

Rezension zu: Pomerantz, Jeffrey. Metadata. The MIT Press Essential Knowledge Series. The MIT Press, 2015. ISBN: 978-0-262-52851-1. \$15,95

Najko Jahn

Die handliche Reihe „The MIT PRESS Essential Knowledge Series“ widmet eine ihrer neuesten Einführungen dem Thema Metadaten. Jeffrey Pomerantz, als Informationswissenschaftler zuletzt an der School of Information and Library Science der University of North Carolina in Chapel Hill tätig, stellt in acht übersichtlichen Kapiteln die Frage nach Gegenstand und Zweck von Metadaten. Dabei beschränkt er sich in seinem Buch, das aus einem MassiveOpen Online Course (MOOC) heraus entstanden ist, nicht nur auf die Erschließung wissenschaftlicher Veröffentlichungen oder des Kulturerbes. Vielmehr, so seine These, seien Metadaten aufgrund des Mengenwachstums digitaler Ressourcen sowie ihrer Vielfalt für deren Auffindbarkeit und Nachnutzung relevanter denn je zuvor.

Zu Beginn grenzt Pomerantz Metadaten von den Begriffen Information und Daten ab. Er definiert Metadaten als Aussagen über möglicherweise informative Gegenstände („Metadata is a statement about potentially informative objects“). Mit dem Bezug auf Aussagen, welche aus Subjekt, Prädikat und Objekt bestehen, möchte er sich einen Freiraum verschaffen, um im Verlaufe des Buches auf die Beschreibungsstandards Dublin Core und das Resource Description Framework (RDF) gesondert einzugehen. Zu den weiteren Verfahren der Erschließung, in die Pomerantz zu Beginn einführt und die er in Form eines Glossars übersichtlich zusammenfasst, gehören Normdateien, Thesauri und Ontologien sowie die freie Stichwortvergabe mittels Tags.

Nach der Definition der Grundbegriffe und Verfahren der Erschließung unterscheidet die Einführung zwischen deskriptiven, administrativen und Nutzungs-Metadaten. Es stellt Dublin Core als Beispiel für einen deskriptiven Metadatenstandard vor. Erdacht als kleinster gemeinsamer Nenner während eines Workshops in Dublin (Ohio), dem Heimatort des Bibliotheksdienstleisters OCLC, umfasst Dublin Core fünfzehn Elemente zur Beschreibung von Ressourcen im Web. Zwar sei laut Pomerantz Dublin Core für die Websuche heutzutage kaum mehr relevant. Allerdings zeigt er, dass sich gegenwärtige Standards in ihrer Arbeit weiterhin auf Dublin Core und dessen Gestaltungsprinzipien wie dem universellen Anspruch und die technisch einfache Implementierung beziehen.

Administrative Metadaten sind Metadaten, die den Umgang mit einer Ressource dokumentieren. Sie umfassen Aussagen zu Technik, Bewahrung oder den rechtlichen Rahmenbedingungen, unter denen eine Ressource verwendet werden darf. Mit der Provenienz führt das Buch in einen wichtigen Aspekt für die Bewertung der Vertrauenswürdigkeit einer Ressource ein. Wer hat die

Ressource erstellt, welche Personen, Institutionen oder technischen Maßnahmen waren ebenfalls beteiligt oder wo ist die Ressource erschienen, sind Fragen, die insbesondere in den Wissenschaften zur Bewertung einer Quelle relevant sind. Provenienzaussagen können sich hierbei auf eine Sammlung, ein Werk oder auch auf einzelne Messwerte in einem experimentellen Datensatz beziehen.

Während Pomerantz deskriptive und administrative Metadaten überzeugend einführt, ist das Kapitel zu Nutzungsmetadaten („Use metadata“) weniger gelungen. Zwar weist er darauf hin, dass Nutzungsereignisse in verschiedenen Bereichen wie dem Handel, der Sicherheitspolitik oder Bildung für analytische Zwecke abgeschöpft werden. Allerdings mangelt es, wie Pomerantz eingesteht, an terminologischer Klarheit, was unter „Nutzungsmetadaten“ zu verstehen sei. Möglicherweise wäre hier der Rückgriff auf die „Altmetrics“ und deren Verwendung für die Auffindbarkeit wissenschaftlicher Werke hilfreich gewesen, welche allerdings keine Erwähnung im Buch findet.

