

Daniel Dahl und Gottfried Vossen

Lernobjekt-Metadatenerstellung in Zeiten des Web 2.0

Learning Object Metadata Generation in the Web 2.0 Era

e-Learning_Lernobjekte_Metadaten_LOM_Web

Zusammenfassung. Die Ziele, die mit der Einführung des Metadatenstandards LOM verfolgt wurden (z. B. die Auffindbarkeit und die Wiederverwendbarkeit von Lernobjekten) werden aktuell trotz einer weit reichenden Verbreitung und Akzeptanz nur in geringem Maße erreicht: Aufgrund der großen Komplexität des Standards steht dem in der Theorie immensen Potential eine deutliche Zurückhaltung in der Praxis gegenüber. Dies lässt sich u. a. darauf zurückführen, dass der Prozess der Metadatenerstellung (wer erstellt wann wie welche Attribute) bislang nicht im Detail definiert ist. Dieses Papier zeigt einen Weg auf, der gleichzeitig eine hohe Quantität wie Qualität bei der Erstellung von Lernobjekt-Metadaten sicherstellt. Dafür werden bekannte Ansätze der Metadatenerstellung kombiniert und erstmals die Potentiale des im Kontext der Web 2.0-Entwicklungen vorbildlich umgesetzten User-zentrierten Paradigmas zur Community-getriebenen Metadatenerstellung von Lernobjekten genutzt. Am konkreten Beispiel des Einsatzes einer Lernplattform wird definiert, wer wann welche Metadatenattribute des Standards auf welche Art und Weise füllt.

Summary. When introducing the metadata standard LOM, objectives such as the ability to find or reuse learning objects were followed. These objectives are actually achieved in LOM only to a limited degree, in spite of the designation as a de-facto standard for the description of electronic learning content: Based on the complexity of the standard, a high theoretical potential faces rejection in practice. One reason for this is that the process of metadata generation – for example, who creates which metadata attributes – is not defined in detail yet. This paper illustrates an approach which guarantees a high quantity as well as a high quality of learning object metadata records, bringing together known ways of metadata creation and the new paradigm of users describing content as implemented in recent Web 2.0 applications. In the context of a concrete e-learning platform we exemplarily define who creates which metadata records of LOM in which way at what time.

1. Einleitung

Elektronisches Lernen findet heute insbesondere als Blended Learning eine immer weiter reichende Verbreitung. Durch die Umsetzung des Konzeptes der Lernobjekte (Wiley 2002) soll in Folge der Einführung von Metadaten (Daten über Daten) die Auffindbarkeit und die Wiederverwendbarkeit von einzelnen Lerneinheiten sichergestellt bzw. verbessert werden. Als de facto Standard im Bereich der Lernobjekt-Metadaten kann aufgrund seiner weiten Verbreitung das IEEE Learning Object Metadata (LOM, siehe <http://ltsc.ieee.org/wg12/20020612-Final-LOM-Draft.html>) bezeichnet werden, das mit seinen über 40, in neun Subkategorien aufgeteilten Metadatenfeldern eine umfassende Auszeichnung von Lernobjekten ermöglicht. An zentraler Stelle werden die Metadaten gesammelt und in e-Learning Repositories (Lernobjekte und Metadaten) bzw. e-Learning Metadaten Repositories (Metadaten und Verweise auf Lernobjekte) potentiellen Nutzern zugänglich gemacht, wodurch eine Transparenz und Integration bestehender Lernangebote in unterschiedlichen Kontexten erzielt wird (Dahl und Vossen 2007). Während die über 40 Attribute des LOM in der Theorie eine äußerst detaillierte Beschreibung ei-

nes Lernobjekts ermöglichen, findet eine umfassende Auszeichnung von Inhalten in der Praxis de facto nur selten statt. Untersuchungen zeigen, dass naheliegende Attribute wie „Titel“, „Format“ oder „Sprache“ relativ häufig gefüllt werden, während Felder wie „Schwierigkeit“ oder „Struktur des Lernobjektes“ wenig bis keine Beachtung finden (Friesen 2004). Solange die Metadaten lediglich in einem Kontext bzw. innerhalb eines Systems von Interesse sind, kann eine gezielte Reduzierung des Attributumfanges sogar sinnvoll sein, da eine Ausrichtung auf den Endnutzer vorgenommen und die Komplexität deutlich reduziert werden kann (Dahl, Vossen und Westerkamp 2006).

