | EPUB/MOBI-ISBN   | ISBN der elektronischen EPUB-Ausgabe, z.B. bei Titelmeldung KNV  |
| 1.15                                 | ePDF-ISBN  | ISBN der elektronischen PDF-Ausgabe, z.B. bei Titelmeldung KNV   |
| 1.16                                 | Produktsprache   | Angabe der Produktsprache bei Titelmeldungen, Verwendung des ISO-Standard für bibliografischen Sprachcode <sup>111</sup> |
| 1.17                                 | Art der Hochschulschrift (Bachelor/Master/Diplom/Forschungsarbeit) | Hinweis auf die Art der Hochschulschrift, z.B. bei Titelmeldungen mit ONIX   |

<sup>111</sup> siehe Liste der ISO-639-1-Codes: [https://de.wikipedia.org/wiki/Liste\\_der\\_ISO-639-1-Codes](https://de.wikipedia.org/wiki/Liste_der_ISO-639-1-Codes)

| <b>Nr.</b> | <b>Datenelement</b>              | <b>Beschreibung</b>   |
|------------|----------------------------------|---|
| 1.18       | Ladenpreis-Hardcover             | Angabe des Hardcover-Ladenpreises, z.B. bei Titelmeldung VLB  |
| 1.19       | Ladenpreis-Softcover             | Angabe des Softcover-Ladenpreises, z.B. bei Titelmeldung VLB  |
| 1.20       | Ladenpreis elektronische Formate | Angabe des Ladenpreises für elektronische Formate (Open Access = kostenfrei)  |
| 1.21       | Auflage                          | Angabe der Auflagennummer (2. Auflage), z.B. bei Titelmeldung VLB (i. d. R. keine Angabe zur Print-Auflagenhöhe möglich)  |
| 1.22       | Art der Auflage                  | Angabe zur Art der Auflage, z.B. bei Titelmeldung VLB   |
| 1.23       | Publikationsort                  | hier wird der Ort angegeben, an dem der Titel veröffentlicht wird (Angabe des Publikationsortes ist wiederholbar, da ein Titel an mehreren Orten gleichzeitig erscheinen kann) <sup>112</sup> , z.B. bei Titelmeldung VLB   |
| 1.24       | Erscheinungsland                 | kann alternativ oder ergänzend zum Publikationsort angegeben werden (Angabe Erscheinungslandes ist nicht wiederholbar) <sup>113</sup> , z.B. bei Titelmeldung VLB   |
| 1.25       | Publikationsstatus               | dient dem Anzeigen der Verfügbarkeit von Produkten gegenüber dem Handel sowie Kunden und ist eng mit dem Erscheinungstermin verbunden (wenn der Erscheinungstermin eines Produktes erreicht ist, sollte der Publikationsstatus als "erschienen" gemeldet werden) <sup>114</sup> , z.B. bei Titelmeldung VLB |
| 1.26       | Förderer                         | Angabe zum Förderer der Publikation   |
| 1.27       | Cover-Bild                       | Einbettung des Cover-Bilds oder der URL zum Cover-Bild  |
| 1.28       | Copyright-Information            | Angabe zu Copyright-Informationen   |
| 1.29       | CC-Lizenz                        | Angabe zur verwendeten CC-Lizenz  |
| 1.30       | Access-Status                    | Angabe zum Access-Status (open / closed)  |
| 1.31       | Art des Urhebers                 | Angabe zur Art des Urhebers (Autor, Herausgeber, Übersetzer)  |
| 1.32       | Rechteinhaber                    | Angabe zum Urheber  |
| 1.33       | Auftragsnummer                   | Nummer zur Identifizierung des Auftrages  |

---

112 VLB 2023a

113 ebd.

114 VLB 2023b

| <b>Nr.</b>                      | <b>Datenelement</b>  | <b>Beschreibung</b>  |
|---------------------------------|----------------------|--|
| <b>2. Daten zu Mitwirkenden</b> |                      |  |
| 2.01                            | Autor-Name           | Name des Autors  |
| 2.02                            | Autor-Vorname        | Vorname des Autors   |
| 2.03                            | Autor-Titel          | akademischer Titel des Autors  |
| 2.04                            | Autor-Wohnadresse    | Anschrift des Autors für für Publikationsvertrag und Zusendung von Autorenexemplaren                           |
| 2.05                            | Autor-PLZ            | PLZ des Autors für Publikationsvertrag und Zusendung von Autorenexemplaren                                     |
| 2.06                            | Autor-Wohnort        | Wohnort des Autors für Publikationsvertrag und Zusendung von Autorenexemplaren                                 |
| 2.07                            | Autor-Institution    | Institution des Autors   |
| 2.08                            | Autor-Telefon-Nr.    | Telefonnummer des Autors zur Kontaktierung bei Fragen im Publikationsprozess                                   |
| 2.09                            | Autor-E-Mail-Adresse | E-Mail-Adresse des Autors zur Kontaktierung bei Fragen im Publikationsprozess                                  |
| 2.10                            | Autor-Kurzvita       | kurzer Beschreibungstext zum Autor, der bspw. im Werk abgedruckt sein kann                                     |
| 2.11                            | Autor-ORCID-ID       | ORCID-ID zur eindeutigen Autor-Identifikation, welcher bei Titelmeldungen, wie z.B. VLB, angegeben werden kann |
| 2.12                            | Autor-ISNI           | ISNI zur eindeutigen Autor-Identifikation, welcher bei Titelmeldungen, wie z.B. VLB, angegeben werden kann     |
| 2.13                            | Autor-Autorenbild    | Bild des Autors (wird manchmal auch im Werk abgedruckt)  |
| 2.14                            | Autor-Bankverbindung | Bankverbindung des Autors für die Auszahlung eines möglichen Autorenhonorars                                   |
| 2.15                            | Beteiligte-Name      | Name des am Werk Beteiligten (Illustrator, Gutachter, Lektor usw.)   |
| 2.16                            | Beteiligte-Vorname   | Vorname des am Werk Beteiligten (Illustrator, Gutachter, Lektor usw.)  |
| 2.17                            | Beteiligte-Titel     | akad. Titel des am Werk Beteiligten (Illustrator, Gutachter, Lektor usw.)                                      |
| 2.18                            | Beteiligte-E-Mail    | E-Mail-Adresse des am Werk Beteiligten (Illustrator, Gutachter, Lektor usw.)                                   |

| Nr.  | Datenelement            | Beschreibung  |
|------|-------------------------|---|
| 2.19 | Beteiligte-Institution  | Institution des am Werk Beteiligten (Illustrator, Gutachter, Lektor usw.) |
| 2.20 | Herausgeber-Name        | Name des Herausgebers   |
| 2.21 | Herausgeber-Adresse     | Anschrift des Herausgebers  |
| 2.22 | Herausgeber-PLZ         | PLZ des Herausgebers  |
| 2.23 | Herausgeber-Land        | Land des Herausgebers   |
| 2.24 | Herausgeber-E-Mail      | E-Mail-Adresse des Herausgebers   |
| 2.25 | Herausgeber-Website-URL | Verlagswebsite des Herausgebers   |

### 3. Daten zur Produktion

|      |                                    |  |
|------|------------------------------------|--|
| 3.01 | Produktformen                      | Beschreibung der Produktformen der Publikation (Hardcover, Softcover, EPUB, PDF); bei Titelanmeldungen helfen die Angaben dem Handel, um die Art des Produktes zu erkennen und eine kundenorientierte Produkteinordnung bzw. -platzierung vorzunehmen <sup>115</sup> |
| 3.02 | Datei der elektronischen Ressource | Angabe der URL der elektronischen Ressource des Werkes für das automatische Abliefern über OAI-Harvesting-Verfahren, welche direkt zur elektronischen Ressource führt (Transfer-URL) <sup>116</sup> , Angabe bei Titelmeldung dient zur Abholung der Netzpublikation |
| 3.03 | EPUB-Version                       | Version des EPUB-Formats (EPUB 2, EPUB 3 usw.)   |
| 3.04 | Umfang Print Insgesamt             | Umfang der gedruckten Seiten (Angabe beim Druckdienstleister)  |
| 3.05 | Umfang Print Römisch               | Umfang der römischen, gedruckten Seiten (Verzeichnisse, Frontmatter)   |
| 3.06 | Umfang Print Arabisch              | Umfang der arabischen, gedruckten Seiten (eigentlicher Inhalt)   |
| 3.07 | Abbildungsanzahl-S/W               | Anzahl an SW-Abbildungen (Angabe beim Druckdienstleister)  |
| 3.08 | Abbildungsanzahl-farbig            | Anzahl an farbigen Abbildungen (Angabe beim Druckdienstleister)  |
| 3.09 | Print-Format-Breite                | Breite des gedruckten Formats  |

<sup>115</sup> VLB 2023c

<sup>116</sup> DNB 2018, S.10

## 112 Anhang C

| <b>Nr.</b> | <b>Datenelement</b>       | <b>Beschreibung</b>   |
|------------|---------------------------|---|
| 3.10       | Print-Format-Höhe         | Höhe des gedruckten Formats                                   |
| 3.11       | Print-Farbigkeit          | Farbigkeit der gedruckten Inhaltsseiten                       |
| 3.12       | Papierqualität Innenteil  | Art des Inhaltspapiers, flächenbez. Masse                     |
| 3.13       | Papierqualität Umschlag   | Art des Umschlagpapiers, flächenbez. Masse                    |
| 3.14       | Angabe zur Veredelung     | Angabe zur Veredelung des Buchumschlags (z.B. Spotlackierung) |
| 3.15       | Bindungsart               | Angabe zur Bindung (z.B. Fadenheftung, Klebebindung)          |
| 3.16       | Erstauflage Hardcover     | Auflage der gedruckten Hardcover                              |
| 3.17       | Erstauflage Softcover     | Auflage der gedruckten Softcover                              |
| 3.18       | Erscheinungstermin-Print  | Erscheinungsdatum der gedruckten Ausgabeform                  |
| 3.19       | Erscheinungstermin-Online | Erscheinungsdatum der elektronischen Ausgabeform              |
| 3.20       | Lieferbarkeitsstatus      | Angabe zur Lieferbarkeit (noch nicht lieferbar, lieferbar)    |
| 3.21       | Tabellenanzahl S/W        | Angabe zur Anzahl einfarbiger (1c) Tabellen                   |
| 3.22       | Tabellenanzahl farbig     | Angabe zur Anzahl farbiger (4c) Tabellen                      |
| 3.23       | Rückenstärke              | Angabe zur Stärke des Buchrückens                             |

