

## Fachbeitrag

Philipp Zumstein und Matti Stöhr

# Zur Nachnutzung von bibliographischen Katalog- und Normdaten für die persönliche Literaturverwaltung und Wissensorganisation

DOI 10.1515/abitech-2015-0037

**Zusammenfassung:** Literaturverwaltungsprogramme haben sich zu funktionsmächtigen, alltäglichen Begleitern des wissenschaftlichen Arbeitens sowie auch allgemein für die persönliche Wissensorganisation entwickelt. Eine seit jeher zentrale Funktion dieser Tools ist die Übernahme bibliographischer Daten aus Datenbanken, insbesondere aus elektronischen Bibliothekskatalogen (OPACs). Im Artikel werden die Möglichkeiten und die Qualität des Datenimports aus OPACs in die persönliche Literaturverwaltung diskutiert. Als ein wesentliches Desiderat wird dabei die Nachnutzung von Normdaten herausgearbeitet. Zudem wird eine Sensibilisierung für Anwendungen der persönlichen Literaturverwaltung bei der Katalogisierung oder Konzeption des Webkatalogs als notwendig erachtet.

**Schlüsselwörter:** Bibliographische Daten, Literaturverwaltung, Normdateien

## Reusing bibliographical metadata from library catalogues and data from authority files for personal reference management and knowledge organization

**Abstract:** In recent years reference management systems evolved into highly functional tools for the entire process of scholarly work. A common core feature is the import of bibliographical metadata from databases, specifically from the online library catalogue (OPAC). This article discusses the possibilities and quality of the import of bibliographical datasets from OPACs into the chosen personal reference management software. In this case we see the integration of information from authority files as a desideratum. Furthermore we want to stress sensitivity for reference management in the process of cataloguing as well as in the development of catalogues.

**Keywords:** Authority Files, Bibliographical Metadata, Reference Management

## 1 Grundzüge der Literaturverwaltung

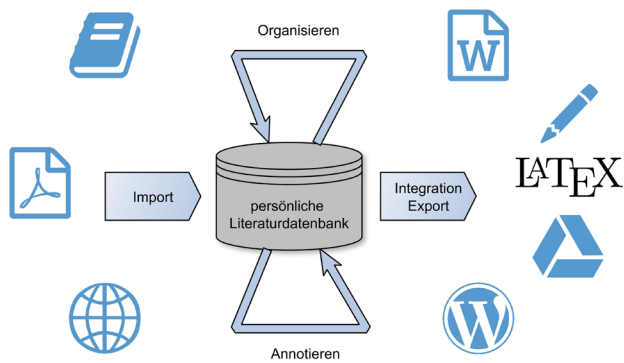
Die Tätigkeit der Literaturverwaltung ist ein zentrales Mittel zur Begleitung von literaturgestützten Erkenntnis- bzw. Wissensprozessen und stellt dabei eine essenzielle Handlung im wissenschaftlichen Arbeiten dar.<sup>1</sup> Im Zentrum steht dabei die persönliche Literaturdatenbank, welche neben den reinen bibliographischen Daten auch Wissens Elemente (z. B. Zitate, Zusammenfassungen, Informationen zur Strukturierung) und vor allem Volltexte enthalten kann. Die Arbeiten bei der Literaturverwaltung können dabei in drei Bereiche aufgeteilt werden (vgl. Abb. 1): strukturierter Import von unterschiedlichen Quellen, Organisieren und Annotieren der einzelnen Einträge sowie die Integration der Literaturdatenbank in den Schreibprozess für zitierstilgerecht formatierte Literaturnachweise respektive Literaturlisten.

Wie unter anderem Matthias Nagelschmidt und Thomas Stöber nachvollziehbar feststellten, ist die Literaturverwaltung nicht isoliert zu betrachten: Sie ist eine selbstverständliche Teilaufgabe im „Kontext der Informationszirkulation“<sup>2</sup> in engem Zusammenhang mit der Literaturrecherche, -beschaffung, -rezeption sowie anderen Teilschritten der Literatuarbeit.<sup>3</sup>

1 Vgl. Stöhr, Matti. *Bibliothekarische Dienstleistungen für Literaturverwaltung: Eine vergleichende Analyse des Angebots wissenschaftlicher Bibliotheken unter besonderer Berücksichtigung der Nutzerperspektive*. Magisterarbeit, Institut für Bibliotheks- und Informationswissenschaft der Humboldt-Universität zu Berlin. Berlin: Berliner Handreichungen zur Bibliotheks- und Informationswissenschaft 283, 2010. 14–16. urn:nbn:de:kobv:11-100178393.

2 Nagelschmidt, Matthias. „Literaturverwaltungsprogramme: Das funktionale Minimum.“ *ABI-Technik* 30, Nr. 2 (2010): 94–99. S. 96. doi:10.1515/ABITECH.2010.30.2.94.

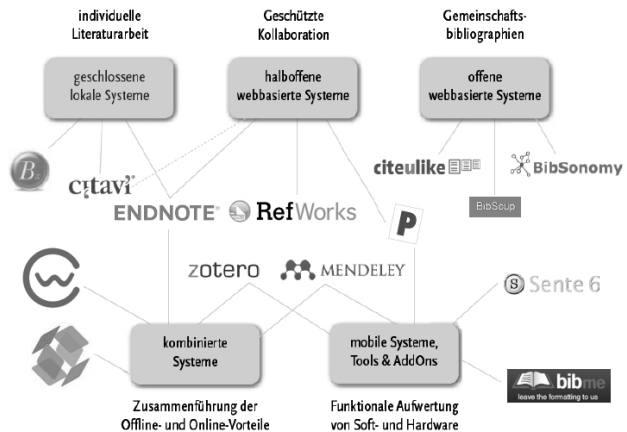
3 Vgl. Stöber, Thomas. „Informationskompetenz und Literaturverwaltung.“ In *Handbuch Informationskompetenz*. Hrsg. von Wilfried Sühli-Strohmenger, 281–289. Berlin u. a.: De Gruyter, 2012. S. 286. doi:10.1515/9783110255188.281.



**Abb. 1:** Grundfunktionalitäten von Literaturverwaltungsprogrammen lassen sich in drei Bereiche unterteilen: Import, Organisieren/Annotieren, Integration/Export

Es existieren zur Unterstützung und Effektivierung der Literaturverwaltung eine Reihe unterschiedlichster Softwarelösungen. Dazu gehören beispielsweise Citavi, EndNote, Mendeley, RefWorks und Zotero. Literaturverwaltungsprogramme wurden in den letzten Jahren als Instrumente weiterentwickelt, die den gesamten wissenschaftlichen Arbeits- bzw. Forschungsprozess von der Ideenfindung bis hin zur Manuskripterstellung und Publikation unterstützen. Folglich sind solche Programme treffender als „Research Management“-Tools beschrieben, da sie weit mehr bieten als eine „klassische Literaturverwaltung“. Dies geht mit dem Angebot differenzierter, mächtiger Funktionsmöglichkeiten einher, die verschiedenste Anwendungsszenarien erlauben – typologisch dargestellt in Abb. 2. Flexible Importmechanismen zur Übernahme und systematischen (kollaborativen) Weiterverarbeitung bibliographischer Daten sind inzwischen bei den meisten Marktvertretern Standard. State of the Art ist die funktionale Kombination von individueller und kollaborativer Literaturarbeit über Synchronisationsmechanismen zwischen Desktop- und Webkomponenten („kombinierte Systeme“), entsprechenden Applikationen für mobile Endgeräte sowie die Einbindung von Literaturverwaltungsfunktionen in verschiedenste Webumgebungen, wie über Plugins in Content Management Systeme, etwa im Falle von Mendeley und Zotero für WordPress.<sup>4</sup>

In den Wissenschaftsdisziplinen bilden sich, teils bereits seit den 1990er Jahren, elektronische Publikationskulturen heraus. Gleichzeitig wächst die Verfügbarkeit und Nachnutzbarkeit bibliographischer (Meta-)Daten aus den unterschiedlichsten (Online-)Datenquellen, welche für die persönliche



**Abb. 2:** Funktional-anwendungsbezogene Typologie von Literaturverwaltungssoftware<sup>5</sup>

Literaturverwaltung und Wissensorganisation genutzt werden können. Eine der wichtigsten Quellen für bibliographische Daten von selbstständigen Werken und deren Übernahme sind elektronische Bibliothekskataloge, die OPACs. Inwiefern diese tatsächlich eine attraktive Datenquelle darstellen und wie es um die Datenqualität in Bezug auf die Nachnutzbarkeit innerhalb des wissenschaftlichen Schreibens steht, soll im Folgenden weiter ausgeleuchtet werden.