nes Lernobjekts ermöglichen, findet eine umfassende Auszeichnung von Inhalten in der Praxis de facto nur selten statt. Untersuchungen zeigen, dass naheliegende Attribute wie „Titel“, „Format“ oder „Sprache“ relativ häufig gefüllt werden, während Felder wie „Schwierigkeit“ oder „Struktur des Lernobjektes“ wenig bis keine Beachtung finden (Friesen 2004). Solange die Metadaten lediglich in einem Kontext bzw. innerhalb eines Systems von Interesse sind, kann eine gezielte Reduzierung des Attributumfanges sogar sinnvoll sein, da eine Ausrichtung auf den Endnutzer vorgenommen und die Komplexität deutlich reduziert werden kann (Dahl, Vossen und Westerkamp 2006).

Probleme treten jedoch auf, wenn Repositories miteinander kommunizieren sollen und Suchanfragen in verteilten Verzeichnissen durchgeführt werden: Während auf der einen Seite ein Metadatenattribut als wichtig erachtet wird und z. B. zentraler Bestandteil einer Suchmaske ist, wird es in anderen Repositories unter Umständen nie verwendet. Bei einer geringen Überschneidung der jeweils gefüllten Metadatenfelder sind die ursprünglichen Ziele wie die Auffindbarkeit und die darauf aufbauende Wiederverwendbarkeit von Lernobjekten nicht mehr zu erreichen. Zudem besteht bei einer großen Menge an zu füllenden Metadatenfeldern die Gefahr einer oberflächlichen Metadatenerstellung, in Folge dessen zwar eine hohe Quantität resultieren kann, Defizite in der Qualität jedoch wahrscheinlich sind. Zur Erreichung der systemübergreifenden Auffindbarkeit und Wiederverwendbarkeit von Lernobjekten muss also eine hohe Quantität einhergehend mit einer hohen Qualität der Lernobjekt-Metadaten sichergestellt sein; dies ist aktuell jedoch nur selten der Fall.

Das zentrale Dilemma der Lernobjekt-Metadatengenerierung ist also, dass auf der einen Seite mit LOM in der Theorie eine äußerst detaillierte Beschreibung von Lernobjekten möglich ist, diese aufgrund des großen Umfangs und der Komplexität des IEEE-Standards jedoch eine zu geringe praktische Umsetzung erfährt. Wir führen dies darauf zurück, dass bislang vier wichtige „W-Fragen“ nicht klar beantwortet sind:

Wer erstellt wann wie (auf welche Art) was (welche Metadatenattribute)?

Es erscheint nahe liegend, dass eine einzelne Person kaum alle Metadatenattribute in Form einer Liste mit leeren Textfeldern füllen kann. Trotzdem ist dies ein in der Praxis häufig zu beobachtendes Verfahren. Einbußen in Form von Qualitäts- und Quantitätsverlusten sind dabei vorprogrammiert.

Demgegenüber fassen wir die Metadatenerstellung als einen Prozess auf, im Rahmen dessen verschiedene Akteure interagieren und unterschiedliche Quellen für Metadaten herangezogen werden. Damit die eingangs angesprochenen Ziele erreicht werden können, muss allerdings genau definiert werden, welche Attribute zu welcher Zeit von wem auf welche Art und Weise generiert werden können. Nur so ist eine hohe Metadaten-

quantität einhergehend mit einer hohen Qualität zu erzielen.

Mit dem Ziel, den Prozess der Lernobjekt-Metadatengenerierung für den konkreten Lernkontext an einer Hochschule im Detail zu definieren, ist diese Arbeit wie folgt aufgebaut: Abschnitt 2 zeigt zunächst auf, welche Wege der Metadatenerstellung bislang generell verfolgt werden; zudem werden aus Web 2.0-Anwendungen bekannte Ansätze vorgestellt, in denen der Systembenutzer als Metadaten-Ersteller innerhalb einer Community in Erscheinung tritt. Abschnitt 3 führt zunächst die bekannten Wege der Metadatenerstellung in einem Modell zusammen und zieht erste Schlüsse bezüglich der Akteure im Erstellungsprozess bei Lernobjekten (Wer-Frage). Durch eine klassische Metadaten-Perspektive auf LOM, die sich von dem hierarchischen Ausgangsmodell löst, lassen sich Gruppierungen von LOM-Attributen identifizieren, die auf die gleiche Art generiert werden können (Was-Frage). Abschnitt 4 beschreibt in einem Real-Welt-Szenario durch den Einsatz einer Lerner-zentrierten Plattform, an welchen Stellen vor und während des Einsatzes von Lernobjekten Metadaten anfallen bzw. generiert werden können (Wann-/Wie-Frage). In einer Zusammenführung aller Erkenntnisse lassen sich die eingangs gestellten Fragen beantworten. Abschnitt 5 fasst die Arbeit zusammen und gibt einen Ausblick auf weitere Forschungsfragen.

