

# Forum

MARKUS LINTEN / BODO RÖDEL / CHRISTIAN WOLL

## **Wo und wie finde ich die Nadel im Heuhaufen? Repositorien als Wissensspeicher für die Berufsbildung**

Haben Sie schon einmal versucht, ihre 25 Jahre alten VHS-Video-Kassetten zu digitalisieren, oder waren die Bilder schon verschwunden, da sich mit jedem Abspielen der VHS-Kassetten die Qualität minimal verschlechtert? Sollten Sie ihre Diplomarbeit vor 15 Jahren vielleicht auf einem Atari-PC geschrieben haben, dürfte es ebenfalls schwierig werden, diesen Text heute noch digital weiter zu bearbeiten.

Was hat das nun mit Berufsbildungsforschung zu tun? Diese alltäglichen Beispiele zeigen zum einen, wie flüchtig unsere Wissensspeicher heute sind und zum anderen, welche Herausforderung es ist, Daten über einen langen Zeitraum verfügbar zu halten. Zu diesen zwei Problemen kommt ein drittes: Die Flut an Informationen steigt seit Jahren exponentiell an. Im Bibliothekswesen geht man von einer quantitativen Verdopplung (wissenschaftlicher) Publikationen etwa alle 10 bis 15 Jahre aus (vgl. GANTZ 2011). Damit steht auch die Berufsbildungsforschung vor einer zentralen Herausforderung: Wie bleiben ihre wissenschaftliche Veröffentlichungen langfristig verfügbar und wo finden Forscher die ‚Nadel im Heuhaufen‘, sprich relevante Fachinformationen? Bereits 1970 sprach ALVIN TOFFLER vom „information overload“ und bezeichnete damit den Zustand des „zu viel an Informationen“, der letztlich Entscheidungsunfähigkeit hervorruft (vgl. TOFFLER 1970). Mit Blick auf das Internet ist festzustellen, dass nur ein geringer Teil des Web über Suchmaschinen überhaupt recherchierbar ist – wie groß dieser Anteil ist und welches Ausmaß das sogenannte ‚invisible web‘ hat, darüber streiten sich die Expertinnen und Experten.

Aus dem „Nadel-im-Heuhaufen-Problem“ ergibt sich eine neue Schlüsselqualifikation für Wissenschaftler: Geeignete Informationsquellen im Internet verifizieren und so nutzen können, dass ein Informationsbedarf bei Anwendung effektiver Suchstrategien erfolgreich befriedigt wird. Betrifft dies zum einen die Methodenkompetenz der Nutzer des Informationsangebotes, wird auch für die Community der Berufsbildungsforschung und -praxis die eingangs skizzierte grundlegende Fragestellung zunehmend virulent: Wie wird wissenschaftliche Literatur über einen langen Zeitraum so archiviert, dass sie problemlos wiedergefunden werden kann? Ist das Verlagern dieser Daten in so genannte Clouds eine Alternative, mit denen das Problem von ‚Big Data‘ gleich mitgelöst wird? Unter Big Data werden neue Methoden und Technologien subsumiert, die es erlauben, eine riesige Anzahl von Daten zu erfassen, zu speichern und zu analysieren (vgl. HORVATH 2013). Parallel zu dieser Frage findet die Diskussion um den freien Zugang zu wissenschaftlicher Literatur (open access) und frei verfügbarem Lehr-Lern-Material (open educational resources) statt. Damit verbunden stellen sich unweigerlich Fragen des Urheber- und Verwertungsrechts sowie der Finanzierung bzw. Vergütung solcher Publikationen.

Versuchen wir nun mit Blick auf die Methodenkompetenz und zumindest den ersten Teil der Langzeitarchivierungsfrage eine Antwort.

## Repositorien sichern eine Langzeitarchivierung und fördern das Open-Access-Prinzip

Die Langzeitarchivierung wissenschaftlicher Literatur wird heute in zahlreichen Wissenschaftsdisziplinen durch sogenannte Repositorien sichergestellt. Hierbei handelt es sich um Volltextserver, über die systematisch qualitätsgesicherte wissenschaftliche Literatur bereitgestellt wird. In Deutschland gibt es zurzeit rund 180 Repositorien (vgl. DINI 2014), weltweit über 2.600.