## 2 Datenformate: In welcher Sprache kommuniziert der Bibliothekskatalog mit Literaturverwaltungsprogrammen?

Bibliotheken katalogisieren mit einem erheblichen Aufwand ihre Bestände, folgen dabei gewissenhaft den entsprechenden Regelwerken und pflegen gemeinsam bibliographische Titeldaten wie auch Normdaten. Die Aufbereitung dieser Daten für den Webkatalog läuft meist über mehrere Schritte und auch durch verschiedene Systeme. Häufig bietet der Katalog wiederum Exportmöglichkeiten in weitere (Web-)Formate. Daneben existieren auch spezielle Austauschformate und Schnittstellen. Somit gibt es zumindest theoretisch eine Vielzahl von unterschiedlichen Datenformaten für bibliographische Daten zur Nutzung in Literaturverwaltungstools – siehe Abb. 3.

<sup>4</sup> Vgl. O. A. Zotpress. *WordPress.org*. <https://de.wordpress.org/plugins/zotpress/> (Zugegriffen: 7. Oktober 2015) und O. A. Mendeley Plugin. *WordPress.org*. <https://de.wordpress.org/plugins/mendeleyplugin/> (Zugegriffen: 7. Oktober 2015).

<sup>5</sup> Eine erste Version des Typologieschemas mit ausführlicher Beschreibung ist unlängst erschienen: Stöhr, Matti. 2012. Steckbrief: Literaturverwaltungsprogramme – Eine Typologie. *cms-journal*, Nr. 35: 46–47. urn:nbn:de:kobv:11-100200911. S. 47.

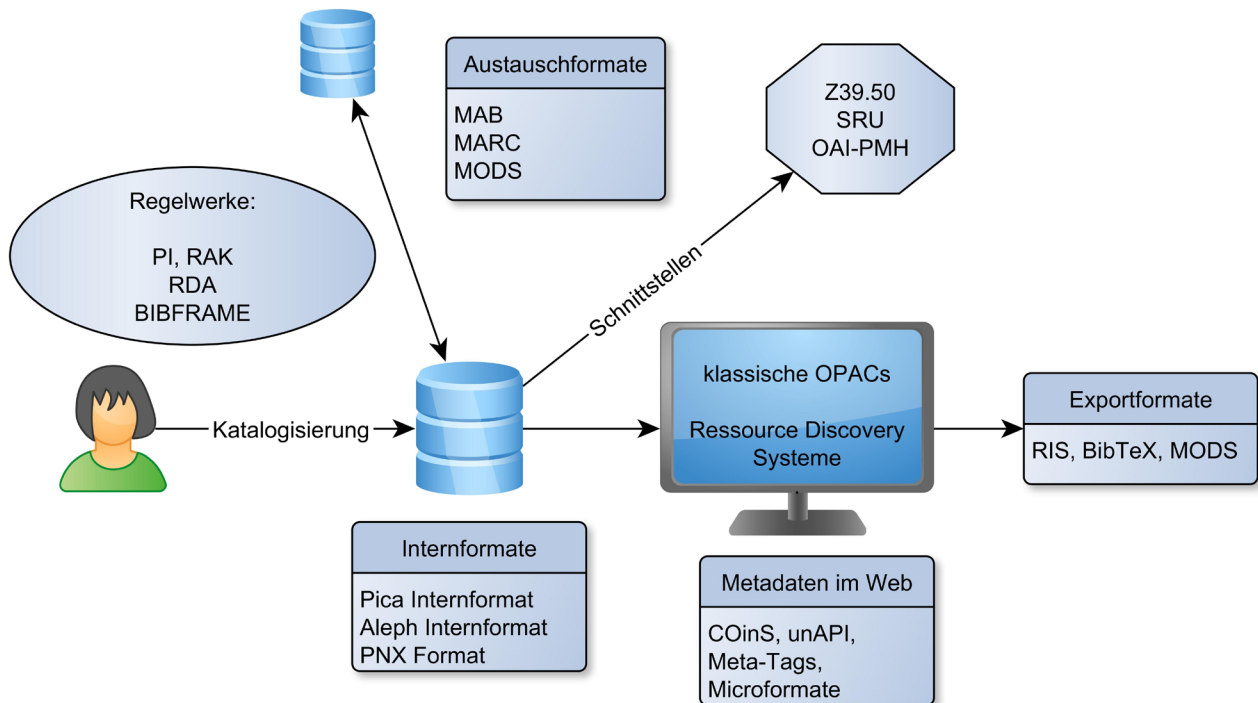


Abb. 3: Datenformate im Umfeld eines Bibliothekskataloges

## 2.1 Bibliothekarische Formate und Regelwerke

Das neu in die Erschließungspraxis eingeführte Regelwerk Resource Description and Access (RDA) will von Anfang an auf die Benutzer ausgerichtet sein, welche dann insbesondere durch OPACs im klassischen Sinne oder auch Ressource Discovery Systeme mit den Bibliotheksdaten in Berührung kommen: „Daten, die unter Verwendung von RDA zur Beschreibung einer Ressource erzeugt werden, sollen die Benutzer bei den folgenden Anforderungen unterstützen:

- Finden [...],
- Identifizieren [...],
- Auswählen [...] und
- Zugang erhalten [...].“<sup>6</sup>

<sup>6</sup> RDA 0.0 was in Übereinstimmung mit den Anforderungen an bibliographische Datensätze im Modell der Functional Requirements for Bibliographic Records (FRBR) steht. Vgl. Oehlschläger, Susanne, International Federation of Library Associations and Institutions und Deutsche Nationalbibliothek, Hrsg. *Funktionale Anforderungen an bibliografische Datensätze: Abschlussbericht der IFLA Study Group on the Functional Requirements for Bibliographic Records*. Geänderte und korrigierte Fassung, Stand: Februar 2009. Leipzig, Frankfurt a.M., Berlin: Dt. Nationalbibliothek, 2009. S. 8. urn:nbn:de:101-2009022600.

Diese vier Punkte umschreiben treffend verschiedene Aufgaben, welche während einer Literaturrecherche auftreten: Thematische Suche, Known-Item-Search, zwischen verschiedenen Auflagen bzw. Medienformaten auswählen, im Bibliothekskatalog den Standort eines Buches nachschlagen bzw. den Ausleihvorgang anstoßen, Zugriff zu lizenzierten E-Medien bekommen.

Sobald man sich bei einer Literaturrecherche entschieden hat, dass eine bestimmte Quelle für die wissenschaftliche Arbeit interessant ist, wird man diese oft auch in das genutzte Literaturverwaltungsprogramm speichern wollen. Dazu benötigt man mindestens eine Möglichkeit, die bibliographischen Daten zu übernehmen. Eine Übernahme bibliographischer Daten durch die Benutzer ist aber kein erklärtes Ziel von RDA. Dies heißt nicht, dass dies nicht möglich ist, sondern nur dass das Katalogisierungsregelwerk nicht auf dieses Anforderung hin ausgerichtet ist. Inwiefern sich dies auswirkt und wie die Benutzer dennoch Daten aus dem Katalog in die eigene Literaturverwaltung übernehmen können, werden wir im Folgenden weiter untersuchen.

Ebenso ist RDA nicht ausdrücklich darauf ausgerichtet, Bibliotheksdaten als Open Data für die Nachnutzung auch außerhalb der Bibliothekswelt zu veröffentlichen. Dies steht im spannungsvollen Gegensatz zum Potential von BIBFRAME: „With BIBFRAME, the library community has an opportunity to make its controlled and well-crafted bibliographic data accessible to a global audience. Wider

accessibility of a library's bibliographic data makes the library's resources and holdings known and available to 'outsiders'.“<sup>7</sup>

Es gibt aber einen regen Austausch von bibliographischen Metadaten der Bibliotheken untereinander. Insbesondere durch die Internationalisierung des Regelwerkes soll es noch einfacher sein, bereits vorhandene Daten zu übernehmen – mit dem Ziel, Effizienzsteigerungen beim Katalogisieren zu erreichen. Die Austauschformate dafür, wie etwa MAB2, MODS, MARC,<sup>8</sup> können sehr detailliert die bibliographischen Daten speichern, was für einen möglichst verlustfreien Austausch von Daten zwischen Bibliotheken auch notwendig ist.

Diese Austauschformate werden aber nur in der Bibliothekswelt verwendet. Beispielsweise wird MARC von Bibliotheken, Bibliotheksverbünden und Verlagen zum Austausch verwendet, aber weitere Datenlieferanten für MARC sind eher rar. Daher ist es nicht weiter verwunderlich, dass Literaturverwaltungsprogramme diese bibliothekarischen Formate meist nicht importieren können. Aber selbst wenn die Programme die Formate verstehen, gibt es ein weiteres Problem, dass die Benutzer diese Formate aus dem Katalog oft gar nicht angeboten bekommen.