### 4. Daten zu Klassifikation

|      |                      |  |
|------|----------------------|--|
| 4.01 | Warengruppe-Index    | Angabe zur Editionsform der Publikation (Bestandteil der Warengruppe), z.B. 1 = Hardcover, Softcover, 7 = Kalender   |
| 4.02 | Warengruppe          | dient der statistischen Analyse unterschiedlicher Titel-Segmente im Buchhandel und besteht aus einer vierstelligen Zahl (1. Ziffer = Warengruppen-Index / 2.-4. Ziffer = Warengruppennummer) |
| 4.03 | thema-Klassifikation | internationale Thema-Klassifikation mit derzeit knapp 7.200 Kategorien, um Produkt passgenau zu beschreiben  |
| 4.04 | BIC                  | Angabe, mit der ein Titel klassifiziert werden kann ist eine in Großbritannien genutzte Klassifikation, welche als Grundlage für die thema-Klassifikation diene                              |

| <b>Nr.</b> | <b>Datenelement</b>         | <b>Beschreibung</b>   |
|------------|-----------------------------|---|
| 4.05       | DDC-Sachgruppe              | Notation für klassifikatorische Inhaltserschließung in der Deutschen Nationalbibliothek, DDC-Sachgruppen gliedern die Deutsche Nationalbibliografie |
| 4.06       | Kundengruppe / Zielgruppe   | Angabe zur Zielgruppe (z.B. Fachpublikum/Wissenschaftler; junge Erwachsene)   |
| 4.07       | URL primäres Repositorium   | Link zum primären Repositorium  |
| 4.08       | Schlagwörter (Hauptsprache) | Schlagwörter in der Hauptsprache  |
| 4.09       | Keywords (Zweitsprache)     | Schlagwörter in der Zweitsprache  |

### 5. Daten zu Peer Review

|      |  |   |
|------|--|---|
| 5.01 | Peer-Review-Titel                        | Titel des Peer-Review   |
| 5.02 | Peer-Review-Nr.                          | Interne Nummer/Kennung zur Identifizierung einer bestimmten Bewertung   |
| 5.03 | Peer-Review-Verfahren                    | Angabe zum gewählten Verfahren des Peer-Review (z.B. double-blind review)   |
| 5.04 | Peer-Review-Sprache                      | Angabe zur Sprache, in der das Peer-Review durchgeführt wurde   |
| 5.05 | Peer-Review-Überprüfungsdatum            | Datum der Überprüfung, einschließlich Monat, Tag, Jahr  |
| 5.06 | Gutachter-Name                           | Name des Gutachters   |
| 5.07 | Gutachter-Rolle                          | Rolle des Gutachters (Rollen können sein: Reviewer, Review-Assistent, Statistik-Reviewer, Reviewer-extern, Leser, Übersetzer, Autor, Redakteur) |
| 5.08 | Institution (die Begutachtung einreicht) | Angabe zur Institution, die Begutachtung einreicht  |

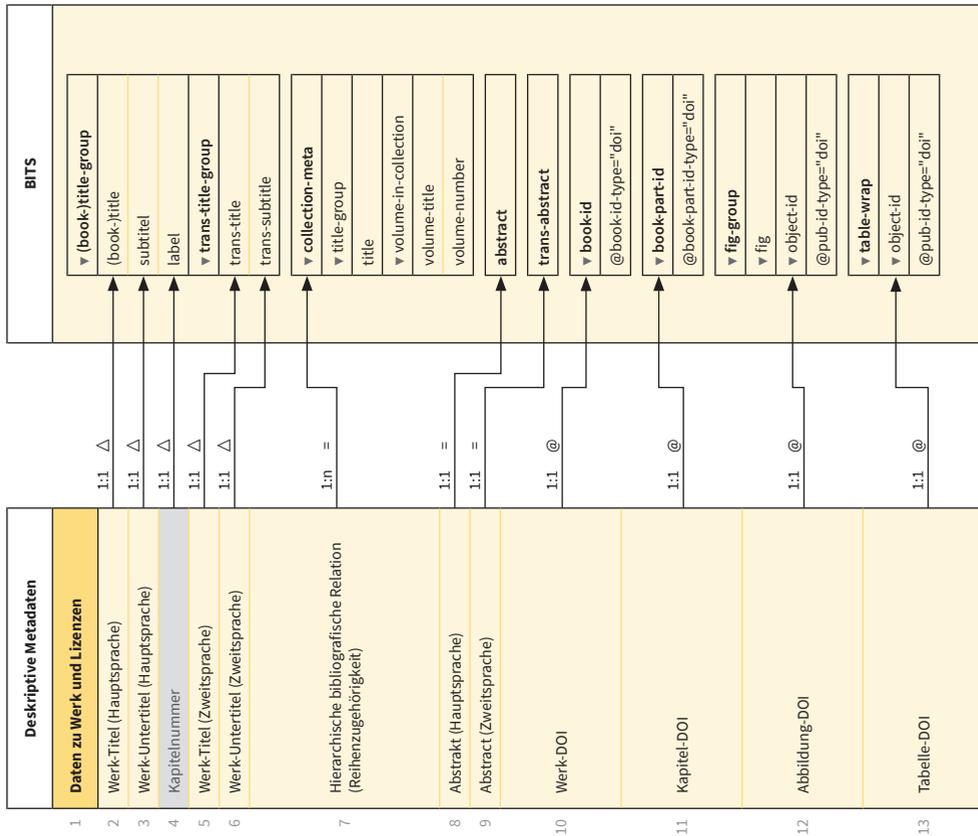
| Nr.                                 | Datenelement   | Beschreibung  |
|-------------------------------------|--|---|
| <b>6. Daten zu Barrierefreiheit</b> |  |   |
| 6.01                                | Zugänglichkeitsbeschreibung  | Zusammenfassung der Gesamtzugänglichkeit der Publikation  |
| 6.02                                | Angabe zu menschlichen Sinneswahrnehmungen (Text, visuell)                                   | Angabe zu menschliche Sinneswahrnehmungen, die standardmäßig zur Wahrnehmung der Inhalte einer Publikation benötigt werden              |
| 6.03                                | Angabe zu physiologischen Gefahren   | beschreibt die Eigenschaft der Ressource, die für einige Benutzer physiologisch gefährlich ist (z.B. Flimmern)                          |
| 6.04                                | Konformität zu WCAG 2.0  | Angabe zur Konformität zu WCAG, ein internationaler Standard zur barrierefreien Gestaltung von Internetangeboten der Europäischen Union |
| 6.05                                | Konformität zu ISO 14289-1:2014  | beschreibt die Anforderungen von ISO 32000-1:2008 plus ISO 14289-1:2014 (PDF/A), nur zur Verwendung in ONIX 3.0                         |
| 6.06                                | Anpassungen zur Optimierung der Zugänglichkeit (z.B. vollständig alternative Beschreibungen) | Angabe zu Anpassungen zur Optimierung der Zugänglichkeit, z.B. Alternativtexte  |