2. Verwandte Arbeiten

Die Erstellung von Lernobjekt-Metadaten kann auf verschiedenen Wegen erfolgen, wobei jeder Weg individuelle Vor- und Nachteile mit sich bringt. Primär lassen sich zwei unterschiedliche Ansätze identifizieren: Ein Top-Down-Vorgehen zielt darauf ab, explizit Metadatenattribute, z. B. des LOM-Standards, zu füllen. Bei einem Bottom-Up-Vorgehen werden vorhandene Informationen (die möglicherweise in großen Mengen vorliegen) analysiert und durch Aufbereitung und Selektion zu Metadaten verarbeitet. Im Folgenden werden zunächst einige Top-Down-Vorgehen sowie ein Bottom-Up-Ansatz, aus dem Web 2.0-Umfeld vorgeht, dargestellt.

Naheliegender ist, die Aufgabe der Metadatenerstellung an eine Person, sei es

den Autor eines Lernobjekts oder einen explizit bestimmten Metadatenersteller, im besten Fall einen Experten der jeweiligen Lerndomäne, zu übertragen. Zur Unterstützung bieten sich so genannte LOM-Editoren (z. B. Cebeci und Erdogan 2005 oder <http://www.multibook.de/lom/de>) an, die das hierarchische Modell des Standards z. B. in Form eines Baumes abbilden. Über Freitextfelder und Drop-Down-Listen kann nach und nach die gesamte LOM-Komplexität genutzt und ein hochwertiger Metadatensatz angelegt werden. Allerdings ist es einer einzelnen Person alleine aus Zeit- und Kosten Gründen nahezu unmöglich, quantitativ wie qualitativ hochwertige Metadaten-sätze für eine größere Menge von Lernobjekten zu erstellen. Eine kooperative Form der Metadatenerstellung ist für LOM-Editoren in der Regel weder definiert noch durch Software unterstützt.

Ein anderer Ansatz nutzt die Unterstützung des Computers in einem sehr viel höheren Maß: Die automatische Metadatengenerierung extrahiert aus den Lernobjekten und deren Kontext für die Metadatenfelder nützliche Informationen (Cardiaels, Meire und Duval 2005). Während auf der einen Seite durch die Einsparung von Personal und Arbeitsaufwand Kosten gesenkt werden, muss auf der anderen Seite mit ungenauen bis fehlerhaften Metadatenattributen gerechnet werden. Während z. B. aus textlastigen Lernobjekten noch relativ zuverlässig nutzbare Metadaten extrahiert werden können, sieht dies bei audiovisuellen Inhalten schon anders aus. Aktuell ist bereits das Auslesen eines Lernobjekttitels z. B. aus einer Powerpoint-Folie eine nicht zu unterschätzende Herausforderung (sofern der Titel nicht bereits explizit als Metainformation eingetragen wurde), so dass der Ansatz der automatischen Metadatenerstellung als äußerst vielversprechend angesehen werden kann, der ausschließliche Einsatz ohne menschliche Interaktion jedoch als fehleranfällig zu bezeichnen ist.

Einen Mittelweg zwischen der automatischen und der manuellen Metadatenerstellung zeigen Hybride Systeme (Motolet und Baloian 2006) auf. Basierend auf einer automatischen Analyse des Lernobjekts werden z. B. drei Gruppen von Informationen gebildet: Sehr wahrscheinliche Werte, wahrscheinliche Werte und Restriktionen von möglichen Wer-

ten. Während erstere Gruppe, der das Format oder die Größe eines Lernobjektes angehören, in der Regel keine manuelle Kontrolle erfordern, wird für die zweite Gruppe eine Überarbeitung empfohlen. Die dritte Gruppe besteht aus Restriktionen, die das wohldefinierte Vokabular von Metadatenfeldern einschränkt, selbst jedoch keine expliziten Einträge vorgibt. Der Einsatz von Hybriden Systeme erleichtert die abschließend durch eine Person (Autor oder Experte) durchgeführte Metadatenerstellung ungemein, indem Aspekte der automatischen Generierung eingebracht werden. Was jedoch nach wie vor fehlt ist eine detaillierte Zuordnung, die einige der eingangs gestellten W-Fragen beantworten kann. Nachteilig ist ebenfalls, dass z. B. eine kollaborative Erstellung von Metadaten nicht unterstützt wird.