Häufig bieten Bibliotheken bzw. die Verbünde Schnittstellen wie etwa Z39.50, SRU, OAI-PMH für den Datenaustausch an. Die übertragenen Daten haben dabei aber meist die gleichen bibliothekarischen Austauschformate. Einige Literaturverwaltungsprogramme, etwa Citavi, bieten daher die Möglichkeit, aus dem Programm heraus die Z39.50-Schnittstelle abzufragen. Dies ist in den meisten Fällen jedoch nur dann sinnvoll, wenn man weiß, was man sucht: einen bestimmten Titel oder die Publikationen einer bestimmten Autorin. Die für eine Recherche verfügbaren Kategorien sowie die Darstellung der Suchergebnisse im Literaturverwaltungsprogramm sind in der Regel ein Minimalangebot. Dagegen werden in einem Resource Discovery System oder auch schon in einem OPAC natürlich noch sehr viele weitere Recherchefunktionalitäten bereitgestellt. Mögliche Fehler beim Import der Metadaten in das Literaturverwaltungsprogramm sind jedoch meist erst im Nachhinein manuell korrigierbar.

<sup>7</sup> Frank, Paul. *BIBFRAME: Why? What? Who?* Library of Congress: 2014. S. 2. <http://www.loc.gov/aba/pcc/bibframe/BIBFRAME%20paper%2020140501.docx> (6. Oktober 2015).

<sup>8</sup> MARC ist nur ein Container und kann unterschiedlich gefüllt werden (MARC21, UNIMARC, BELMARC) bzw. unterschiedlichen Katalogisierungsregeln (AACR2, RAK, RDA) folgen.

## 2.2 Katalogdaten im Web

Die Darstellung der Katalogdaten im Web nennen wir hier der Einfachheit halber Webkataloge und fassen darunter folgende Begriffe zusammen: klassische OPACs, Resource Discovery Systeme, Next Generation Systems, OPAC 2.0. Insbesondere bei den neueren Webkatalogen werden die Daten in einem ersten Schritt normalisiert, indexiert und daraus ein Such-Index (meist auf Grundlage von Lucene/Solr) aufgebaut. Dadurch können auch wiederum neue Formate, beispielsweise das Primo Normalized XML (PNX) Format, entstehen.

Die Ziele des Webkataloges sind prinzipiell identisch mit den oben genannten vier Zielen nach dem FRBR-Modell: Finden, Identifizieren, Auswählen und Zugang erhalten. Es gibt aber zumindest in den IFLA-Richtlinien für die OPAC-Anzeige von 2010 die Empfehlung 5.12: „Geben Sie dem Benutzer die Möglichkeit, eine vollständig kodierte Datensatzanzeige anzufordern.“<sup>9</sup>

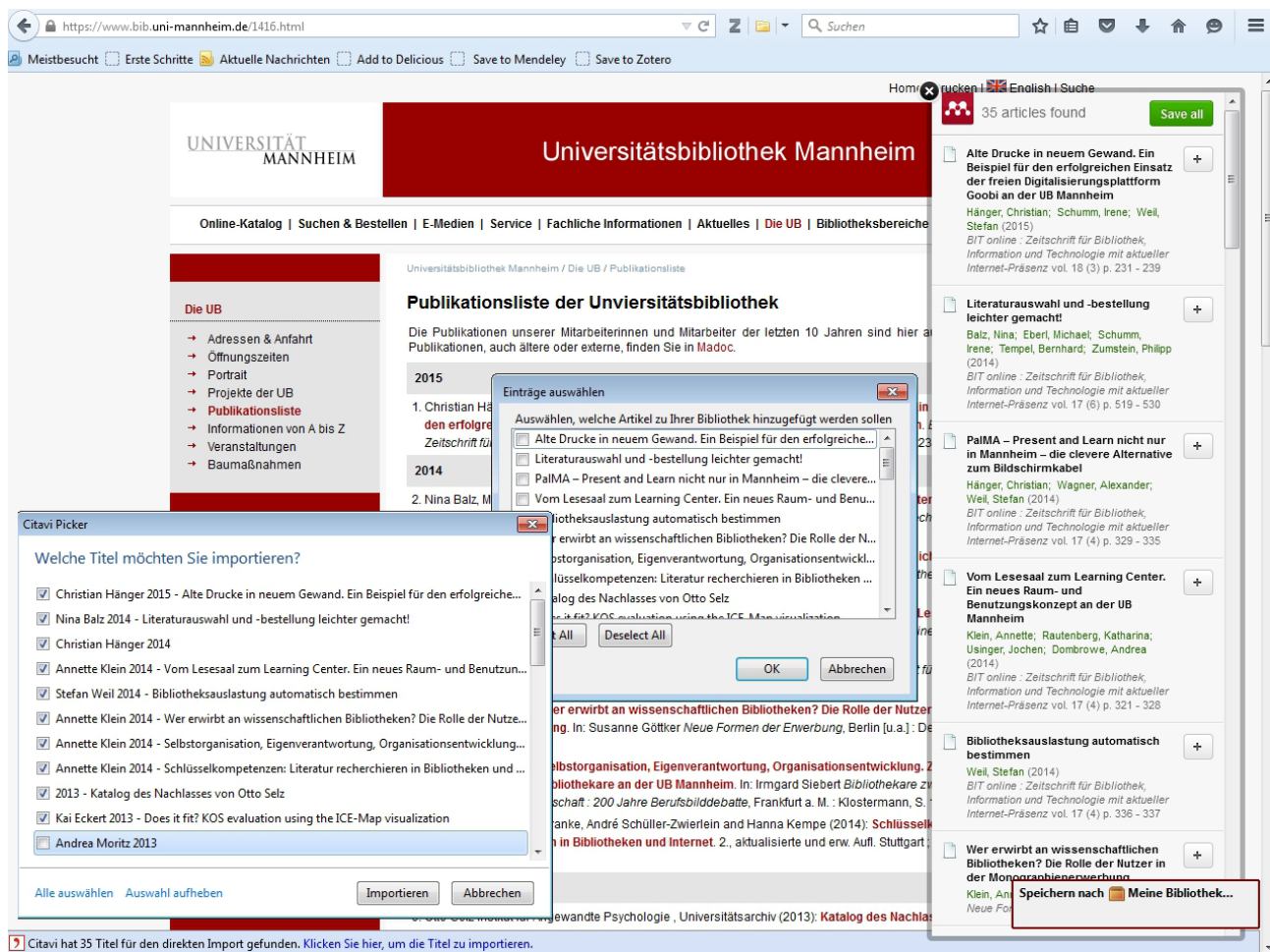
In den Richtlinien-Beispielen wird dies jeweils mit einem Link auf den komplementären MARC-Eintrag realisiert. Dies wäre zumindest eine Möglichkeit für den Benutzer, die bibliothekarisch-bibliographischen Daten während einer Katalogrecherche zu speichern. Inwiefern der Nutzer mit dem MARC-Datensatz oder ob beispielsweise das gewählte Literaturverwaltungsprogramm damit etwas anfangen kann, bleibt offen. Ebenfalls fehlt dabei die Vorstellung, dass man mit der Software auch über andere Wege kommunizieren könnte und Daten austauschen kann, so dass nicht alles auf reine „Datensatzanzeige“ hinausläuft.

Die meisten Webkataloge zeigen die Daten für einen menschlichen Benutzer möglichst ansprechend und übersichtlich an, so dass der Mensch alle oben aufgeführten Aufgaben möglichst einfach erledigen kann. Beispielsweise wird bei Webkatalogen mit neuen Features wie etwa mit dem einfachen Kopieren von fertig formatierten Zitationen in einer Handvoll von Zitierstilen geworben.<sup>10</sup> Dies zeigt aber, dass Webkataloge auf die grafisch ansprechende Präsentation für einen Menschen ausgerichtet sind und nicht auf das Übertragen der dahinterliegenden

<sup>9</sup> Vgl. Task Force on Guidelines for OPAC Displays, Hrsg. *Richtlinien für die OPAC-Anzeige*. Übers. von Friederike Schimmelpfennig. IFLA Series on Bibliographic Control 40. Berlin: De Gruyter Saur, 2010. S. 23. doi:10.1515/9783110232493.

<sup>10</sup> Die fertigen Formatierungen sind jedoch häufig fehlerhaft und können auch teilweise ohne Kontext gar nicht korrekt formatiert werden, z. B. wenn im Zitierstil bei direkten Wiederholungen von Literaturstellen „ibid“ verwendet wird oder bei Wiederholungen von Autoren „idem“ o. ä., vgl. o. A. „How can subsequent occurrences of the same author be replaced by a fixed term/symbol?“ [Zotero Documentation]. <https://www.zotero.org/support/kb/idem> (6. Oktober 2015).





**Abb. 4:** Die Datenübernahme mittels Zotero Ein-Klick-Button, Citavi Picker, Mendeley Bookmarklet werden in dieser Webseite durch die eingebauten COinS-Daten ermöglicht

bibliographischen Daten. Eine übersichtliche Kombination sowohl menschen- wie auch maschinenlesbarer Übernahmeoptionen bibliographischer Daten aus Webkatalogen ist durchaus möglich. Dies beweisen z. B. Beluga<sup>11</sup> (Katalog der Hamburger Bibliotheken) und der Katalog.plus!<sup>12</sup> der Universitätsbibliothek Bielefeld, auch wenn es im Detail Verbesserungspotential gibt, etwa in der fehlenden Berücksichtigung von Herausgebern.