Die Gewährleistung langfristiger Verfügbarkeit der dort abgelegten Volltexte ist zentrale Funktion von Repositorien. Mit Blick auf die Einhaltung von Regeln guter wissenschaftlicher Praxis ist auch die Frage der Zitierfähigkeit digitaler Publikationen von Bedeutung. Die herkömmlichen Webadressen (URLs) reichen für diesen Zweck allein nicht aus. Grund hierfür ist, dass über eine URL nicht das digitale Objekt selbst, sondern nur der „Ort“, an dem es liegt, spezifiziert wird. Bei Verlagerung der Inhalte wird eine URL ungültig – das bekannte Phänomen des „toten Links“. Ein weiteres Problem stellt die einfache Veränderbarkeit elektronischer Dokumente dar. Diese birgt die Gefahr, dass unter derselben URL zu einem späteren Zeitpunkt eine andere Version der gleichen Publikation zu finden ist, schlimmstenfalls vielleicht sogar eine gänzlich andere Publikation. Dies erschwert die Nachprüfbarkeit von Zitaten in erheblichem Maße.

Ein Lösungsansatz für diese beiden Probleme – Flüchtigkeit und Veränderbarkeit – sind sogenannte Persistent Identifier. Sie sind vergleichbar mit den Standardnummern „ISBN“ und „ISSN“ für gedruckte Literatur. Durch sie wird sichergestellt, dass auf eine Veröffentlichung auch dann noch zugegriffen werden kann, wenn sich ihr physikalischer Speicherort geändert hat. Deshalb werden die Volltexte in Repositorien grundsätzlich mit einem solchen Identifier, beispielsweise einem Uniform Resource Name (URN) oder einem Digital Object Identifier (DOI), versehen. Auf diese Weise erhalten die digitalen Publikationen eine weltweit eindeutige Kennung und sind damit zuverlässig zitierbar.

Repositorien tragen somit auch zur Umsetzung des Open Access-Prinzips bei. Viele Institutionen, die ein Repository betreiben, haben die Berliner Erklärung unterzeichnet bzw. eine Open Access Policy verabschiedet.

## Repositorien mit unterschiedlichen Ausrichtungen

Bezogen auf die inhaltliche Ausrichtung sind zwei Grundformen von Repositorien zu unterscheiden: institutionelle und fachliche Repositorien.

In institutionellen Repositorien werden überwiegend die Publikationen der Mitarbeiterinnen und Mitarbeiter einer Einrichtung gespeichert. Beispiele hierfür sind die Dokumentenserver von Hochschulen sowie außeruniversitärer Forschungsorganisationen (Max-Planck-Gesellschaft, Fraunhofer-Gesellschaft und Helmholtz-Gemeinschaft). Im Unterschied dazu sind fachliche Repositorien institutionsübergreifend ausgerichtet. Hier fließen sämtliche Publikationen einer bestimmten Disziplin oder Teildisziplin ein. Beispiele für fachbezogene Repositorien aus den Bezugswissenschaften der Berufsbildung sind PEDOCS (Erziehungswissenschaften), EconStor (Wirtschaftswissenschaften) und SSOAR (Sozialwissenschaften). Die meisten Repositorien, insbesondere die fachlichen, zeichnen sich außerdem durch eine spezifische Recherchefunktionalität aus. Sie sind über eine eigene Webpräsenz erreichbar, auf welcher verschiedene Suchmöglichkeiten angeboten werden. Diese ermöglichen sowohl ein gezieltes Auffinden spezifischer Dokumente, als auch ein eher intuitives „Durchstöbern“ des Datenbestands.

Ein Repository speziell für die Berufsbildungsforschung fehlt bislang. Dies ist umso beachtlicher, als die drei genannten Repositorien in Bezug auf den Themenbereich „Berufsbildung“ nur einen sehr geringen Abdeckungsgrad aufweisen, was sich an verschiedenen Parametern festmachen lässt. Analysiert man beispielsweise, inwieweit Artikel der für die Berufsbildung einschlägigen Zeitschriften in diesen Repositorien zu finden sind, zeigt sich ein ernüchterndes Bild: die Anzahl der Beiträge aus online verfügbaren Bildungszeitschrif-

ten wie *BWP*, *bwp@*, *denk-doch-mal* oder *Die berufsbildende Schule* tendiert in allen drei Repositorien gegen Null. Ähnlich sieht es bei der vergleichenden Suche nach einschlägigen Schlagwörtern aus dem Berufsbildungskontext aus (vgl. Tab.).