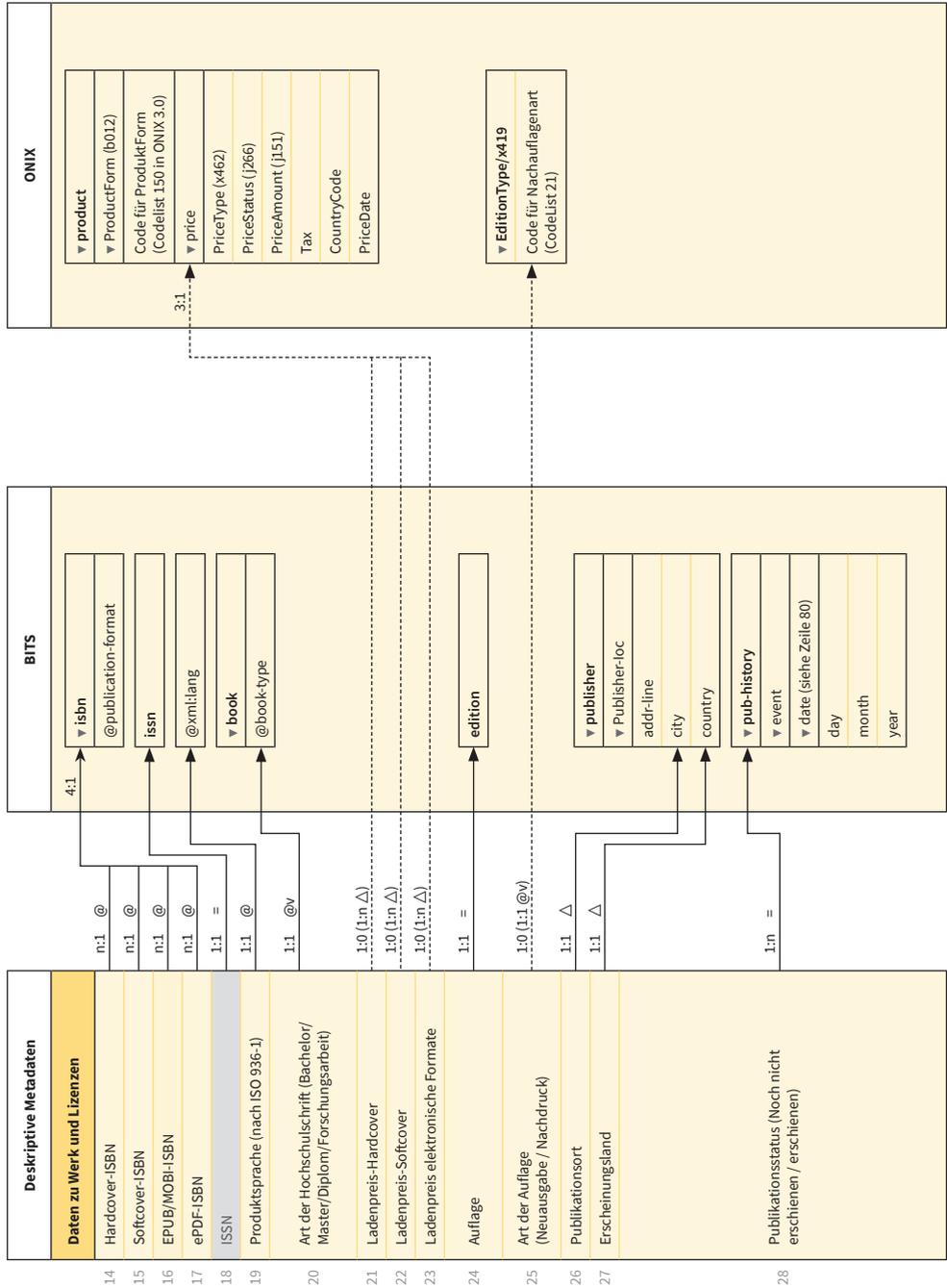


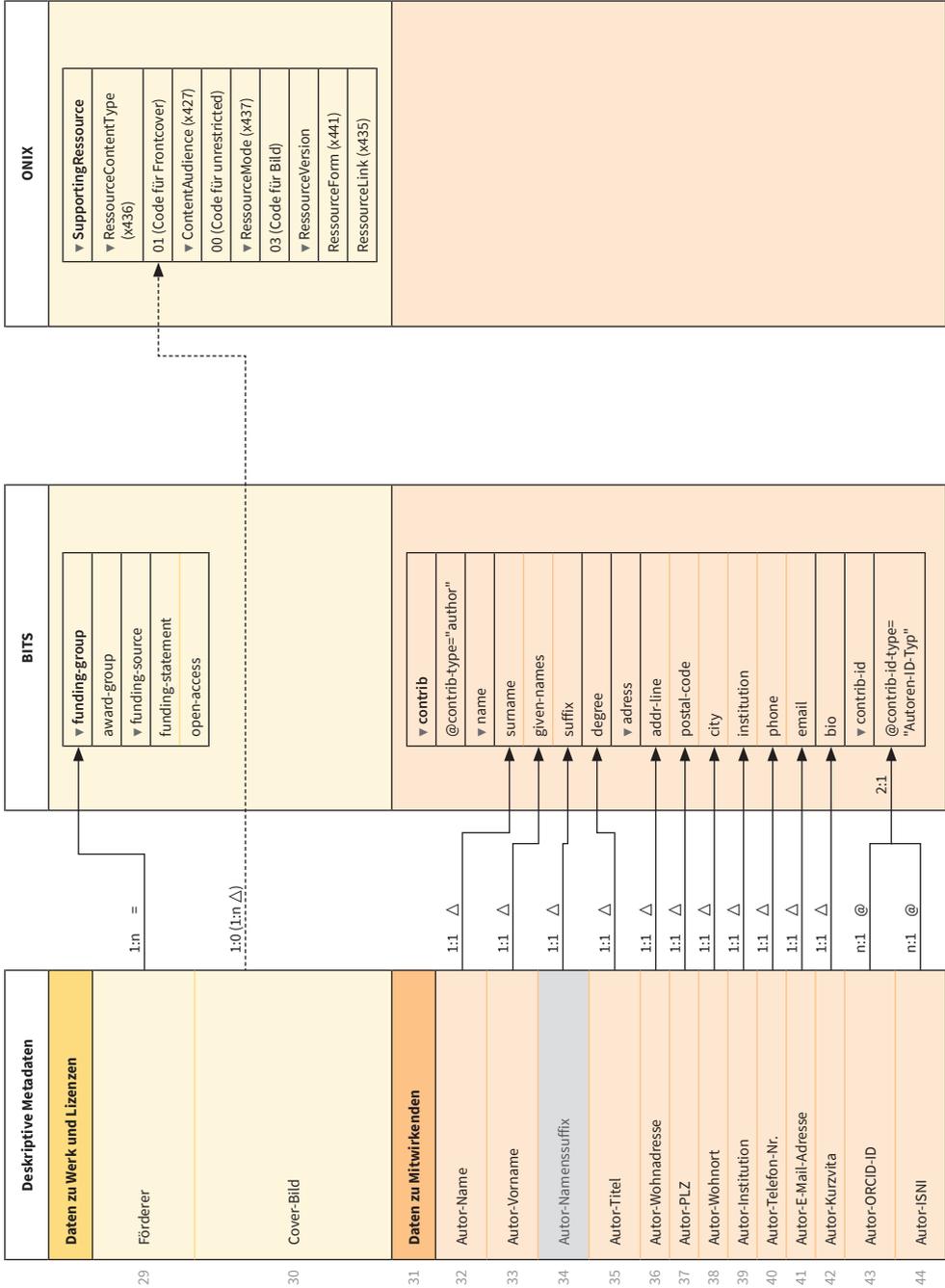
# Anhang D

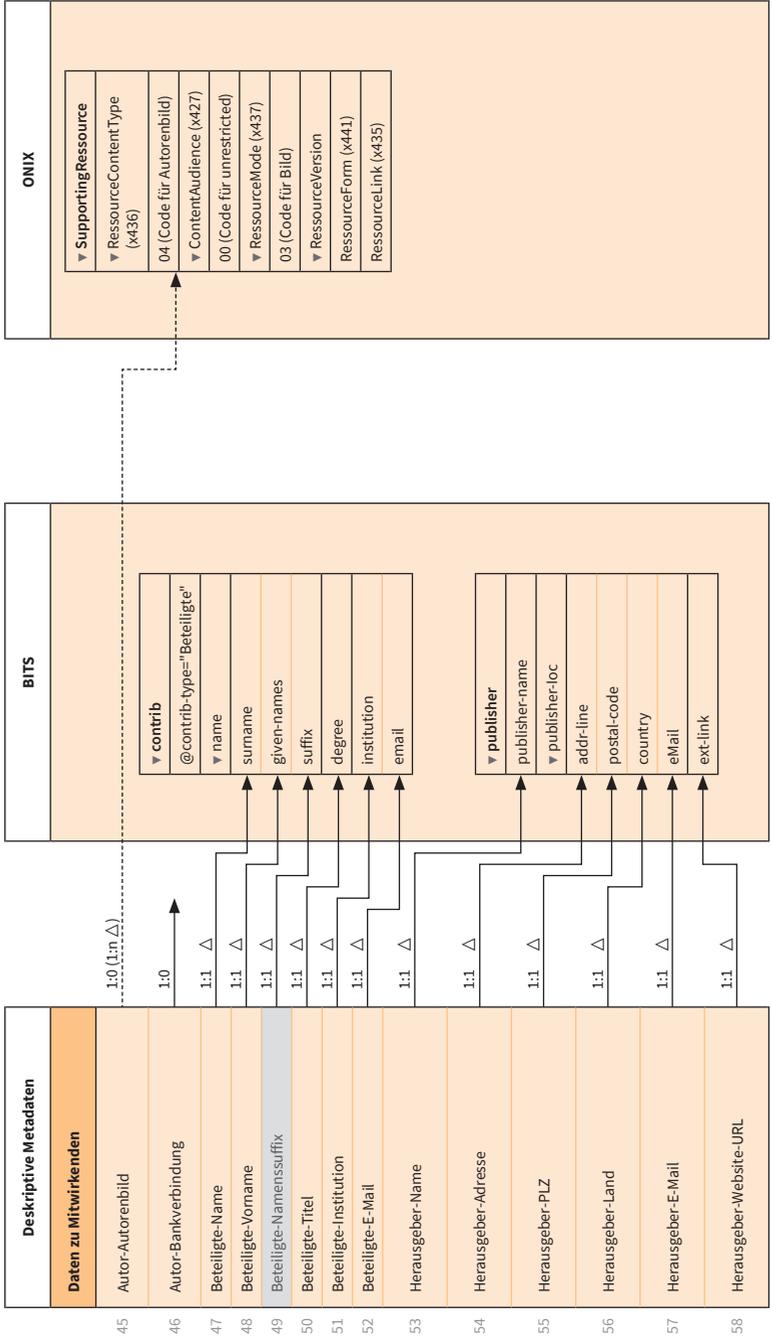
---

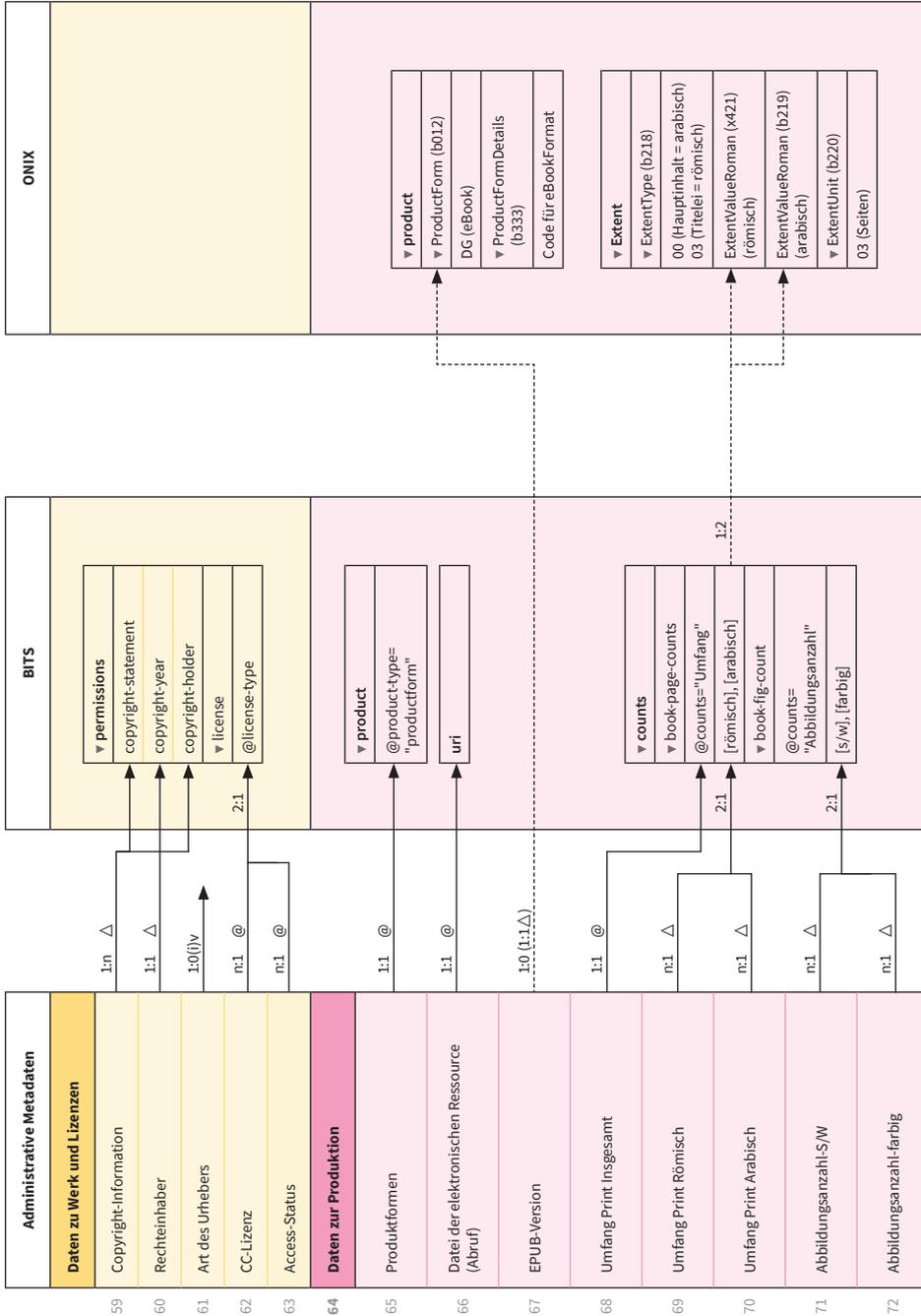
## Transformationsregeln

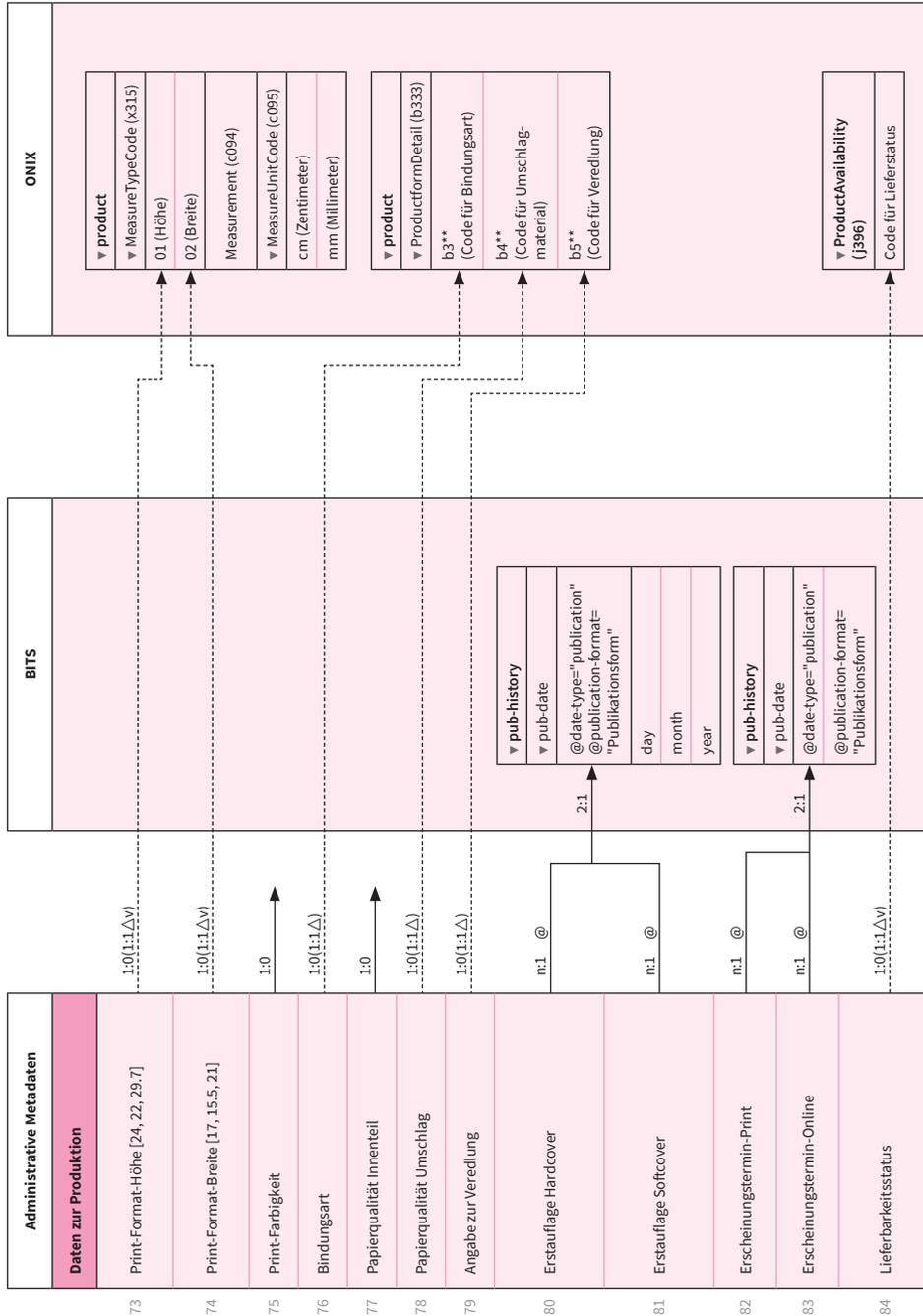


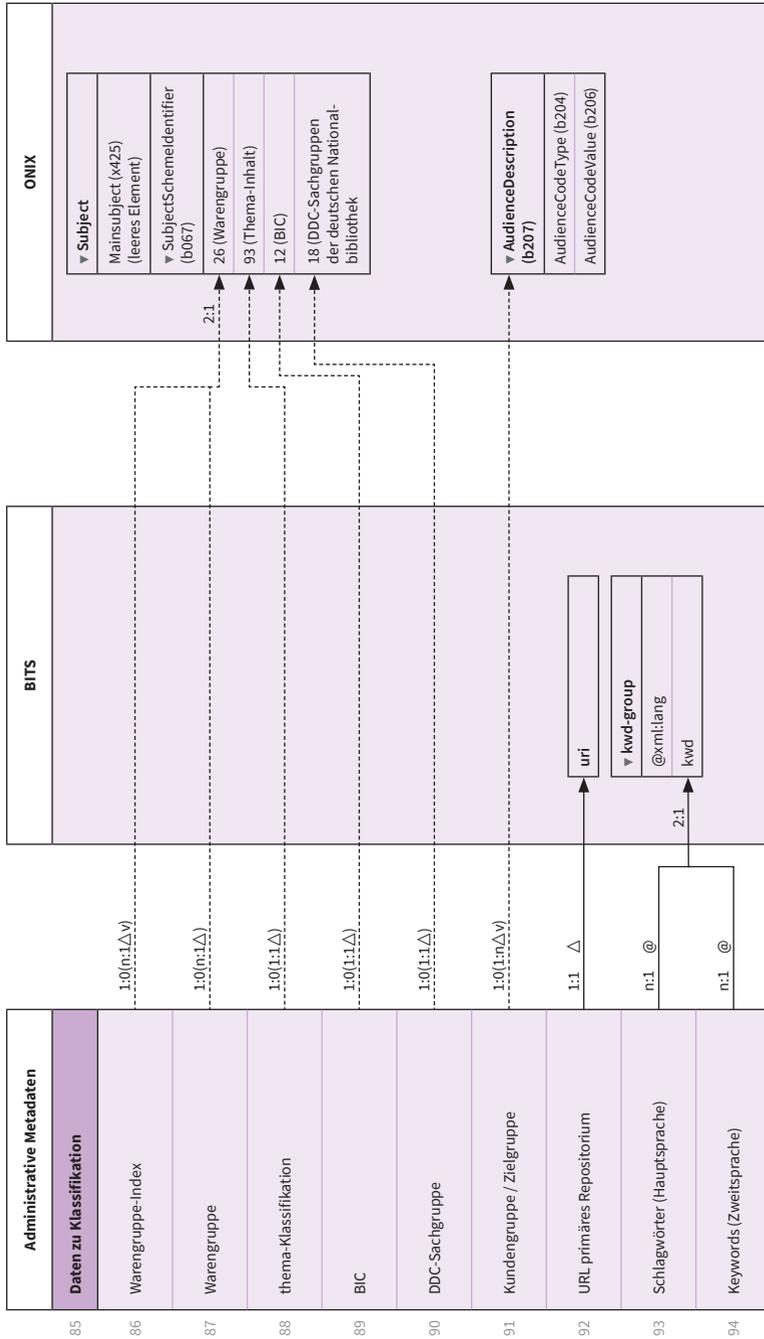


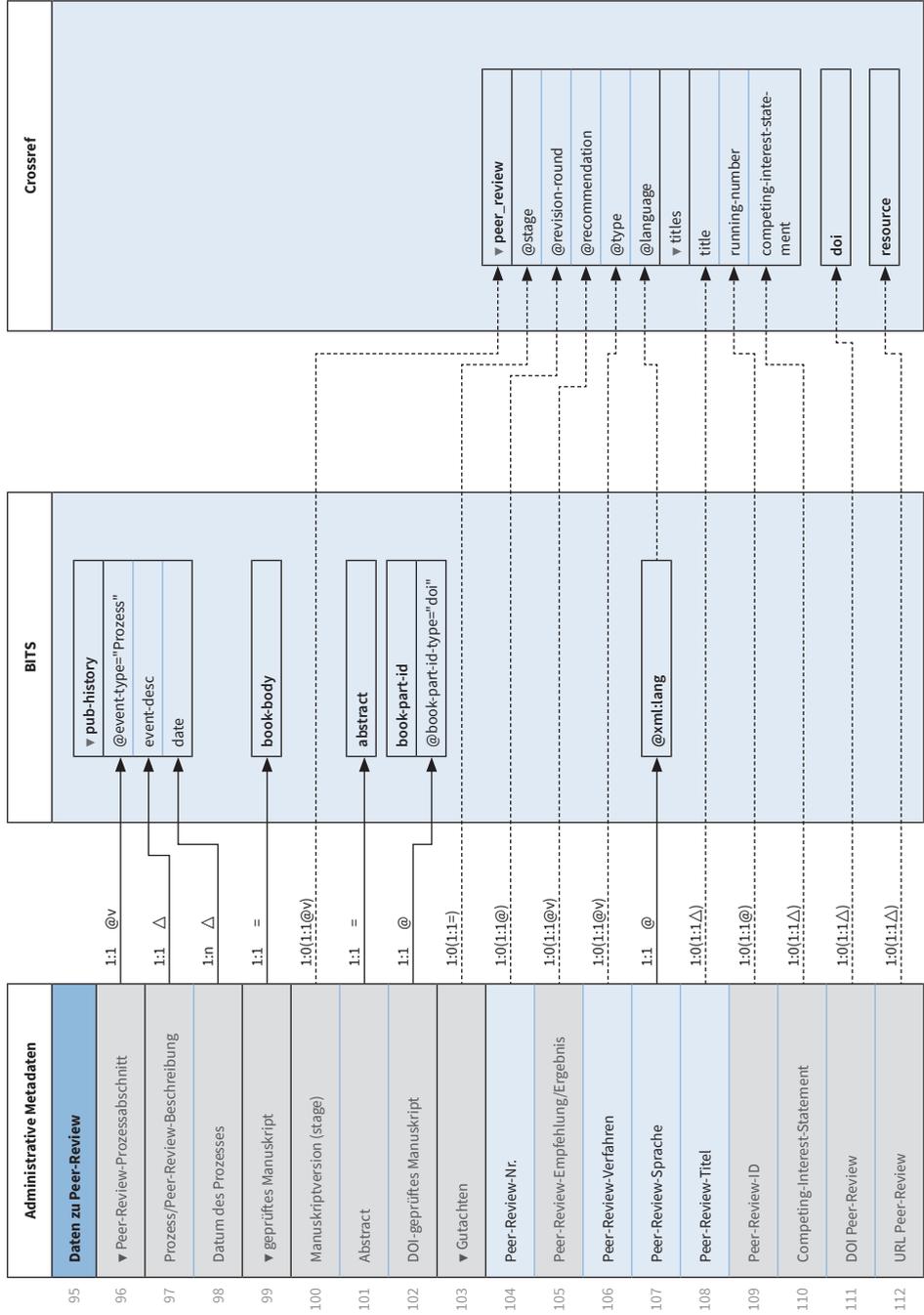


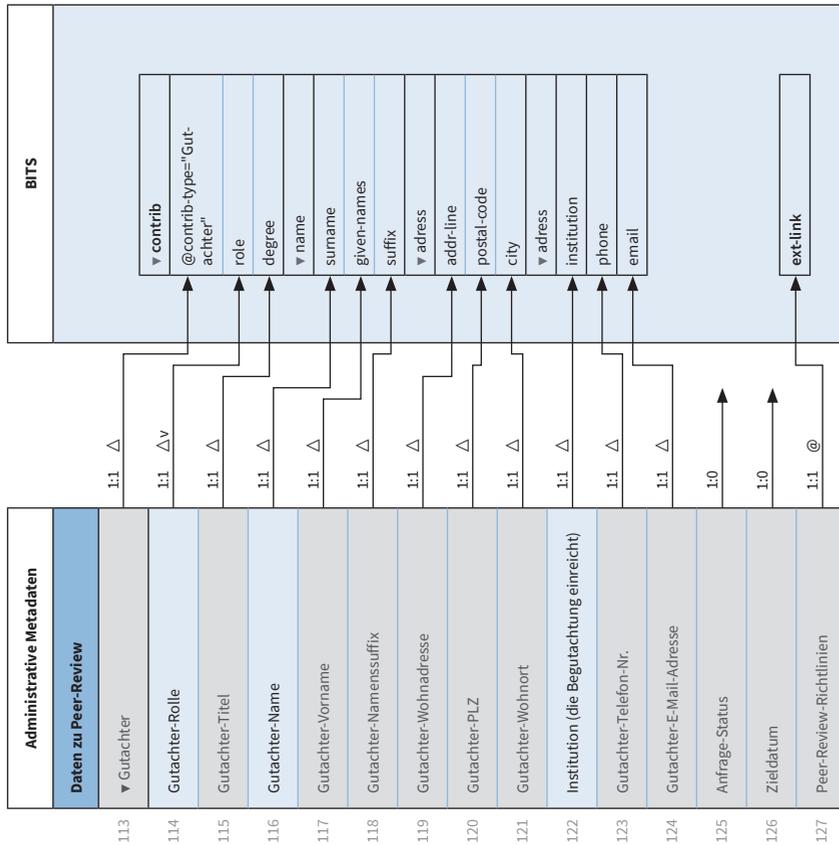


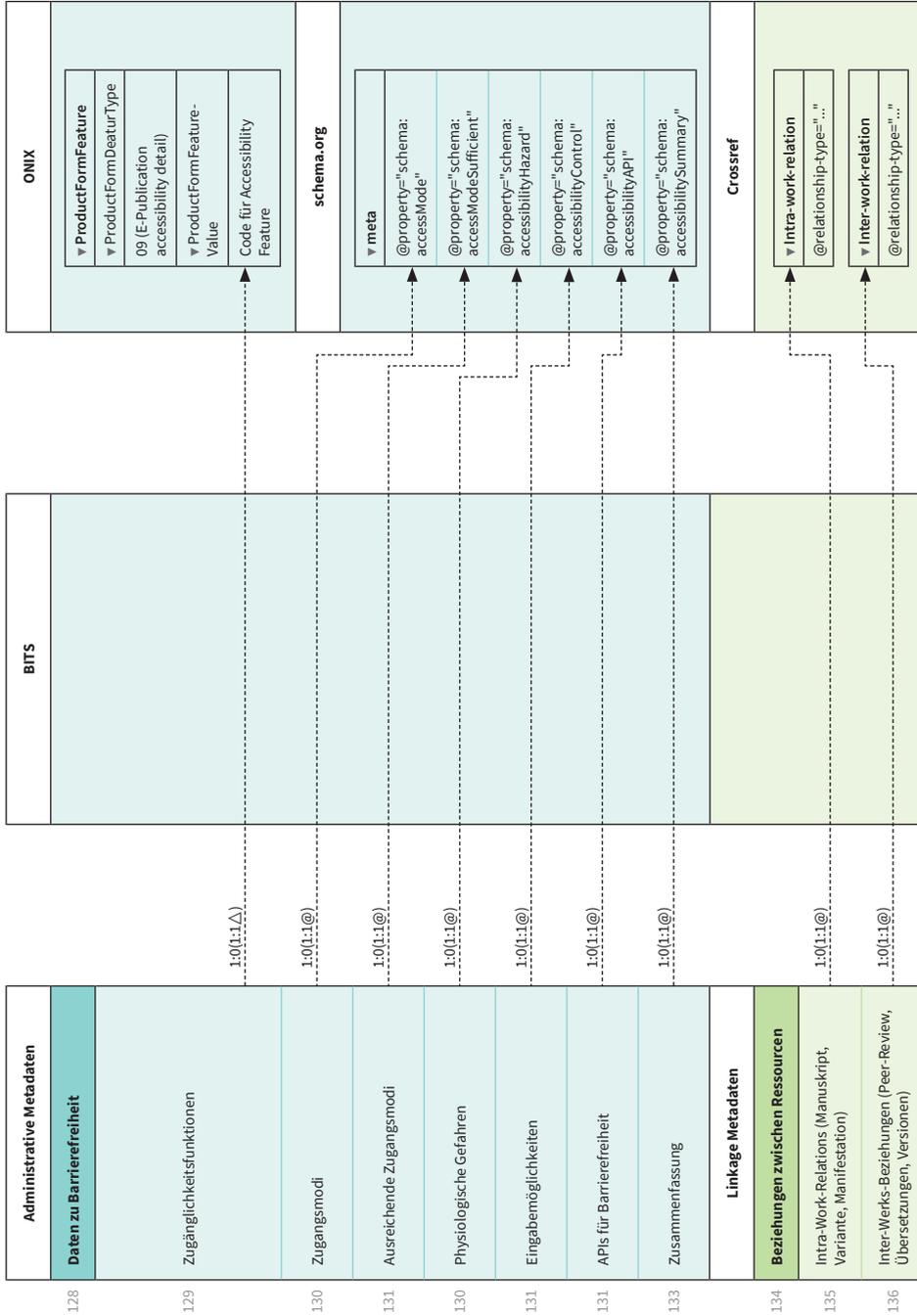














# Anhang E

---

Projektsteuerdatei projekt.xml

```

1 <?xml version="1.0" encoding="utf-8"?>
2 <projekt xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
3     xsi:noNamespaceSchemaLocation="Struktur/projekt.xsd">
4
5     <projekt-metadaten>
6         <historie>
7             <event type="Erstellung" datum="2021-07-29"/>
8         </historie>
9     </projekt-metadaten>
10
11     <content>
12         <instanz id="1"
13             href="Content/BITS_Musterbuch.xml"
14             schema="Struktur/BITS-book2.dtd" />
15         <instanz id="2"
16             href="Content/Onix2.1_Musterbuch_Monografie.xml"
17             schema="Struktur/ONIX_Release2.1_reference.xsd" />
18         <instanz id="3"
19             href="Content/Crossref_Musterbuch_Monografie.xml"
20             schema="Struktur/crossref4.4.0.xsd" />
21         <instanz id="9"
22             href="Content/Crossref_Peer-Review_Musterbuch_Monografie.xml"
23             schema="Struktur/crossref4.4.2.xsd" />
24     </content>
25
26     <metadaten>
27         <meta var="101" name="Werk-Titel-Hauptsprache" id_ref="1">