In Webkatalogen können die Daten über verschiedene Webformate maschinenlesbar und -auffindbar gemacht werden. Durch Verwendung von Webformaten sind die Daten auch für direkte maschinelle Auswertung abgreifbar, wie es beispielsweise bei Suchmaschinen oder Literaturverwaltungsprogrammen genutzt wird. Dabei müssen diese

Webformate zwei Funktionen unterstützen: Erstens müssen sie standardisiert in der Webseite hinterlegt sein. Zweitens müssen die Daten maschinenlesbar und auffindbar sein. Sind beide Bedingungen erfüllt, können Literaturverwaltungsprogramme während der Online-Recherche eine einfache Übernahmemöglichkeit für die Daten anbieten. Beispiele hierfür sind das Zotero Ein-Klick-Symbol, der Citavi Picker und das Mendeley Bookmarklet – vgl. Abb. 4.

In der Praxis wird häufig das Webformat ContextObjects in Spans<sup>13</sup> (COinS) verwendet, um eine Anbindung an Literaturverwaltungsprogramme zu erreichen. Dabei wird im HTML-Dokument ein unsichtbares SPAN-Element mit der Klasse „Z3988“ verwendet, welches die

<sup>11</sup> Beispiel: <https://beluga.sub.uni-hamburg.de/vufind/Record/804067864> (7. Oktober 2015).

<sup>12</sup> Beispiel: <https://katalogplus.ub.uni-bielefeld.de/title/2426306> (7. Oktober 2015).

<sup>13</sup> Gemeinschaftliche Spezifikation von COinS: Hellman, Eric. *OpenURL COinS: A Convention to Embed Bibliographic Metadata in HTML. Stable version 1.0*. OCLC New Jersey, 16. Juni 2009. <http://ocoins.info/> (6. Oktober 2015).

bibliographischen Informationen im Attribut TITLE speichert. Beispiel:

```
<span class='Z3988' title='url_ver=Z39.88-2004
&ctx_ver=Z39.88-2004
&rft_id=urn%3Aisbn%3A978-3-11-025473-0
&rft_val_fmt=info%3Aofi%2Ffmt%3Akev%3Amtx%3Abook
&rft.genre=book
&rft.btitle=Handbuch%20Informationskompetenz
&rft.place=Berlin
&rft.publisher=De%20Gruyter%20Saur
&rft.date=2012
&rft.isbn=978-3-11-025473-0
&rft.language=ger'></span>
```

Durch diese speziell gewählte Klassenbezeichnung ist es für Literaturverwaltungsprogramme leicht möglich, automatisiert die bibliographischen Daten innerhalb der HTML-Datei zu finden und in die programmeigene Feldstruktur zu transferieren. Das COinS-Format basiert auf dem OpenURL-Standard zur Angabe von Metadaten in einer URL, was in Linkresolving Systemen wie z. B. SFX verwendet wird. Das Ziel von OpenURL ist es, eine konkrete Resource, welche anhand von bibliographischen Metadaten beschrieben wird, auf den von einer Institution lizenzierten Zugang (direkt beim Verlag oder über einen Aggregator) abzubilden. Für diesen Zweck müssen die Metadaten nicht unbedingt vollständig sein. Beispielsweise gibt es im COinS-Format nur die Möglichkeiten Autoren zu speichern, nicht aber Herausgeber, Übersetzer oder weitere Funktionen. Ebenfalls kann bei Büchern die Bandnummer und Nummer der Buchreihe nicht gespeichert werden.<sup>14</sup>

Anstatt die Daten in ein spezielles Webformat umzuwandeln und in die Webseite zu integrieren, ist es auch möglich die Daten in den üblichen Formaten (z. B. MARC, MODS) bereitzustellen und maschinenverständlich und zielgenau in der Webseite darauf hinzuweisen. Als Beispiel kann man hier unAPI anführen, wobei das Projekt inzwischen verwaist zu sein scheint.<sup>15</sup> Eine andere Möglichkeit

ist es, den LINK-Tag zu benutzen, um auf die Metadaten hinzuweisen. Durch die genaue Angabe des TYPE Attributes kann diese Information einfach gefunden werden. Beispielsweise bietet die Katalogsoftware Blacklight<sup>16</sup> diese Funktionalität an, was dann etwa so aussieht:

```
<link rel="alternate" title="marcxml"
type="application/marcxml+xml"
href="http://demo.projectblacklight.org/catalog/
2009001699.marcxml"
/>
```

Die META Tags in den Kopfdaten einer HTML-Datei können ebenfalls genutzt werden, um bibliographische Metadaten zu speichern, beispielsweise mit dem DublinCore Standard. Zusätzlich existieren bibliographische Mikroformate,<sup>17</sup> HTML5 Microdata,<sup>18</sup> RDFa, JSON-LD oder RDF-Daten, welche über Content Negotiation zugänglich sein können. Diese innovativen Datenformate und Konzepte aus dem Semantic Web haben aber bisher nur begrenzten Einfluss auf die Welt der Literaturverwaltungsprogramme.

## 2.3 Exportformate: RIS, BibTeX

Die beiden prominentesten, bei Literaturverwaltungsprogrammen etablierten bibliographischen Metadatenformate sind RIS und BibTeX. Beide Formate besitzen eine Liste von möglichen Publikationstypen und speichern die Informationen als Paare von Feldnamen und Werten ab. Einige Werte müssen in der richtigen Form eingegeben werden, beispielsweise trennt BibTeX verschiedene Autoren durch das Schlüsselwort „and“ und die einzelnen Autoren müssen in der Form Nachname(n), Vorname(n) eingetragen werden.

Das Akronym RIS geht auf den ursprünglichen Firmennamen Research Information Systems Inc. zurück, eine Firma die den Reference Manager (RefMan) und ProCite herausgegeben hat.<sup>19</sup> Das Unternehmen ist jetzt Teil von Thomson Reuters, die auch EndNote vertreiben.

<sup>14</sup> Hellman, Eric. *Brief guide to Implementing OpenURL 1.0 Context-Object for Books*. OCLC New Jersey, 16. Juni 2009. <http://ocoinfos.info/cobgbook.html> (6. Oktober 2015).

<sup>15</sup> Die Webseite <http://unapi.info/> gibt es nicht mehr und die alten Versionen sind nur noch über die WayBackmaschine zugänglich: Chudnov, Daniel. *unAPI.info – for all your unAPI needs*. 2006. <https://web.archive.org/web/2015011204516/http://unapi.info/> (6. Oktober 2015). Eine Implementation des Services findet man im Wiki des Gemeinsamen Bibliotheksverbunds GBV: o. A. *unAPI – Verbund-Wiki GBV*. 12. Januar 2015. <https://www.gbv.de/wikis/cls/index.php?title=UnAPI&oldid=17706> (6. Oktober 2015).

<sup>16</sup> O. A. *Blacklight*. <http://projectblacklight.org/> (6. Oktober 2015).

<sup>17</sup> O. A. *Citation microformat efforts Microformats Wiki*. <http://microformats.org/wiki/index.php?title=citation&oldid=49131> (6. Oktober 2015).

<sup>18</sup> O. A. *5 Microdata – HTML 5*. <http://www.w3.org/TR/2009/WD-html5-20090825/microdata.html> (6. Oktober 2015).

<sup>19</sup> O. A. „RIS (file format).“ *Wikipedia, the free encyclopedia*. 13. Januar 2015. [https://en.wikipedia.org/w/index.php?title=RIS\\_\(file\\_format\)&oldid=642289508](https://en.wikipedia.org/w/index.php?title=RIS_(file_format)&oldid=642289508) (6. Oktober 2015).

Die Entwicklung des Reference Manager wurde unlängst zugunsten von EndNote eingestellt, und der Hersteller empfiehlt offensiv den Umstieg.<sup>20</sup> Es gibt einen Standard zum RIS Format,<sup>21</sup> aber dieser lässt einen gewissen Interpretationsspielraum beim Gebrauch des Formates, und es findet auch keine aktive Weiterentwicklung des Formates statt. Daher werden auch neue Felder von den Anbietern benutzt, ohne dies mit anderen Entwicklungen abzustimmen. Eine RIS-Datei des einen Anbieters muss daher nicht kompatibel mit der Interpretation eines anderen sein. Dadurch kommt es zu einem Wildwuchs an Ausprägungen des RIS-Formates und Anwendungsprofilen, die dann von Literaturverwaltungsprogrammen unterschiedlich gut importiert werden können.<sup>22</sup>

Hier ist es natürlich auch für Webkataloge schwierig, einen optimalen RIS-Export bereitzustellen. Wir halten es aber nicht für zielführend, diesen Export nur auf ein Programm zu kanalisieren, z. B. auf das Programm, welches von der Bibliothek momentan unterstützt wird. Ebenfalls sieht man in Katalogen teilweise anstatt „RIS-Export“ die Bezeichnung des Exports in ein spezielles Programm. Dies mag zwar für die lokalen Benutzer temporär einfacher zu verstehen sein, aber ist dann nicht konform mit anderen Informationsdiensten. Bei einem Wechsel des Literaturverwaltungsprogramms oder auch für andere Nutzergruppen ist die Bezeichnung verwirrend. Es ist im Sinne der Benutzerinnen und verschiedenen Anwendungsszenarien, im Exportdialog die Datenformate (etwa RIS und BibTeX) explizit zu benennen. Konkrete Produktnamen können als ergänzende Beispiele fungieren.