Einen kollaborativen Weg der Metadatenerstellung von Lernobjekten schlägt (Or-Bach 2005) vor: Mit dem pädagogischen Ziel der Reflexion von Lerninhalten erstellen Studenten computergestützt und zielgerichtet Metadaten. Durch das Gruppenbewusstsein wird ein begründendes Arbeiten erfordert. Am Beispiel von Programmier-Lerneinheiten wird deutlich, dass durch das Abstrahieren und Erkennen von Konzepten oder das Zusammenführen von Einzelteilen und Beispielen – was für die Erstellung hochwertiger Metadaten sinnvoll ist – ein Lernfortschritt erzielt werden kann. Dieser Ansatz setzt den Benutzer eines Systems, in diesem Fall den Lernenden, in das Zentrum der Metadatenerstellung. Allerdings wird nicht explizit das Ziel verfolgt, vollständige und hochwertige Metadaten für ein Lernobjekt im Sinne eines e-Learning Repositories (Auffindbarkeit, Wiederverwendbarkeit) zu gewinnen. Dies zeigt sich auch daran, dass der Ge-

danke der benutzergetriebenen Metadatenerstellung und der daraus resultierende Mehrwert nicht konsequent fortgeführt wurden: Es ist nicht klar, wie eine Aufgabenteilung von Autor bzw. Lehrendem und Lernenden aussehen kann.

Die Herausforderung, multimediale Inhalte auffindbar zu machen ist auch losgelöst vom Lernkontext von großem Interesse: Wie kann man beispielsweise allgemein Fotos kennzeichnen, so dass ein schnelles Wiederfinden sichergestellt ist? Im Zuge des Web 2.0-Phänomens (O'Reilly 2006, Vossen und Hagemann 2007) haben sich Folksonomies (Vanderwal 2006) (in Anlehnung an das „Folk“ als Ersteller und die klassische Taxonomie), basierend auf der Verschlagwortung von Inhalten als vielversprechend herausgestellt. Hier sind die Ersteller der Metadaten keine Experten (wie z. B. in Bibliothekssystemen) oder Autoren (wie z. B. im Sinne des Dublin Core Metadatenstandards). Das Erstellen von Metadaten durch Systembenutzer (Mathes 2006) kann zunächst individuelle Gründe, wie die Auffindbarkeit von interessanten Fotos, haben. Aus diesen persönlichen Metadaten lassen sich auch für die Gruppe der Benutzer interessante Informationen gewinnen, indem z. B. die Metadaten aller Anwender zusammengeführt oder Objekte kollaborativ verschlagwortet werden. Unter bestimmten Voraussetzungen nimmt also der Benutzer eines Systems die Rolle des Metadatenerstellers ein, ohne dass ihm dies während seines Handelns bewusst ist. Während bei der kollaborativen Metadatenerstellung jederzeit bekannt ist, dass zielgerichtet Metadatenfelder gefüllt werden, sieht der Web 2.0-Benutzer zunächst individuelle Vorteile in der Anreicherung von Inhalten. Zu beachten ist hier, dass die Metadaten nicht von Experten stammen und

somit die Qualitätsfrage gestellt werden darf.

Zusammenfassend halten wir fest, dass unterschiedliche Wege der Metadatenerstellung existieren, die individuelle Vorteile wie Nachteile mit sich bringen. Der Bottom-Up-Ansatz besitzt auch im Kontext des Lernens und der Erstellung von Lernobjekt-Metadaten ein enormes Potential, das bislang noch nicht in Kombination mit bekannten Top-Down-Ansätzen genutzt wurde. Im Folgenden soll daher das Ziel sein, die einzelnen Ansätze genauer zu betrachten und zunächst ein umfassendes Modell der Lernobjekt-Metatdatenerstellung zu beschreiben. Im Anschluss daran werden gezielt Ansätze kombiniert, wodurch einzelne Vorteile hervorgehoben und Nachteile aufgehoben werden.