28             book/book-meta/book-title-group/book-title</meta>
29         <meta var="102" name="Werk-Untertitel-Hauptsprache" id_ref="1">
30             book/book-meta/book-title-group/subtitle</meta>
31         <meta var="103" name="Werk-Titel-Zweitsprache" id_ref="1">
32             book/book-meta/book-title-group/trans-title-group/trans-title</meta>
33         <meta var="104" name="Werk-Untertitel-Zweitsprache" id_ref="1">
34             book/book-meta/book-title-group/trans-title-group/trans-subtitle</meta>
35         <meta var="105" name="Abstrakt-Hauptsprache" id_ref="2">
36             ONIXMessage/Product/OtherText[@language='ger']</meta>
37         <meta var="106" name="Abstrakt-Zweitsprache" id_ref="2">
38             ONIXMessage/Product/OtherText[@language='eng']</meta>
39         <meta var="108" name="Werk-DOI" id_ref="1">book/book-meta/book-id[@book-id-type='doi']
40         </meta>
41         <meta var="109" name="Kapitel1-DOI" id_ref="3">
42             doi_batch/body/content_item[@component_type='chapter']/doi_data/doi</meta>
43         <meta var="109" name="Kapitel2-DOI" id_ref="3">
44             doi_batch/body/content_item[@component_type='chapter']/doi_data/doi</meta>
45         <meta var="109" name="Kapitel3-DOI" id_ref="3">
46             doi_batch/body/content_item[@component_type='chapter']/doi_data/doi</meta>
47         <meta var="109" name="Kapitel4-DOI" id_ref="3">
48             doi_batch/body/content_item[@component_type='chapter']/doi_data/doi</meta>