BibTeX ist als bibliographisches Format für das Textsatzsystem LaTeX eingeführt worden. Die Bezeichnung BibTeX wird sowohl für das Metadatenformat wie auch für das gleichnamige Programm verwendet. Das Metadatenformat BibTeX besteht ebenfalls aus verschiedenen Feldern im Kontext von verschiedenen Publikationstypen. Jegliche weitere Information kann in zusätzliche Felder geschrieben werden, wie es im Handbuch von 1988 heißt: „BibTeX ignores any field that is not required or optional, so you can include any fields you want in a bib file entry.“<sup>23</sup>

Diese zusätzlichen Felder ermöglichen einen sehr flexiblen und individualisierbaren Einsatz des BibTeX-Formates, aber erschweren gleichzeitig die Anstrengungen, sich auf einen gemeinsamen Standard zu einigen. Seit der Veröffentlichung Mitte der 1980er Jahre kam es zu verschiedenen Neuimplementationen, z. B. biblatex, biber oder bibtex8. Einerseits unterscheiden sich diese Neuimplementationen bzgl. der Zeichenkodierung bei der Eingabe oder auch Sortierreihenfolge, da BibTeX nur einfache ASCII-Zeichen unterstützt und somit alle Sonderzeichen speziell kodiert werden müssen. Andererseits werden aber auch andere Felder benutzt, um die Informationen zu speichern. Die dazugehörigen bibliographischen Metadaten sind daher oft nicht kompatibel.

Zwar sind RIS oder BibTeX die am häufigsten verwendeten Datenformate, um bibliographische Daten für die Literaturverwaltung auszutauschen. Aber wie wir gesehen haben, sind diese Datenformate häufig unvollständig und es fehlt eine standardisierte Verwendung in der Praxis. Daher gehen bei solchen Importen in ein Literaturverwaltungsprogramm häufig einige Details verloren, was sich in der Folge als unbefriedigend darstellt.

### 3 Datenqualität: Wie gut eignen sich Katalogdaten für die Literaturverwaltung?

In diesem Abschnitt wollen wir uns auf die Datenqualität der Katalogdaten in Bezug auf die Literaturverwaltung konzentrieren. Dazu fokussieren wir uns zum ersten auf die Katalogisierungsregeln und Gepflogenheiten beim Katalogisieren in Bibliotheken – unabhängig vom konkreten Datenformat. Die Bibliotheken legen fest, welche Informationen überhaupt erfasst werden und auf welche Weise dies geschieht. Zum Zweiten sind die Anforderungen der Literaturverwaltung in Hinblick auf das wissenschaftliche Schreiben maßgebend. Das Literaturverzeichnis und die Zitationen folgen in einer wissenschaftlichen Arbeit einem bestimmten Zitierstil, welcher beispielsweise die Reihenfolge und Formatierung der einzelnen bibliographischen Elemente einer Fundstelle festlegt. Es gibt eine Unmenge von verschiedenen Zitierstilen mit oft sehr vielen Detailregelungen.<sup>24</sup> Die im Literaturverwaltungsprogramm verarbeiteten Daten sollten die benötigten Informationen für möglichst alle Zitierstile enthalten.

<sup>20</sup> O. A. *EndNote | Reference Manager | Thomson Reuters*. <http://refman.com/switch> (6. Oktober 2015).

<sup>21</sup> Als aufbereitete Form vgl. Vinkevicius, Aurimas. *RIS Tag Map aurimasv/translators Wiki GitHub*. 18. April 2012. <https://github.com/aurimasv/translators/wiki/RIS-Tag-Map> (6. Oktober 2015).

<sup>22</sup> Beispielsweise kann man die unterschiedlichen Routinen beim Importieren von einer RIS-Datei im Quellcode von Zotero sehen: O. A. *translators/RIS.js at master zotero/translators GitHub*. 22. August 2015. <https://github.com/zotero/translators/blob/master/RIS.js> (6. Oktober 2015).

<sup>23</sup> Patashnik, Oren. *BIBTeXing*. 8. Februar 1988. S. 8. <http://bibtexml.sourceforge.net/btxdoc.pdf> (6. Oktober 2015).

<sup>24</sup> Beispielsweise listet momentan das CSL-Verzeichnis beinahe 8 000 Zitierstile, von denen ca. 1 200 als grundsätzlich verschieden gelten: O. A. *Zotero Style Repository*. <https://www.zotero.org/styles/> (6. Oktober 2015).

### 3.1 Publikationstypen und Detailierungstiefe

Die verschiedenen Publikationstypen werden in den Bibliotheksdaten sehr fein unterschieden. In RDA wird dies zusätzlich zu den Informationen aus dem MARC-Leader in den Feldern für Inhaltstyp, Medientyp und Datenträgertyp ausgeführt, was eine sehr genaue Beschreibung des Publikationstyps und dessen Erscheinungsform erlaubt. Diese Komplexität wird meist nur teilweise für die Literaturverwaltung benötigt.<sup>25</sup> Die Katalogisierung in Bibliotheken im engeren Sinne beschränkt sich hierbei (bis auf wenige Ausnahmen, etwa in Spezialbibliotheken) auf selbstständige Werke und erfasst in der Regel bis heute keine einzelnen Zeitschriftenartikel oder Buchkapitel, welche allerdings häufig zitiert werden.

Dissertationen werden in vielen Zitierstilen mit Angabe von Ort und Universität angegeben. Diese Angaben finden sich in Bibliotheksdaten sauber eingetragen und fehlen etwa bei Amazon, wo über den Buchhandel vertriebene Dissertationen meist auch im Katalog erscheinen. Bei Kongressschriften wird ebenfalls der Ort und Name der Konferenz gespeichert und kann somit für die Literaturverwaltung und schließlich für das Zitieren benutzt werden. Diese Informationen werden bei den üblichen Zitierstilen meist nur genutzt, wenn der Konferenzband nicht über einen Verlag publiziert wurde. Somit enthalten Katalogdaten für diese speziellen Publikationsarten wie etwa Hochschulschriften und Konferenzbände einen nachweisbaren Mehrwert, welcher auch bei der Literaturverwaltung relevant ist. Jedoch wächst meist gleichzeitig die Komplexität dieser Katalogisate, und somit ist eine einfache maschinelle Interpretierbarkeit nicht ohne weiteres gegeben.

### 3.2 Titelangaben

Der Titel eines Werkes setzt sich im Allgemeinen zusammen aus dem eigentlichen Titel, einem Untertitel, einem Gesamttitel oder Titel der Reihe sowie einer möglichen Bandangabe. In Zitierstilen werden diese Elemente häufig unterschiedlich formatiert, z. B. wird der Gesamttitel kursiv oder in Klammern gesetzt und daher benötigt man diese Informationen als atomare (nicht weiter aufteilbare)

<sup>25</sup> Beispielsweise wird in Zotero momentan nur der MARC-Leader herangezogen um den Publikationstyp zu bestimmen (ggf. Dissertation gesondert): O. A. *translators/MARC.js at master zotero/translators* GitHub. <https://github.com/zotero/translators/blob/master/MARC.js#L319> (6. Oktober 2015).

Elemente jeweils einzeln: „Für eine optimale automatische Verarbeitung bibliographischer Daten zu DIN 1505 konformen Bibliographien ist eine atomare und ausdrucksstarke Datenstruktur die Voraussetzung.“<sup>26</sup>

Durch die unterschiedlichen Felder und Unterfelder ist dies in bibliothekarischen Dateiformaten für die Titelangaben gegeben. Bei der Katalogisierung werden aber teilweise auch sogenannte Titelaufsätze (z. B. „mit 99 Tabellen“) als Untertitel angegeben. Solche Informationen müssen dann händisch beim Metadatenimport in ein Literaturverwaltungsprogramm gelöscht werden.

Bei hierarchischen Beschreibungen von mehrteiligen Monografien in RDA sind die für die Literaturverwaltung relevanten Daten einerseits in der untergeordneten Aufnahme für den Teil (entspricht Stückerfassung oder Bandaufführung in RAK) und andererseits in der übergeordneten Aufnahme für das Ganze enthalten. Damit können zwar die Daten sauber eingetragen werden, wird aber die automatisierte Interpretation wiederum durch die zusätzliche Komplexitätsstufe erschwert.