In den nachfolgenden Kapiteln stellt Pomerantz kurz und allgemeinverständlich die technischen Grundlagen für die Erstellung von Metadaten vor. Ein besonderes Augenmerk legt er hierbei auf das Semantic Web, und zwar insbesondere auf die Frage, wie Gedächtnisinstitutionen ihre Nachweise und Erschließungsverfahren so erweitert haben, damit ihre Werke im Web auffindbar sind und sich mit anderen Quellen wie der DBpedia vernetzen lassen. Er illustriert zudem, wie der Webstandard Schema.org zur strukturierten Suche von Ressourcen über Suchmaschinen wie Google oder Bing beiträgt.

Die Einführung schließt mit einen Ausblick: Domänenspezifische Metadaten, die sich an den allgemeinen Standards orientieren, werden zur Nachnutzung vielfältiger digitaler Ressourcen ebenso beitragen wie der mittlerweile obligatorische Zugriff auf diese Metadaten mittels Webschnittstellen wie die Application Programming Interface (API). Ein Gebiet, auf das Pomerantz gesondert eingeht, ist die eScience. Hier sieht er einen besonderen Bedarf an Metadaten, die den Erstellungsverlauf beschreiben, um Vertrauen in die Provenienz einer wissenschaftlichen Ressource zu schaffen. Kritisch sieht Pomerantz zum Abschluss die Aufzeichnung und Weitergabe von Metadaten, die die private Kommunikation betreffen. Ohne etwa den Inhalt von Telefongesprächen aufzuzeichnen, ließen sich über Metadaten zum Gesprächsverlauf Rückschlüsse auf persönliche Präferenzen ziehen. Aus diesem Grunde würden seines Erachtens Metadaten in Zukunft verstärkt an gesellschaftlicher Relevanz gewinnen und Gegenstand rechtlicher und politischer Auseinandersetzungen.

Die abschließende Betonung normativer Aspekte im Umgang mit Metadaten weist auf ein Desiderat der Einführung hin: Es fehlt ein eigenständiges Kapitel über die Rolle von Werten im Erschließungs- und Bewahrungsprozess. Entscheidungen darüber, wie Bücher oder wissenschaftliche Datensätze erschlossen werden, können maßgeblichen Einfluss auf deren Auffindbarkeit und Rezeption haben. Sowohl die Bibliotheks- und Informationswissenschaft als auch gesellschaftliche Initiativen wie die „Radical Reference Collectives“ in den USA problematisieren, dass Metadaten trotz ihres objektiven Anspruchs nicht davor gefeit sind, Annahmen oder befangene Sichtweisen wiederzugeben, welche Erkenntnisinteressen oder gesellschaftliche Gruppen marginalisieren. Ein Kapitel, das insbesondere die Rolle von sozialen Werten bei der Erschließung im Informationswesen und der Wissenschaft aufgreift, wäre daher eine verdienstvolle Ergänzung zur ansonsten lesenswerten Einführung.

Insgesamt bietet die handliche Einführung leicht verständliche Antworten darauf, wie Metadaten zur Auffindbarkeit und Nachnutzung von Ressourcen im Web beitragen. Es ist daher nicht nur Studierenden der Bibliotheks- und Informationswissenschaft zu empfehlen, die nach einem lesenswerten Einstieg in die Materie der Erschließung suchen. Vielmehr ist es ein Verdienst der Einführung, dass sie vielfältige Erschließungsverfahren vorstellt und hierbei besonders auf aktuelle Webstandards eingeht. Jeffrey Pomerantz' „Metadata“ ist daher auch den Praktikerinnen und Praktikern sehr zu empfehlen.

Najko Jahn ist Referent für Projektentwicklung und Innovationsmanagement an der Universitätsbibliothek Bielefeld. Derzeit besucht er an der University of Exeter die Data Studies Group von Dr. Sabina Leonelli, die die epistemischen Bedingungen datenbasierter Wissenschaft erforscht.