49         <meta var="109" name="Kapitel5-DOI" id_ref="3">
50             doi_batch/body/content_item[@component_type='chapter']/doi_data/doi</meta>
51         <meta var="110" name="Abbildung1-DOI" id_ref="3">
52             /component_list/component[@parent_relation='isPartOf']/doi_data/doi</meta>
53         <meta var="110" name="Abbildung2-DOI" id_ref="3">
54             doi_batch/body/content_item[@component_type='chapter']
55             /component_list/component[@parent_relation='isPartOf']/doi_data/doi</meta>
56         <meta var="111" name="Tabelle1-DOI" id_ref="3">
57             doi_batch/body/content_item[@component_type='chapter']
58             /component_list/component[@parent_relation='isPartOf']/doi_data/doi</meta>
59         <meta var="111" name="Tabelle2-DOI" id_ref="3">
60             doi_batch/body/content_item[@component_type='chapter']
61             /component_list/component[@parent_relation='isPartOf']/doi_data/doi</meta>
62         <meta var="112" name="Hardcover-ISBN" id_ref="1">
63             book/book-meta/isbn[@content-type='hardcover']</meta>
64         <meta var="113" name="Softcover-ISBN" id_ref="1">
65             book/book-meta/isbn[@content-type='softcover']</meta>
66         <meta var="114" name="EPUB-MOBI-ISBN" id_ref="1">
67             book/book-meta/isbn[@content-type='epub']</meta>
68         <meta var="115" name="ePDF-ISBN" id_ref="1">book/book-meta/isbn[@content-type='pdf']
69     </metadaten>
70 </projekt>