### 3.3 Angaben von Personen und Beziehungen

Wie viele Autoren werden beim Zitieren bzw. für das Literaturverzeichnis angegeben? Es gibt Zitierstile, in denen alle Autoren angegeben werden, egal ob dies 3 oder über 100 sind. Häufig wird aber mit „et al.“ gearbeitet. Beispielsweise listet der weithin bekannte Zitierstil der American Psychological Association (APA) im Literaturverzeichnis bis zu 7 Autoren auf, und bei 8 oder mehr werden nur noch die ersten 6 Autoren angegeben und mit et al.<sup>27</sup> angedeutet, dass es noch andere Autoren gibt. Bei Kurzreferenz im Text werden in APA bis zu 5 Autoren angegeben, aber nur bei der ersten Erwähnung, bei weiteren Erwähnungen werden nur noch 1 oder 2 Autoren angegeben.<sup>28</sup> Bei der Verwendung von Literaturverwaltungsprogrammen sind diese Regeln bereits in den jeweiligen Zitierstilen implementiert.

<sup>26</sup> Kolatzek, Robert. *Effizientere Nutzung bibliographischer Metadaten im wissenschaftlichen Bereich durch Verbesserung der Datenqualität*. Dissertation. Saarbrücken: Universität des Saarlandes, 2012. S. 89. urn:nbn:de:bsz:291-scidok-48839.

<sup>27</sup> Üblicherweise wird „et al.“ in Plural für „und andere“ benutzt und subsumiert somit mindestens zwei Autoren.

<sup>28</sup> Diese Regeln können in Ausnahmefällen noch etwas komplizierter sein. Eine schöne Übersicht und Erklärung: Lee, Chelsea. „The Proper Use of Et Al. in APA Style.“ *APA Style Blog*: 3. November 2011. <http://blog.apastyle.org/apastyle/2011/11/the-proper-use-of-et-al-in-apa-style.html> (6. Oktober 2015).



Als Endnutzerin kann man dann den Zitierstil komfortabel mit einem Klick ändern und bereits gesetzte Zitate bzw. das Literaturverzeichnis werden automatisch angepasst. Dies funktioniert aber natürlich nur, solange die in der Literaturverwaltung gepflegten Daten „vollständig“ sind. Mit RDA können im Katalog jetzt alle Autoren eingegeben werden, müssen es aber nicht. Schlussendlich liegt es auch im Ermessen der Katalogisiererin bzw. des Katalogisierers, wie viele Autoren erfasst werden. Aus Sicht der Literaturverwaltung müssten konsequenterweise alle Autoren gespeichert werden, um jegliche Zitierstile korrekt abzubilden.

Neben Autoren werden zum Zitieren auch Herausgeber, Übersetzer oder andere Beteiligte in ihren jeweiligen Rollen angegeben. Daher ist es auch wichtig, die Funktion der sonstigen Beteiligten möglichst einfach anzugeben. Maschinenlesbare, standardisierte Beziehungskennzeichnungen,<sup>29</sup> z. B. als MARC Code, können genau dies erfüllen, auch wenn man von der Menge der verschiedenen Funktionen etwas erschlagen wird.

## 4 Nachnutzung von Normdaten als Desiderat und Chance

In den meisten Literaturverwaltungsprogrammen ist es selbstverständlich möglich, durch Nutzung von dezidierten (Teil-)Feldern sowohl zwischen Vor- und Nachnamen zu unterscheiden, als auch zwischen den Funktionen in der jeweiligen Publikation zu differenzieren. Das Erfassen von Namenspräfixen,<sup>30</sup> von vormodernen Namensformen, von Titeln geistiger und weltlicher Würdenträger oder gar vom Geschlecht wird dagegen häufig nicht unterstützt. Dabei ist insbesondere letzteres gerade im deutschen Sprachraum nicht unwesentlich, soll das Literaturverwaltungsprogramm über den gewählten Zitierstil in einer Fußnote etwa bei aufeinanderfolgenden Zitaten mit der gleichen Autorenschaft automatisch zwischen „ders.“ und „dies.“

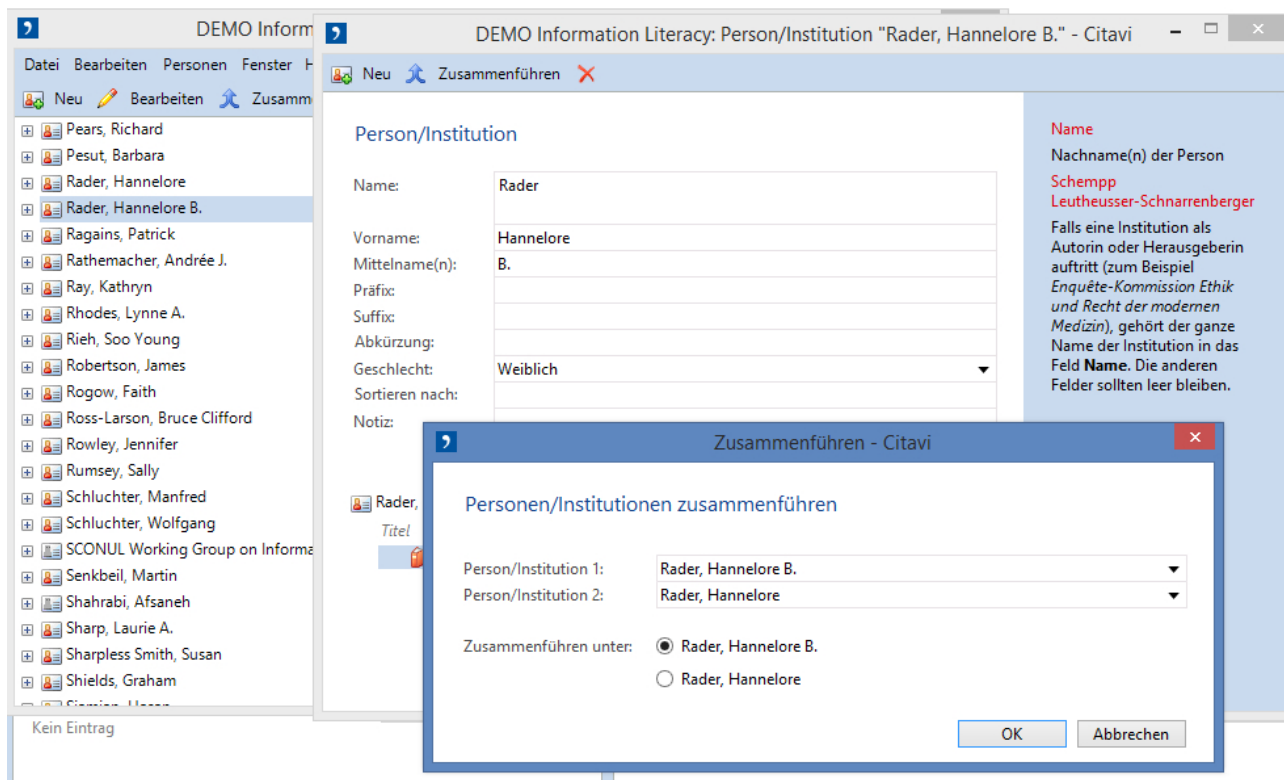
unterscheiden. Ein positives Beispiel liefert hier Citavi. Über die Liste „Personen und Institutionen“ lassen sich Detailinformationen zu in den Titeldatensätzen gespeicherten (Verfasser-)Namen hinterlegen, die wiederum zitierstilrelevant sind – vgl. die Abb. 5.

Durch Titeldatenübernahmen aus den unterschiedlichsten Quellen und manuellen Eintragungen ist es keine Seltenheit, dass dieselbe(n) Person(en) bzw. Personenna(m)e(n) formal unterschiedlich angesetzt in das genutzte Literaturverwaltungsprogramm importiert werden. Der häufigste Fall ist dabei das Fehlen des vollständigen (ersten) Vornamens. Unterschiedliche Ansetzungen stehen somit nebeneinander. Spätestens beim Zitieren bzw. bei der Erstellung eines Literaturverzeichnisses fallen entsprechende Uneinheitlichkeiten auf und müssen im Programm mühevoll korrigiert werden. Citavi bietet hier etwa als zeitsparende Hilfe die Möglichkeit, verschiedene Ansetzungen zu einer gültigen zusammenzuführen, was sich gleichzeitig auch korrigierend in den betreffenden Titeldatensätzen niederschlägt. Eine solche Funktion führt gleichzeitig jedoch als ungünstiger Nebeneffekt dazu, dass im Grunde unterschiedliche und auch gebräuchliche Namensvarianten eliminiert und für die Identifikation von Personen nicht mehr herangezogen werden können. Im Sinne einer möglichst zweifelsfreien Identifikation bzw. Unterscheidbarkeit von gleichnamigen Person ist dies ähnlich wichtig wie die Identifikation von Personen über ihre Lebensdaten, Berufe, Geburts-, Sterbe- und Wirkungsorte sowie Beziehungen zu anderen Personen, etwa Verwandtschaften. Gerade für umfangreichere Publikationsvorhaben wie etwa Dissertationen ist die Unterscheidbarkeit zwischen Personen nicht unerheblich, können doch Autoren gleichen Namens in ähnlichen Forschungsgebieten publizistisch aktiv sein. Es ist jedoch ein deutlicher Mehraufwand, situationsabhängig im gewählten Literaturverwaltungsprogramm – sofern überhaupt möglich – manuell personenbezogene differenzierte Informationen strukturiert zu hinterlegen, welche entweder inhaltliche Relevanz haben und/oder für das korrekte Zitieren benötigt werden.

