

■ **Miller, Stephen J.: Metadata for digital collections – a how-to-do-it manual. London: Facet 2011. XXIII, 343 S. ISBN 978-1-85604-771-5. £ 54,95**

Der anglo-amerikanische Raum, v.a. die USA, sind in Sachen digitale Bibliotheken in praktischer wie theoretischer Hinsicht Vorreiter mit im Vergleich zu Europa langer Tradition¹. Zahlreiche Universitäts- und öffentliche Bibliotheken haben dort seit den 1990er Jahren Digitalisierungsvorhaben umgesetzt und „digital collections“ verschiedenster Materialarten aufgebaut, die sie online zugänglich machen.

Das hier besprochene englischsprachige Buch ist hilfreicher Begleiter für jeden, der digitale Bestände anbieten möchte. Der Autor wählt drei der meistgenutzten Metadaten-schemata (Dublin Core, MODS - Metadata Object Description Schema und VRA - Schema der Visual Resources Association) und kann diese dadurch eingehend und praxisorientiert behandeln. Die begleitende Website (<http://www.neal-schuman.com/metadata-digital-collections>) enthält vertiefende Fragen und Übungen sowie weitere Literaturangaben zu den einzelnen Kapiteln.

Kapitel 1 setzt den thematischen Rahmen. Grundlegende Definitionen erklären die Begriffe Metadaten (für Aspekte wie Beschreibung, Verwaltung, Langzeitarchivierung, Struktur) und deren Standards und umreißen den Prozess der Metadatenplanung und -erstellung für digitale Sammlungen. Charakteristisch für die Anwendung beschreibender Metadaten, im Gegensatz zur klassischen Formalerschließung, ist größere Flexibilität, da in der Regel für jede digitale Sammlung die etablierten Standardschemata durch speziell auf den jeweiligen Schwerpunkt zugeschnittenen Feldern abgewandelt werden, die sich idealerweise auf Felder eines Standards (meist Dublin Core als „Lingua Franca“) abbilden lassen. Mehrfach wird auf dieses Spannungsfeld zwischen lokalen Bedürfnissen, an die die Metadaten angepasst werden, und internationalen Standards, die für Datenaustausch und Harvesting wichtig sind, hingewiesen.



Die Kapitel 2, 3 und 4 stellen das meistgenutzte Schema Dublin Core vor. Der Autor konzentriert sich dabei zunächst auf Elemente wie Titel, Jahr bzw. Datum, Sprache, Name, Objekttyp, Format und Thema, und legt, nach allgemeinen Bemerkungen über deren Verwendung, dar, was speziell in Dublin Core zu beachten ist. Gleichzeitig bezieht er die später eigens im Buch behandelten Schemata MODS und VRA zum Vergleich mit ein und zeigt anhand von Übersichtstabellen, welche Elemente sich in den drei Schemata jeweils entsprechen. Hier sind klar die unterschiedlichen Detailebenen zu erkennen — MODS bietet z.B. zur Darstellung des Titels vier Felder, während Dublin Core nur eines aufweist.

In Kapitel 5 geht es um kontrollierte Vokabularien wie Thesauri oder Taxonomien und deren Anwendung und Entwicklung.

Kapitel 6 beschäftigt sich mit XML als Kodierungsformat der Metadatenschemata. Selbstverständlich lässt sich XML nicht in einem Kapitel erschöpfend behandeln, aber Ziel ist es vielmehr, ein Verständnis dafür zu schaffen, wie Informationen kodiert und die behandelten Schemata in XML dargestellt werden, also XML „lesen“ und verstehen zu können. Sehr hilfreich auch hier die Aufstellungen mit detaillierter Erläuterung der einzelnen Elemente. Wer sich eingehender mit XML beschäftigen möchte, dem gibt die oben angesprochene Verlags-Website zusätzliche Übungen zum Einstieg an die Hand; z.B. kann man mit Hilfe eines XML-Editors MODS-Aufnahmen analysieren.

Thema von Kapitel 7 ist MODS (Metadata Object Description Schema). Es handelt sich um ein XML-basiertes, im Vergleich zum flachen Dublin Core ausdrucksstärkeres und komplexeres Schema. Tabellarisch werden Mappings zwischen DC und MODS abgebildet und Aufnahmen Seite an Seite gestellt, was die unterschiedliche Komplexität aufzeigt und Fragen der Interoperabilität beleuchtet. Hier wird deutlich, daß DC gute Dienste als kleinster gemeinsamer Nenner leisten kann, andere Schemata aber reicher in der Informationsmodellierung sind. Screenshots vermitteln einen Eindruck von verschiedenen MODS-XML-Editoren, die zur Erstellung von MODS im Einsatz sind.