```

```

51     </meta>
52     <meta var="116" name="Produktsprache" id_ref="1">
53     ONIXMessage/Product/Language/LanguageCode</meta>
54     <meta var="117" name="Art-der-Hochschulschrift" id_ref="2">
55     ONIXMessage/Product/ThesisType</meta>
56     <meta var="118" name="Ladenpreis-Hardcover" id_ref="2">
57     ONIXMessage/Product/SupplyDetail/Price/PriceAmount</meta>
58     <meta var="119" name="Ladenpreis-Softcover" id_ref="2">
59     ONIXMessage/Product/SupplyDetail/Price/PriceAmount</meta>
60     <meta var="120" name="Ladenpreis-elektronische-Formate" id_ref="2">
61     ONIXMessage/Product/SupplyDetail/Price/PriceAmount</meta>
62     <meta var="121" name="Auflage" id_ref="2">ONIXMessage/Product/EditionNumber</meta>
63     <meta var="122" name="Art-der-Auflage" id_ref="2">ONIXMessage/Product/EditionTypeCode
64     </meta>
65     <meta var="123" name="Publikationsort" id_ref="2">ONIXMessage/Product/CityOfPublication
66     </meta>
67     <meta var="124" name="Erscheinungsland" id_ref="2">
68     ONIXMessage/Product/CountryOfPublication</meta>
69     <meta var="125" name="Publikationsstatus" id_ref="2">
70     ONIXMessage/Product/PublishingStatus</meta>
71     <meta var="126" name="Förderer" id_ref="2">ONIXMessage/Product/Publisher/NameCodeValue
72     </meta> <!--PublishingRole='4' (Förderer)-->
73     <meta var="127" name="Coverbild" id_ref="2">ONIXMessage/Product/MediaFile/MediaFileLink
74     </meta>
75     <meta var="128" name="Reihenzugehörigkeit" id_ref="1">
76     book/book-meta/collection-meta/@collection-type</meta>
77     <meta var="129" name="Buch-Typ" id_ref="1">book/@book-type</meta>
78     <meta var="130" name="Buch-URL" id_ref="1">book/book-meta/self-uri</meta>
79
80     <meta var="131" name="Kapitel-Titel" id_ref="1">
81     book/book-body//book-part[@book-part-type='chapter']/book-part-meta/title-group/title
82     </meta>
83
84     <meta var="201" name="Autor-Name" id_ref="1">
85     book/book-meta/contrib-group/contrib[@contrib-type='content-supplier']/name/surname
86     </meta>
87     <meta var="202" name="Autor-Vorname" id_ref="1">
88     book/book-meta/contrib-group/contrib[@contrib-type='content-supplier']/name/given-names
89     </meta>
90     <meta var="203" name="Autor-Titel" id_ref="1">
91     book/book-meta/contrib-group/contrib[@contrib-type='content-supplier']/name/prefix
92     </meta>
93     <meta var="209" name="Autor-E-Mail-Adresse" id_ref="1">
94     book/book-meta/contrib-group/contrib[@contrib-type='content-supplier']/email</meta>
95     <meta var="210" name="Autor-Kurzvita" id_ref="1">
96     book/book-meta/contrib-group/contrib[@contrib-type='content-supplier']/bio/p</meta>
97     <meta var="212" name="Autor-ISBN" id_ref="2">
98     ONIXMessage/Product/PersonNameIdentifier/IDValue</meta>
99     <meta var="301" name="Beteiligte-Name" id_ref="1">
100    book/book-meta/contrib-group/contrib[@contrib-type='reviewer_1']/name/surname</meta>
101    <meta var="302" name="Beteiligte-Nachname" id_ref="1">
102    book/book-meta/contrib-group/contrib[@contrib-type='reviewer_1']/name/given-names</meta>
103    <meta var="303" name="Beteiligte-Titel" id_ref="1">
104    book/book-meta/contrib-group/contrib[@contrib-type='reviewer_1']/degrees</meta>
105    <meta var="304" name="Beteiligte-E-Mail-Adresse" id_ref="1">
106    book/book-meta/contrib-group/contrib[@contrib-type='reviewer_1']/email</meta>
107    <meta var="305" name="Beteiligte-Institution" id_ref="1">
108    book/book-meta/contrib-group/contrib[@contrib-type='reviewer_1']/xref[@id]</meta>
109    <meta var="401" name="Herausgeber-Name" id_ref="1">
110    book/book-meta/publisher/publisher-name</meta>
111    <meta var="402" name="Herausgeber-Adresse" id_ref="1">
112    book/book-meta/publisher/publisher-loc/addr-line</meta>
113    <meta var="403" name="Herausgeber-PLZ" id_ref="1">
114    book/book-meta/publisher/publisher-loc/postal-code</meta>
115    <meta var="404" name="Herausgeber-Land" id_ref="1">
116    book/book-meta/publisher/publisher-loc/country</meta>

```

```

87 <meta var="405" name="Herausgeber-E-Mail" id_ref="1">
book/book-meta/publisher/publisher-loc/email</meta>
88 <meta var="502" name="Datei-der-eElektronischen-Ressource" id_ref="3">
doi_batch/body/book[@book_type='monograph']/book_metadata/doi_data/resource</meta>
89 <meta var="504" name="Umfang-Print-Gesamt" id_ref="2">ONIXMessage/Product/NumberOfPages
</meta>
90 <meta var="509" name="Print-Format-Breite" id_ref="2">
ONIXMessage/Product/Measure/Measurement</meta>
91 <meta var="510" name="Print-Format-Höhe" id_ref="2">
ONIXMessage/Product/Measure/Measurement</meta>
92
93 <meta var="601" name="Erscheinungstermin-Print" id_ref="1">
book/book-meta/pub-date[@date-type='print']</meta>
94 <meta var="602" name="Erscheinungstermin-Online" id_ref="1">
book/book-meta/pub-date[@date-type='online']</meta>
95
96
97 <meta var="603" name="Lieferbarkeitsstatus" id_ref="2">
ONIXMessage/Product/SupplyDetail/ProductAvailability</meta>
98 <meta var="604" name="Warengruppe-Index" id_ref="2">
ONIXMessage/Product/mainsubject/SubjectCode</meta>
<!--MainSubjectSchemeIdentifier='26'-->
99 <meta var="605" name="Warengruppe" id_ref="2">
ONIXMessage/Product/mainsubject/SubjectHeadingText</meta>
<!--MainSubjectSchemeIdentifier='26'-->
100 <meta var="606" name="thema-Klassifikation" id_ref="2">
ONIXMessage/Product/mainsubject/SubjectCode</meta> <!--SubjectSchemeIdentifier='93'-->
101
102 <meta var="607" name="BIC" id_ref="2">ONIXMessage/Product/mainsubject/SubjectCode</meta>
<!--SubjectSchemeIdentifier='12'-->
102 <meta var="608" name="DDC-Sachgruppe" id_ref="2">
ONIXMessage/Product/Subject[@sourcename='DNB']/SubjectCode</meta>
103 <meta var="609" name="Kundengruppe" id_ref="2">ONIXMessage/Product/AudienceCodeType
</meta>
104 <meta var="611" name="Schlagwörter-Hauptsprache" id_ref="1">
book/book-meta/kwd-group[@xml:lang='de']/kwd</meta>
105 <meta var="612" name="Schlagwörter-Zweitsprache" id_ref="1">
book/book-meta/kwd-group[@xml:lang='en']/kwd</meta>
106 <meta var="801" name="Copyright-Inhaber" id_ref="1">
book/book-meta/permissions/copyright-holder</meta>
107 <meta var="802" name="Copyright-Jahr" id_ref="1">
book/book-meta/permissions/copyright-year</meta>
108
109
110 <meta var="901" name="Peer-Review-DOI" id_ref="9">
doi_batch/body/peer_review/doi_data/doi</meta>
111 <meta var="902" name="Peer-Review-URL" id_ref="9">
doi_batch/body/peer_review/doi_data/resource</meta>
112 <meta var="903" name="Peer-Review-Datum" id_ref="9">
doi_batch/body/peer_review/review_date</meta>
113 <meta var="904" name="Peer-Review-Titel" id_ref="9">
doi_batch/body/peer_review/titles/title</meta>
114 </metadaten>
115
116 <dokumente>
117 <dokument id="3" href="doc/abstract.doc"/>
118 <dokument id="4" href="doc/gutachten.doc"/>
119 <dokument id="5" href="doc/messreihe1.xls"/>
120 </dokumente>
121
122 <media>
123 <medium id="cover" type="img" href="Media/cover.jpg"/>
124 <medium id="logo" type="img" href="Media/HTWK_Logo.jpg"/>
125 <medium id="qr" type="img" href="Media/qr-doi.jpg"/>
126 <medium id="cc" type="img" href="Media/CC_Hinweis.jpg"/>
127
127 <medium id="b_7890123456789_part_002_chapter_003_fig_001" type="img" href="
"Media/b_7890123456789_fig_001.jpg"/>
128 <medium id="b_7890123456789_part_002_chapter_003_fig_002" type="img" href="