In bibliothekarischen Normdateien wie der Gemeinsamen Normdatei (GND), der Library of Congress Authority File (LC Auth) oder der als Metanormdatei die Normdateien vieler Nationalbibliotheken bündelnde Virtual Authority File (VIAF) werden entsprechende Informationen seit langem nicht nur zu Personen, sondern auch zu Körperschaften, Geographika etc. gepflegt. Es erscheint daher ebenso naheliegend wie sinnvoll, dass Anwenderinnen und Anwender von Literaturverwaltungsprogrammen bereits existierende umfangreiche (Personen-) Normdateien und deren Inhalte gewinnbringend für sich

<sup>29</sup> Arbeitsstelle für Standardisierung (AFS), Hrsg. *Arbeitshilfe. Beziehungskennzeichnungen: Beziehungen zwischen einer Ressource und Personen, Familien und Körperschaften, die mit ihr in Verbindung stehen (Anhang I)*. Deutsche Nationalbibliothek, 24. August 2015. <https://wiki.dnb.de/download/attachments/106042227/AH-017.pdf?version=2&modificationDate=1441269734000&api=v2> (6. Oktober 2015).

<sup>30</sup> Diese Teile werden beispielsweise für die zitierstilgesteuerte Sortierreihenfolge benötigt. Vgl. O. A. *CSL 1.0.1 Specification – Citation Style Language 1.0.1-dev documentation*. <http://docs.citationstyles.org/en/stable/specification.html#name-part-order> (6. Oktober 2015) und O. A. *Citavi im Detail > Titel aufnehmen > Titel manuell aufnehmen > Schwierige Namen erfassen*. [http://www.citavi.com/sub/manual5/de/tips\\_for\\_names\\_of\\_persons.html](http://www.citavi.com/sub/manual5/de/tips_for_names_of_persons.html) (6. Oktober 2015).



**Abb. 5:** Beispiel für einen detaillierten Personnamendatensatz in Citavi

nachnutzen können: für eine schnellere und bessere Einheitlichkeit in der Ansetzung und Verwendung von Namensentitäten in der inhaltlichen wie formalen Literaturarbeit. Ferner können hilfreiche Zusatzinformationen zu Personen und anderen Aspekten im Literaturverwaltungssystem hinterlegt oder zu einer geeigneten Informationsquelle verlinkt bzw. aktualisierbar gemacht werden. Literaturverwaltungsprogramme könnten sich damit zu persönlichen wissenschaftlichen Assistenten mit individuellem Zugang zum Semantic Web entwickeln. Denkbar ist etwa folgendes: Vollständige Normsätze können äquivalent zu Titeldatensätzen bei der Recherche in Normdateien ad hoc strukturiert importiert, mindestens jedoch die eindeutige Identifikationsnummern von Datensätzen aus der jeweiligen Normdatei im Programm bzw. Titeldatensatz hinterlegt und mit der Normdatei verlinkt und dynamisch bei Nutzung, etwa im Zuge von Zitationen, abgefragt werden. An den Normdateien und ihren Anbieterinnen selbst liegt es dabei nicht. So ist etwa die GND unlängst unter der Creative-Commons-Zero-Lizenz (CC0) als Linked Open Data veröffentlicht worden.<sup>31</sup> Man kann die einzelnen

Datensätze als RDF/XML-Repräsentation über Content Negotiation der URI oder einen Direktlink abrufen, einen kompletten Abzug (Dump) der GND herunterladen sowie über die Schnittstellen OAI oder SRU holen.<sup>32</sup>

Es gibt jedoch einen entscheidenden Haken: Normalerweise decken bibliothekarische Normdaten für Personen nicht die Articleebene ab. Autorinnen und Autoren bekommen in der Regel erst einen Normdateneintrag, wenn sie ein Buch geschrieben bzw. veröffentlicht haben. In der Wissenschaft ist dies zumeist die Dissertation. Letztlich ist es den entweder zentral oder – wie im Falle der GND – kooperativ arbeitenden Normierungsredaktionen bzw. den Katalogisiererinnen vorbehalten, wann und in welcher Erschließungstiefe ein Normsatz für eine Person angelegt und gepflegt wird. Wissenschaftlerinnen und Wissenschaftler selbst haben darauf, abgesehen von Korrekturanfragen, keinen Einfluss.

Als Alternative für die Identifikation und Disambiguierung lebender Personen im Wissenschaftsbetrieb existiert beispielsweise die Open Researcher and Contributor ID,

<sup>31</sup> Vgl. Hentschke, Jana. *Linked Data Service der Deutschen Nationalbibliothek*. 1. Oktober 2015. [http://www.dnb.de/DE/Service/Digitale-Dienste/LinkedData/linkedata\\_node.html](http://www.dnb.de/DE/Service/Digitale-Dienste/LinkedData/linkedata_node.html) (6. Oktober 2015).

<sup>32</sup> Vgl. O. A. *Der Linked Data Service der Deutschen Nationalbibliothek: Hinweise zum Zugriff*. Deutsche Nationalbibliothek, 24. Februar 2014. <http://www.dnb.de/SharedDocs/Downloads/DE/DNB/service/linkedataZugriff.pdf> (6. Oktober 2015).

kurz ORCID,<sup>33</sup> welche – neben der eindeutigen Personenidentifikation selbst – die elektronische Zuordnung von Publikation und Autorinnen ermöglichen bzw. stark erleichtern soll. Zu den Gründungsmitgliedern des Trägers, der Open Researcher Contributor Identification Initiative, gehören zahlreiche renommierte wissenschaftliche Verlagsgruppen und Forschungsorganisationen, was der ORCID Seriosität und Beständigkeit mit dem Ziel eines Standards für Autorenidentifikation wissenschaftlicher Publikationen verleiht. Der ORCID-Identifizier besteht wie der International Standard Name Identifier (ISNI, ISO-Standard 27729)<sup>34</sup> aus 16 Zeichen (in Vierer-Gruppen geschrieben), dessen erste fünfzehn Zeichen Dezimalziffern sind und dessen letztes eine Prüfziffer ist, die auch „X“ sein kann.<sup>35</sup> Eine ORCID und das dazugehörige Profil kann von jeder interessierten Person registriert und die Publikationszuordnung unter bestimmten Regeln selbstständig gepflegt werden. Dies kommt insbesondere der eindeutigen Zuordnung bzw. Zitation unselbstständiger Publikationen zu Gute.<sup>36</sup> Dementsprechend macht es nicht zuletzt im Kontext von Literaturverwaltung Sinn, ORCIDs in den einschlägigen Programmen hinterlegen und die damit verbundenen Informationen via Verlinkungen nutzen zu können. Diese Idee ist nicht wirklich neu. So hat Martin Fenner bereits im Jahr 2011 vor dem Launch des ORCID-Services darauf hingewiesen, dass die bereits erwähnten Standardaustauschformate wie BibTeX und RIS standardisierte Identifier zur Autorenidentifikation nicht enthalten (können) und damit unzureichend sind.<sup>37</sup> Mittelbar gilt das auch für Literaturverwaltungspro-

gramme, denn seither hat sich diesbezüglich wenig getan. Eine Integration gibt es gegenwärtig im Falle von EndNote mit der ORCID-kompatiblen ResearcherID.<sup>38</sup> Eine Verbindung von Mendeley zum ORCID-Webservice wurde inzwischen ebenso entwickelt.<sup>39</sup> Diese bezieht sich jedoch nur auf Mendeleys Repository, den Research Catalog. Eine Hinterlegung von ORCIDs in den in Mendeley gespeicherten Metadaten ist hingegen (noch) nicht möglich.