Der Autor befasst sich in Kapitel 8 mit dem VRA (Visual Resources Association) Schema, das vor allem in Museen Anwendung findet und primär für die Beschreibung von Werken der bildenden Kunst und Architektur ausgelegt ist. Daher wird im VRA-Datenmodell zwischen „work“ (das eigentliche, „physische“ Werk) und „image“ (seine Repräsentation als Fotografie, Dia oder elektronische Datei) unterschieden und diese beiden Entitäten in einer Aufnahme durch das Element <imageOf> verknüpft. VRA wird allerdings nicht so ausführlich wie DC und MODS behandelt.

In den Kapiteln 9 bis 11 weitet sich die Perspektive weg von konkreten Schemata hin zu den konzeptuellen Gesichtspunkten von digitalen Sammlungen. Stichworte sind hier u.a. Metadaten-Qualität, Austausch, Harvesting via OAI-PMH und Integration in Portale oder andere aggregierte Dienste und Interoperabilität. Es empfiehlt sich, mögliche lokal definierte Felder auf Standardschemata wie Dublin Core abzubilden, um beim Harvesting Informationsverlust zu vermeiden, die Elemente korrekt anzuwenden und lokale Implementierungen zu dokumentieren.

Kapitel 10 macht die Planung und Entwicklung von Metadaten-Profilen und Anwendungs-Richtlinien („application profile“) zum Thema – einer Analyse der speziellen Sammlung und der sich darauf beziehenden Nutzerbedürfnisse schließen sich Fragen an wie z.B. welche Metadaten-Elemente benötigt werden und wie sie angewendet werden. Der Autor hat eine Auswahl an konkreten Beispielen von „application profiles“ zusammengestellt, um die einzelnen zu beachtenden Punkte zu verdeutlichen.

Im abschließenden Kapitel geht der Autor in einer einführenden Darstellung auf das Semantic Web, Linked Data und das Datenmodell RDF (Resource Description and Framework) ein. Dieser Abschnitt soll vor allem einen kursorischen Überblick über diese neueren Entwicklungen bieten, die möglicherweise zukünftig Einfluss auf die Metadaten-Praxis haben könnten.

Hier liegt in einem übersichtlichen Band all das Material zu Metadaten für digitale Bestände vor, das man sich bisher aus verschiedenen Quellen zusammensuchen mußte. Es finden sich sehr nützliche Definitionen von Begriffen, die zum Verständnis der (vorwiegend englischsprachigen) Fachliteratur zum Thema „digital collections“ nötig sind. Am Ende jedes Kapitels werden die zentralen Punkte in einer Zusammenfassung wiederholt.

Die meisten Beispiele beziehen sich auf digitalisierte Fotografien (die im Vergleich zu text-basiertem Material tatsächlich den größeren Teil in amerikanischen „digital collections“ ausmachen²) – wohl auch ein Grund, warum METS, ein Strukturstandard, der oft für digitalisierte Bücher genutzt wird, keine Erwähnung findet.

Das Layout des Buches mit breitem Rand ist sehr gut geeignet, um Notizen oder Bemerkungen anzubringen; es läßt sich sowohl als Arbeits- und Handbuch als auch als Nachschlagewerk ausgezeichnet verwenden. In grau unterlegten Randnotizen werden ausgewählte wichtige Informationen nochmals gesondert herausgestellt. Anschauliche Illustrationen und Tabellen tragen hervorragend zum Verständnis des Gesagten bei. Ein umfangreiches Literaturverzeichnis ergänzt dieses außerordentlich detailreiche, sehr gut durchdachte und thematisch umfassende Buch, das sich als fundierter Einstieg in die Materie eignet.

Saskia Breitling, Wien

- 1 Vgl. Greenstein, Daniel, Suzanne E. Thorin. *The Digital Library: A Biography*. Washington, D.C.: Digital Library Federation (2. ed., 2002). ISBN 1-887334-95-5. Online zugänglich: <http://www.clir.org/PUBS/reports/pub109/pub109.pdf>.
- 2 Vgl. ibd., S. 66.