```

```
129         "Media/b_7890123456789_fig_002.jpg"/>
130         <medium id="b_9783966270021_eq_001" type="img" href="Media/b_9783966270021_eq_001.png"
131         />
132     </media>
133
134     <peer-review>
135         <gutachten id_ref="4"/>
136     </peer-review>
137
138     <forschungsdaten>
139         <messung id_ref="5"/>
140     </forschungsdaten>
141 </projekt>
```



# Anhang F

---

Generierung des Formulars zur DOI-Aktivierung

## F.1 Stylesheet

```

1  <?xml version="1.0" encoding="utf-8"?>
2
3  <xsl:stylesheet xmlns:xsl="http://www.w3.org/1999/XSL/Transform"
4  xmlns:xlink="http://www.w3.org/1999/xlink"
5  xmlns:fn="http://www.w3.org/2005/xpath-functions"
6  version="2.0"
7  xmlns:onix="http://www.editeur.org/onix/2.1/reference"
8  xmlns:crossref="http://www.crossref.org/schema/4.4.2">
9
10 <!-- Einbinden der Konfigurationsdatei: config_metadata.xml -->
11 <xsl:include href="config_metadata.xml"/>
12
13 <!-- Variablen nur für dieses Dokument -->
14 <xsl:variable name="jetzt" select="fn:current-dateTime()"/>
15
16 <xsl:output method="xml" indent="yes" encoding="utf-8"/>
17
18 <xsl:template match="/">
19
20     <doi_batch
21     version="4.4.2"
22     xmlns="http://www.crossref.org/schema/4.4.2"
23     xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
24     xsi:schemaLocation="http://www.crossref.org/schema/4.4.2
25     http://www.crossref.org/schema/deposit/crossref4.4.2.xsd"
26     xmlns:fn="http://www.w3.org/2005/xpath-functions"
27     xmlns:xlink="http://www.w3.org/1999/xlink"
28     xmlns:rel="http://www.crossref.org/relations.xsd">
29
30 <!-- Protokollausgabe Metadatenquellen -->
31 <xsl:message>
32   <xsl:text> Metadatenquellen included:
33 </xsl:text>
34   <xsl:for-each select="projekt/content/instanz">
35     <xsl:value-of select="$Pfad" />
36     <xsl:text></xsl:text>
37     <xsl:value-of select="@href" />
38     <xsl:text>
39   </xsl:for-each>
40 </xsl:message>
41
42   <!-- Head -->
43   <head>
44
45     <doi_batch_id>
46       <xsl:value-of select="format-dateTime($jetzt,'[D01][M01][Y0001]')"/>
47       <!-- zusätzliche Nummerierung -->
48       <xsl:text>-01</xsl:text>
49       <!-- Protokolleintrag -->
50       <xsl:message>
51         <xsl:text>Batch-ID </xsl:text>
52         <xsl:value-of select="format-dateTime($jetzt,'[D01][M01][Y0001]')"/>
53         <xsl:text>-01</xsl:text>
54       </xsl:message>
55     </doi_batch_id>
56
57     <timestamp>
58       <xsl:value-of select="format-dateTime($jetzt,'[D01][M01][Y0001][H10][m01]')"/>
59       <xsl:message>
60         <xsl:text>Zeitstempel </xsl:text>
61         <xsl:value-of select="format-dateTime($jetzt,'[D01].[M01].[Y0001],
62           [H10]:[m01]')"/>
63       </xsl:message>
64     </timestamp>
65
66     <depositor>
67       <depositor_name>
68         <xsl:variable name="Translate-Meta" select=

```

```

68     "projekt/metadaten/meta[@var='401']"/>
69     <xsl:for-each select="projekt/content/instanz">
70         <xsl:if test="@id=$Translate-Meta/@id_ref">
71             <xsl:variable name="Metadatenquelle">
72                 <xsl:value-of select="$Pfad"/>
73                 <xsl:value-of select="@href"/>
74             </xsl:variable>
75             <xsl:choose>
76                 <xsl:when test=
77                     "document($Metadatenquelle)/book/book-meta/publisher/publi
78                     sher-name[normalize-space(.)]">
79                     <xsl:value-of select=
80                         "document($Metadatenquelle)/book/book-meta/publisher/p
81                         ublisher-name"/>
82                     <xsl:message>
83                         <xsl:text>Hinweis: Als depositor wird der Verlag
84                         eingesetzt</xsl:text>
85                     </xsl:message>
86                 </xsl:when>
87                 <xsl:otherwise>
88                     <xsl:value-of select="$Warning"/>
89                     <xsl:message>
90                         <xsl:text>Fehlerhinweis zu depositor_name: Keine
91                         Informationen vorhanden</xsl:text>
92                     </xsl:message>
93                 </xsl:otherwise>
94             </xsl:choose>
95         </xsl:if>
96     </xsl:for-each>
97 </depositor_name>
98
99 <email_address>
100     <xsl:variable
101         name="Translate-Meta"
102         select="projekt/metadaten/meta[@var='405']" />
103     <xsl:for-each select="projekt/content/instanz">
104         <xsl:if test="@id=$Translate-Meta/@id_ref">
105             <xsl:variable name="Metadatenquelle">
106                 <xsl:value-of select="$Pfad" />
107                 <xsl:value-of select="@href" />
108             </xsl:variable>
109             <xsl:choose>
110                 <xsl:when test=
111                     "document($Metadatenquelle)/book/book-meta/publisher/publi
112                     sher-loc/email[normalize-space(.)]">
113                     <xsl:value-of select=
114                         "document($Metadatenquelle)/book/book-meta/publisher/p
115                         ublisher-loc/email"/>
116                 </xsl:when>
117                 <xsl:otherwise>
118                     <xsl:value-of select="$Warning"/>
119                     <xsl:message>
120                         <xsl:text>Fehlerhinweis zu depositor_eMail: Keine
121                         Informationen vorhanden</xsl:text>
122                     </xsl:message>
123                 </xsl:otherwise>
124             </xsl:choose>
125         </xsl:if>
126     </xsl:for-each>
127 </email_address>
128 </depositor>
129
130 <registrant>
131     <xsl:variable name="Translate-Meta"
132         select="projekt/metadaten/meta[@var='401']" />
133     <xsl:for-each select="projekt/content/instanz">
134         <xsl:if test="@id=$Translate-Meta/@id_ref">
135             <xsl:variable name="Metadatenquelle">
136                 <xsl:value-of select="$Pfad" />

```

```

125         <xsl:value-of select="@href" />
126     </xsl:variable>
127     <xsl:choose>
128         <xsl:when test=
129             "document($Metadatenquelle)/book/book-meta/publisher/publisher
130             -name[normalize-space(.)]">
131             <xsl:value-of select=
132                 "document($Metadatenquelle)/book/book-meta/publisher/publi
133                 sher-name" />
134             <xsl:message>
135                 <xsl:text>Hinweis: Als registrant wird der Verlag
136                 eingesetzt</xsl:text>
137             </xsl:message>
138         </xsl:when>
139         <xsl:otherwise>
140             <xsl:value-of select="$Warning"/>
141             <xsl:message>
142                 <xsl:text>Fehlerhinweis zu registrant: Keine
143                 Informationen vorhanden</xsl:text>
144             </xsl:message>
145         </xsl:otherwise>
146     </xsl:choose>
147 </xsl:if>
148 </xsl:for-each>
149 </registrant>
150 </head>
151 <body>
152     <peer_review>
153         <titles>
154             <title>
155                 <xsl:variable name="Translate-Meta" select=
156                     "projekt/metadaten/meta[@var='904']"/>
157                 <xsl:for-each select="projekt/content/instanz">
158                     <xsl:if test="@id=$Translate-Meta/@id_ref">
159                         <xsl:variable name="Metadatenquelle">
160                             <xsl:value-of select="$Pfad" />
161                             <xsl:value-of select="@href" />
162                         </xsl:variable>
163                         <xsl:choose>
164                             <xsl:when test=
165                                 "document($Metadatenquelle)/crossref:doi_batch/crossre
166                                 f:body/crossref:peer_review/crossref:titles/crossref:t
167                                 itle[normalize-space(.)]">
168                                 <xsl:value-of select=
169                                     "document($Metadatenquelle)/crossref:doi_batch/cro
170                                     ssref:body/crossref:peer_review/crossref:titles/cr
171                                     ossref:title" />
172                             </xsl:when>
173                             <xsl:otherwise>
174                                 <xsl:value-of select="$Warning"/>
175                                 <xsl:comment>Fehlerhinweis zu Peer-Review-Titel:
176                                 Keine Informationen vorhanden (Titel kann
177                                 zusammengesetzt werden aus "Review: " und
178                                 "Werks-Titel")</xsl:comment>
179                                 <!-- Protokollausgabe Fehlermeldung -->
180                                 <xsl:message>
181                                     <xsl:text>Fehlerhinweis zu Peer-Review-Titel:
182                                     Keine Informationen vorhanden (Titel kann
183                                     zusammengesetzt werden aus "Review: " und
184                                     "Werks-Titel")</xsl:text>
185                                 </xsl:message>
186                             </xsl:otherwise>
187                         </xsl:choose>
188                     </xsl:if>
189                 </xsl:for-each>
190             </title>
191         </titles>