Eine Veränderung der Datenstrukturen bzw. -modelle sowie neue funktionale Möglichkeiten von Literaturverwaltungsprogrammen sind jedoch insbesondere an die weitere Entwicklung von elektronischen Publikationskulturen und von Nutzbedürfnissen gebunden. Mehrheitlich imitieren elektronische wissenschaftliche Veröffentlichungen über das PDF-Format noch das gedruckte physische Medium. Das zieht auch das vorläufige Beibehalten weitestgehend traditioneller Formen des Zitierens durch die Angabe von korrekten Seitenzahlen, Heft- und Bandzählungen etc. nach sich. Werden die multimedialen Mittel und Möglichkeiten des World Wide Webs ausgeschöpft, werden sich zwangsläufig Zitationsgewohnheiten und -kulturen noch stärker ändern (müssen). Die Einführung und inzwischen immer gängiger werdende Nutzung bzw. Registrierung von Digital Object Identifiers (DOI) als persistente Identifier für elektronische Publikationen und Objekte, etwa Forschungsdaten, ist hier ein wichtiges Vehikel. DOIs können in den gängigen Literaturverwaltungsneuerdings in expliziten Feldern hinterlegt werden – nicht nur für die Angabe von bzw. Verlinkung zu Volltexten, sondern auch für den Import von Metadaten, z. B. via CrossRef. In Citavi etwa ist es äquivalent zum integrierten ISBN-Update möglich, via DOI-Abfrage ad hoc vorhandene Titeldatensätze zu ergänzen oder zu überschreiben. Für Zotero ist ein ähnlicher Mechanismus in Planung. Der Nutzerin hilft es hier natürlich am meisten, wenn eine DOI bereits in den (aus den Bibliothekskatalogen) importierten Titeldatensätzen enthalten ist, um über diese dann bei Bedarf Aktualisierungen vorzunehmen bzw. in den Zitierungen zu berücksichtigen. Darüber hinaus sollten DOIs für jeden Publikationstyp in Literaturverwaltungen

33 O. A. ORCID. <http://orcid.org/> (6. Oktober 2015).

34 O. A. ISNI. <http://www.isni.org/> (6. Oktober 2015).

35 Vgl. O. A. *Structure of the ORCID Identifier – Feedback & Support for ORCID*. <http://support.orcid.org/knowledgebase/articles/116780-structure-of-the-orcid-identifier> (6. Oktober 2015) sowie Überblicksartig: Impactstory Team. „Ten things you need to know about ORCID right now.“ *Impactstory blog*. 10. April 2014. <http://blog.impactstory.org/ten-things-you-need-to-know-about-orcid-right-now/> (6. Oktober 2015).

36 Im Rahmen des Bibliothekartags 2015 gab es eine dezidierte Session zum Thema: O. A. „Persistent Identifiers für Kultur und Wissenschaft.“ *Workshop auf dem 104. Bibliothekartag in Nürnberg*. 22. Juli 2015. <https://dini.de/service/nachrichten/nachricht/x//persistent-identifiers/> (6. Oktober 2015). In dieser wurde unter anderem angesprochen, dass ORCID in Deutschland noch zu wenig genutzt wird. Erst seit März ist es im Sinne der Verschränkung mit Normdaten möglich ORCIDs in GND-Sätzen zu hinterlegen. Zuletzt gab es Ende August 2015 dazu in der Mailingliste InetBib eine kurze, lebendige Diskussion anlässlich einer ORCID-Umfrage. Vgl. den Diskussionsbeginn: Reimer, Torsten F. [InetBib] *Umfrage zum ORCID Research Identifier*. 26. August 2015. <http://www.ub.uni-dortmund.de/listen/inetbib/msg56257.html> (6. Oktober 2015).

37 Vgl. Fenner, Martin. „Why BibTeX, RIS and Endnote XML will soon be broken.“ *Gobbledygook*. 8. November 2011. <http://blogs.plos.org/mfenner/2011/11/08/why-bibtex-ris-and-endnote-xml-will-soon-be-broken/> (6. Oktober 2015).

38 Vgl. O. A. *ResearcherID | Thomson Reuters*. <http://thomsonreuters.com/en/products-services/scholarly-scientific-research/authoring-and-collaboration-tools/researcherid.html> (6. Oktober 2015).

39 Vgl. Bando, Keita. „Mendeley to ORCID.“ *ORCID Outreach Meeting*. Chicago, Ill., Mai 2014. [http://www.researchgate.net/profile/Keita\\_Bando/publication/263124878\\_Mendeley\\_to\\_ORCID\\_Short\\_paper\\_for\\_ORCID\\_Outreach\\_Meeting\\_May\\_2014/links/54df103b0cf296663788b6b1.pdf](http://www.researchgate.net/profile/Keita_Bando/publication/263124878_Mendeley_to_ORCID_Short_paper_for_ORCID_Outreach_Meeting_May_2014/links/54df103b0cf296663788b6b1.pdf) (6. Oktober 2015).

hinterlegbar sein. Eine konsequente Vergabe und Nutzung von DOIs hat das Potential, sowohl das Zitieren zu vereinfachen, als auch die Literaturverwaltung selbst tiefgreifend zu verändern, wie Lambert Heller 2011 schrieb.<sup>40</sup>

## 5 Fazit

Im vorliegenden Artikel sind wir der Frage nachgegangen, inwiefern elektronische Bibliothekskataloge tatsächlich eine attraktive Quelle zur Übernahme von bibliographischen Daten in Literaturverwaltungsprogramme darstellen. Dies geschah vor dem Hintergrund, dass diese Programme funktional zunehmend den gesamten wissenschaftlichen Arbeitsprozess unterstützen und sich damit zu selbstverständlichen Begleitern im wissenschaftlichen Arbeiten von der Recherche bis zum Publizieren etablieren.

Die **bibliothekarischen Daten** sind zumeist nicht schlechter, sondern vielmehr qualitativ hochwertiger als bibliographische Metadaten anderer Quellen, aber das Potential auch für deren Nachnutzung in Literaturverwaltungsprogrammen erscheint uns noch nicht ausgeschöpft. Die speziellen Publikationsarten sind meist in einer einzigartigen Detailtiefe erfasst, aber in sehr komplexen und schwer verständlichen Datenformaten verborgen.

Es gibt kein ausreichend standardisiertes, vollständiges und weit verbreitetes **bibliographisches Datenformat**. Bei den Formaten RIS und BibTeX fehlt nicht nur die Standardisierung in der Praxis, sondern auch die Vollständigkeit. MARCXML hat nur eine sehr begrenzte Reichweite. Bibliotheken haben Expertise bei der Standardisierung und Pflege von bibliographischen Datenformaten, haben aber bisher keinen intensiveren Dialog mit den Anbietern von Literaturverwaltungsprogrammen gesucht.

Verknüpfte **Normdaten** etwa bei Autoren finden momentan in den Literaturverwaltungsprogrammen noch keine Verwendung. Die Weiterentwicklung dieser Programme zu persönlichen wissenschaftlichen Assistenten könnte insbesondere in Verbindung mit den Techniken des Semantic Webs neue Perspektiven bringen.

Leider ist die Ausrichtung der **bibliothekarischen Regelwerke**, welche sich häufig auch in der Konzeption von Webkatalogen widerspiegelt, noch sehr davon geprägt, dass die Daten nur für die menschlichen Nutzer ansprechend dargestellt werden sollen. Hier fehlt häufig

die Vorstellung, dass auch Software-Tools wie etwa Literaturverwaltungsprogramme die Daten verstehen müssen. Gefragt sind hier also Metadatenmanager, Katalogisierer und all diejenigen, die diese Regelwerke aktiv weiterentwickeln. Bei der praktischen Katalogisierung in Anwendung der geltenden Regelwerke ist unserer Meinung nach stets auch das Bedürfnis von Nutzerinnen und Nutzern, die Katalogisate für die eigene (wissenschaftliche) Arbeit nutzen, zu berücksichtigen. Entsprechend sollten auch bei Entscheidungen „im Ermessen des Katalogisierers“ darauf Rücksicht genommen werden. Die Sensibilisierung für Anwendungen der Literaturverwaltung erscheint uns notwendig, damit deren Entwicklung mittels adäquater Erstellung und Bereitstellung von Katalogdaten durch Bibliotheken als moderne Informationsdienstleisterinnen maßgeblich (mit-)begleitet wird.

**Danksagung:** Der vorliegende Artikel ist aus der Diskussion im Blog Literaturverwaltung zum Post „Wünsch dir was“-Runde zur Literaturverwaltung auf dem 7. Bibcamp in Potsdam heraus entstanden.<sup>41</sup> Sebastian Karcher (adamsmith) hat mitdiskutiert und wertvolle Anregungen gegeben, wofür wir uns herzlich bei ihm bedanken.

## Autoreninformationen



**Matti Stöhr**

Rudolf-Reusch-Str. 23D  
10367 Berlin  
[matti.stoehr@online.de](mailto:matti.stoehr@online.de)  
<http://www.matti-stoehr.de>  
[orcid.org/0000-0002-2874-4832](http://orcid.org/0000-0002-2874-4832)



**Dr. Philipp Zumstein**

Universitätsbibliothek Mannheim  
Schloss Schneckenhof West  
68131 Mannheim  
[philipp.zumstein@bib.uni-mannheim.de](mailto:philipp.zumstein@bib.uni-mannheim.de)  
[orcid.org/0000-0002-6485-9434](http://orcid.org/0000-0002-6485-9434)

<sup>40</sup> Vgl. Heller, Lambert. „Zitieren mit URIs: Anfang vom Ende der Literaturverwaltung.“ *Literaturverwaltung*. 21. September 2011. <https://literaturverwaltung.wordpress.com/2011/09/21/zitieren-mit-uris/> (Zugriffen: 6. Oktober 2015).

<sup>41</sup> Vgl. die Kommentare zu: Stöhr, Matti. „Wünsch dir was“-Runde zur Literaturverwaltung auf dem 7. Bibcamp in Potsdam. *Literaturverwaltung*. 1. Oktober 2014. <https://literaturverwaltung.wordpress.com/2014/10/01/wunsch-dir-was-runde-zur-literaturverwaltung-auf-dem-7-bibcamp-in-potsdam/> (6. Oktober 2015).