Tabelle: Anzahl der Nachweise in den drei bestehenden Repositorien und in der Literaturlatenbank Berufliche Bildung (LDBB); Stand: Februar 2014

Schlagwort	peDOCS	SSOAR	Econstor	LDBB
Berufsbildungssystem	17	0	101	1 304
Aufstiegsfortbildung	0	1	0	34
Berufskonzept	2	7	0	324
Übergangssystem	2	0	0	378

Noch überhaupt nicht in den Blick genommen wurden hierbei internationale Literatur oder Forschungsdaten.

### Die Berufsbildung benötigt ein fachliches Repository

Die im Februar 2014 durchgeführte Analyse zur inhaltlichen Abdeckung durch bestehende Repositorien unterstreicht die Notwendigkeit eines eigenen Repositoriums für die Berufsbildung und den daraus resultierenden Handlungsbedarf. Faktisch verhält es sich derzeit so, dass die über 5.000 in der Literaturlatenbank Berufliche Bildung (LDBB) nachgewiesenen Veröffentlichungen auf Hunderten von Servern zwischen Flensburg und Wien sowie Zürich und Berlin verteilt ‚liegen‘ – bei Instituten, Universitäten, Arbeitgeber- und Berufsverbänden, Parteien, Gewerkschaften oder Verlagen. Eine *dauerhafte* Verfügbarkeit dieser Dokumente kann für die Wissenschaftscommunity nicht sichergestellt werden. Um wissenschaftliche Erkenntnisse der Berufsbildung für alle Interessierten zugänglich zu machen, muss der Zugriff auf wissenschaftliche Volltexte mit einer benutzerfreundlichen Rechercheoberfläche gewährleistet werden. Dauerhafte Verfügbarkeit der Publikationen, ein hoher Abdeckungsgrad des Gegenstandsbereichs und intuitive Bedienbarkeit des Volltextservers sind dabei zentrale Qualitätskriterien. Last but not least würde ein fachliches Repository auch die Sichtbarkeit der Berufsbildungsforschung in der gesamten Forschungscommunity weiter verbessern.

### Literatur

- DINI: <http://dini.de/dini-zertifikat/liste-der-repositorien/> (Stand: 13.05.2014)
- GANTZ, J.; REINSEL, D.: Extracting Value from the Chaos. (2011) 12 S. – URL: <http://germany.emc.com/collateral/analyst-reports/idc-extracting-value-from-chaos-ar.pdf> (Stand: 13.05.2014)
- HORVATH, S.: Aktueller Begriff: Big Data. - Deutscher Bundestag/Wissenschaftliche Dienste. (2013) – URL: [www.bundestag.de/dokumente/analysen/2013/Big\\_Data.pdf](http://www.bundestag.de/dokumente/analysen/2013/Big_Data.pdf) (Stand: 13.05.2014)
- IDC: The digital universe of opportunities. Rich data & the increasing Value of the internet of things. (2014) – URL: [www.emc.com/collateral/analyst-reports/idc-digital-universe-2014.pdf](http://www.emc.com/collateral/analyst-reports/idc-digital-universe-2014.pdf) (Stand: 13.05.2014)
- TOFFLER, A.: Future Shock. Random House 1970

Anschrift der Autoren: Markus Linten, Bundesinstitut für Berufsbildung, Arbeitsbereich „Publikationsmanagement/Bibliothek“, Robert-Schuman-Platz 3, 53175 Bonn, E-Mail: [Linten@bibb.de](mailto:Linten@bibb.de)  
 Dr. Bodo Rödel, Bundesinstitut für Berufsbildung, Arbeitsbereich „Publikationsmanagement/Bibliothek“, Robert-Schuman-Platz 3, 53175 Bonn, E-Mail: [Roedel@bibb.de](mailto:Roedel@bibb.de)  
 Christian Woll, Bundesinstitut für Berufsbildung, Arbeitsbereich „Publikationsmanagement/Bibliothek“, Robert-Schuman-Platz 3, 53175 Bonn, E-Mail: [Woll@bibb.de](mailto:Woll@bibb.de)