```

```

175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221

```

```

<review_date>
  <xsl:variable name="Translate-Meta" select=
    "projekt/metadaten/meta[@var='903']" />
  <xsl:for-each select="projekt/content/instanz">
    <xsl:if test="@id=$Translate-Meta/@id_ref">
      <xsl:variable name="Metadatenquelle">
        <xsl:value-of select="$Pfad" />
        <xsl:value-of select="@href" />
      </xsl:variable>
      <month>
        <xsl:choose>
          <xsl:when test=
            "document($Metadatenquelle)/crossref:doi_batch/crossre
            f:body/crossref:peer_review/crossref:review_date/cross
            ref:month[normalize-space(.)]">
            <xsl:apply-templates select=
              "document($Metadatenquelle)/crossref:doi_batch/cro
              ssref:body/crossref:peer_review/crossref:review_da
              te/crossref:month" />
          </xsl:when>
          <xsl:otherwise>
            <xsl:value-of select="$Warning"/>
            <xsl:message>
              <xsl:text>Fehlerhinweis zu Review-Date
                (Month): keine Informationen vorhanden
              </xsl:text>
            </xsl:message>
          </xsl:otherwise>
        </xsl:choose>
      </month>
      <day>
        <xsl:choose>
          <xsl:when test=
            "document($Metadatenquelle)/crossref:doi_batch/crossre
            f:body/crossref:peer_review/crossref:review_date/cross
            ref:day[normalize-space(.)]">
            <xsl:apply-templates select=
              "document($Metadatenquelle)/crossref:doi_batch/cro
              ssref:body/crossref:peer_review/crossref:review_da
              te/crossref:day" />
          </xsl:when>
          <xsl:otherwise>
            <xsl:value-of select="$Warning"/>
            <xsl:message>
              <xsl:text>Fehlerhinweis zu Review-Date (Day):
                keine Informationen vorhanden</xsl:text>
            </xsl:message>
          </xsl:otherwise>
        </xsl:choose>
      </day>
      <year>
        <xsl:choose>
          <xsl:when test=
            "document($Metadatenquelle)/crossref:doi_batch/crossre
            f:body/crossref:peer_review/crossref:review_date/cross
            ref:year[normalize-space(.)]">
            <xsl:apply-templates select=
              "document($Metadatenquelle)/crossref:doi_batch/cro
              ssref:body/crossref:peer_review/crossref:review_da
              te/crossref:year" />
          </xsl:when>
          <xsl:otherwise>
            <xsl:value-of select="$Warning"/>
            <xsl:message>
              <xsl:text>Fehlerhinweis zu Review-Date

```

```

                (Year): keine Informationen vorhanden
            </xsl:text>
        </xsl:message>
    </xsl:otherwise>
</xsl:choose>
</year>
</xsl:if>
</xsl:for-each>
</review_date>
229
230 <rel:program xmlns:rel="http://www.crossref.org/relations.xsd">
231   <rel:related_item>
232     <rel:inter_work_relation
233       relationship-type="isReviewOf"
234       identifier-type="doi">
235       <xsl:variable name="Translate-Meta"
236         select="projekt/metadaten/meta[@var='108']" />
237       <xsl:for-each select="projekt/content/instanz">
238         <xsl:if test="@id=$Translate-Meta/@id_ref">
239           <xsl:variable name="Metadatenquelle">
240             <xsl:value-of select="$Pfad" />
241             <xsl:value-of select="@href" />
242           </xsl:variable>
243           <xsl:choose>
244             <xsl:when test=
245               "document($Metadatenquelle)/book/book-meta/book-id
246                [@book-id-type='doi'][normalize-space(.)]">
247               <xsl:value-of select=
248                 "document($Metadatenquelle)/book/book-meta/boo
249                 k-id[@book-id-type='doi']"/>
250             </xsl:when>
251             <xsl:otherwise>
252               <xsl:value-of select="$Warning"/>
253             <xsl:message>
254               <xsl:text>Fehlerhinweis zu Relations:
255                 Keine Informationen vorhanden</xsl:text>
256             </xsl:message>
257             </xsl:otherwise>
258           </xsl:choose>
259         </xsl:if>
260       </xsl:for-each>
261     </rel:inter_work_relation>
262   </rel:related_item>
263 </rel:program>
264
265 <doi_data>
266   <doi>
267     <xsl:variable name="Translate-Meta" select=
268       "projekt/metadaten/meta[@var='901']" />
269     <xsl:for-each select="projekt/content/instanz">
270       <xsl:if test="@id=$Translate-Meta/@id_ref">
271         <xsl:variable name="Metadatenquelle">
272           <xsl:value-of select="$Pfad" />
273           <xsl:value-of select="@href" />
274         </xsl:variable>
275         <xsl:choose>
276           <xsl:when test=
277             "document($Metadatenquelle)/crossref:doi_batch/crossre
278             f:body/crossref:peer_review/crossref:doi_data/crossref
279             :doi[normalize-space(.)]">
280             <xsl:value-of select=
281               "document($Metadatenquelle)/crossref:doi_batch/cro
282               sref:body/crossref:peer_review/crossref:doi_data/
283               crossref:doi"/>
284           </xsl:when>
285           <xsl:otherwise>
286             <xsl:value-of select="$Warning"/>
287             <xsl:message>
288               <xsl:text>Fehlerhinweis zu DOI_Data/DOI:

```

```

277         Keine Informationen vorhanden</xsl:text>
278     </xsl:message>
279     </xsl:otherwise>
280 </xsl:choose>
281 </xsl:if>
282 </xsl:for-each>
283 </doi>
284 <resource>
285     <xsl:variable name="Translate-Meta"
286       select="projekt/metadaten/meta[@var='902']" />
287     <xsl:for-each select="projekt/content/instanz">
288       <xsl:if test="@id=$Translate-Meta/@id_ref">
289         <xsl:variable name="Metadatenquelle">
290           <xsl:value-of select="$Pfad" />
291           <xsl:value-of select="@href" />
292         </xsl:variable>
293         <xsl:choose>
294           <xsl:when test=
295             "document($Metadatenquelle)/crossref:doi_batch/crossref:body/crossref:peer_review/crossref:doi_data/crossref:resource[normalize-space(.)]">
296             <xsl:value-of select=
297               "document($Metadatenquelle)/crossref:doi_batch/crossref:body/crossref:peer_review/crossref:doi_data/crossref:resource"/>
298           </xsl:when>
299           <xsl:otherwise>
300             <xsl:value-of select="$Warning"/>
301             <xsl:message>
302               <xsl:text>Fehlerhinweis zu DOI_Data/Resource:
303                 Keine Informationen vorhanden</xsl:text>
304             </xsl:message>
305           </xsl:otherwise>
306         </xsl:choose>
307       </xsl:if>
308     </xsl:for-each>
309 </resource>
310 </doi_data>
311 </peer_review>
312 </body>
313 </doi_batch>
314 </xsl:template>
315 </xsl:stylesheet>

```

## F.2 Formular zur DOI-Aktivierung

### **Vorlage: Webformular zu DOI Aktivierung auf Crossref**

Werktitel: Open-Access-Publikationsworkflow für akademische Bücher

Data Type Selection

**Select Data Type:** Book

Book-Information

**Select Book Deposit Type:** Monograph

**Book Type:** Monograph (Warning: Information could not be found)

**Title:** Open-Access-Publikationsworkflow für akademische Bücher: Ein Handbuch für Hochschulen und Universitäten

**Book DOI+:** 10.33968/9783966270175-00

**URL+:** <https://doi.org/10.33968/9783966270175-00>

**Publisher:** Open-Access-Hochschulverlag

**ISBN:** 978-3-96627-017-5

**Publication dates (YYYY/ MM/ DD)**

**(Type) Print:** 2020/ MM/ DD

**(Type) Online:** 2020/ MM/ DD

Chapter-Information

**Chapter 1**

**Title:** Einleitung

**Book DOI+:** 10.33968/9783966270175-01

**URL+:** <https://doi.org/10.33968/9783966270175-01>

**Chapter 2**

---

**Title:** Modell und Modellierung

**Book DOI+:** 10.33968/9783966270175-02

**URL+:** <https://doi.org/10.33968/9783966270175-02>

**Chapter 3**

---

**Title:** Open-Access-Publikationsworkflow für akademische Bücher

**Book DOI+:** 10.33968/9783966270175-03

**URL+:** <https://doi.org/10.33968/9783966270175-03>

**Chapter 4**

---

**Title:** Anwendung des Publikationsworkflow-Modells

**Book DOI+:** 10.33968/9783966270175-04

**URL+:** <https://doi.org/10.33968/9783966270175-04>

**Chapter 5**

---

**Title:** Schlussbetrachtung

**Book DOI+:** 10.33968/9783966270175-05

**URL+:** <https://doi.org/10.33968/9783966270175-05>



# Anhang G

---

Ergebnisinstanz für die Registrierung von Peer-Review-Daten

```

1 <?xml version="1.0" encoding="utf-8"?>
2 <doi_batch xmlns="http://www.crossref.org/schema/4.4.2"
3           xmlns:crossref="http://www.crossref.org/schema/4.4.2"
4           xmlns:fn="http://www.w3.org/2005/xpath-functions"
5           xmlns:onix="http://www.editeur.org/onix/2.1/reference"
6           xmlns:rel="http://www.crossref.org/relations.xsd"
7           xmlns:xlink="http://www.w3.org/1999/xlink"
8           xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
9           version="4.4.2"
10          xsi:schemaLocation="http://www.crossref.org/schema/4.4.2
11                               http://www.crossref.org/schema/deposit/crossref4.4.2.xsd">
12   <head>
13     <doi_batch_id>08092021-01</doi_batch_id>
14     <timestamp>080920211237</timestamp>
15     <depositor>
16       <depositor_name>Open-Access-Hochschulverlag</depositor_name>
17       <email_address>oa-strukt Komm@htwk-leipzig.de</email_address>
18     </depositor>
19     <registrant>Open-Access-Hochschulverlag</registrant>
20   </head>
21   <body>
22     <peer_review>
23       <titles>
24         <title>Peer-Review zu Open-Access-Publikationsworkflow für akademische Bücher
25         </title>
26       </titles>
27       <review_date>
28         <month>01</month>
29         <day>30</day>
30         <year>2020</year>
31       </review_date>
32       <rel:program>
33         <rel:related_item>
34           <rel:inter_work_relation relationship-type="isReviewOf" identifier-type="doi">
35             10.33968/9783966270175-00</rel:inter_work_relation>
36         </rel:related_item>
37       </rel:program>
38       <doi_data>
39         <doi>10.1234/5678901234567</doi>
40         <resource>https://doi.org/10.1234/5678901234567</resource>
41       </doi_data>
42     </peer_review>
43   </body>
44 </doi_batch>

```