3. LOM-Analyse aus einer klassischen Metadaten-Perspektive

In Abschnitt 2 wurden Ansätze aus unterschiedlichen Szenarien der Metadatenerstellung vorgestellt. Diese sollen nun zunächst zusammengeführt und in Form eines allgemeinen Modells veranschaulicht werden. Bild 1 führt die im vorangegangenen Abschnitt vorgestellten Ansätze zusammen.

Auf erster Ebene wird die klassische Unterteilung in automatisierte (System) und manuelle Metadatenerstellung vorgenommen (siehe u. a. Gill, Gilliland-Swetland und Baca 1998). Zwischen diesen beiden Kategorien vermitteln die in Abschnitt 2 genannten Hybriden Systeme.

Die automatische Metadatengenerierung bei Lernobjekten kann in die Aspekte Inhaltsanalyse, Kontextanalyse, Nutzungsanalyse und Strukturanalyse unterteilt werden (Cardinaels, Meire und Duval 2005). Die Metadatenerstellung durch Menschen lässt sich zunächst klassischerweise in Experten, oftmals professionelle Metadatenersteller, und Autoren unterteilen (Greenberg et al. 2001). Hinzu kommt aus unserer Sicht ein weiterer Akteur: der Benutzer selbst. In Abschnitt 2 wurde das Beispiel der Studenten, die aus pädagogischen Gründen Metadaten für Lerninhalte erstellen, genannt. Hier ist eine kollaborative, explizite Herangehensweise festzustellen. Auf der anderen Seite der benutzergetriebenen Metada-

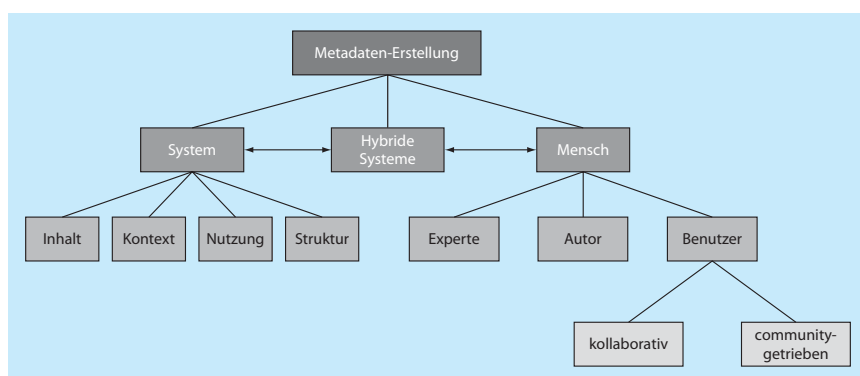


Bild 1: Wege der Metadatenerstellung

tenerstellung lassen sich die aus Web 2.0-Anwendungen bekannten Ansätze einordnen. Wir sprechen hier von einer Community-getriebenen Metadatenerstellung, die im Speziellen im Kontext e-Learning bereits von (Dahl und Vossen 2007) diskutiert wurde.

Welche der hier aufgezeigten Wege sind nun für die Erstellung von Lernobjekt-Metadaten interessant? Unter dem Aspekt der Qualität von Metadaten nehmen (Lux, Klieber und Granitzer 2006) eine Unterscheidung in manuell generierte High Level Metadata und automatisch generierte Low Level Metadata vor. Die unscharfe Grenze ist durch das so genannte Semantic Gap definiert; ohne menschliche Interaktion ist die Überbrückung des selbigen und in der Folge die Erstellung hochwertiger Metadaten, sofern überhaupt möglich, mit immensen Aufwand verbunden. Diese Sichtweise, die, wie eingangs beschrieben, durch aktuelle Probleme der automatischen Metadatenerstellung in der Praxis untermauert wird, lässt den Schluss zu, den automatischen Weg in Kombination mit dem manuellen zu verfolgen. Die vier automatischen Analysetechniken können genutzt werden, um wichtige Informationen über das Lernobjekt zu gewinnen. Im Bereich der menschlichen Metadaten-Erstellung muss im Lernkontext die Rolle des Experten genauer betrachtet werden. In einer Bibliothek z. B. kann, im Gegensatz zu einer Hochschule, diese Rolle explizit besetzt sein, d. h. eine Person ist darauf geschult, Objekte in vorgegebene Kategorien (z. B. in das Dewey Dezimal System) einzuordnen. An einer Universität kann aufgrund seines Fachwissens der Lehrende diese Aufgabe übernehmen. Da dieser in der Regel ebenfalls Autor bzw. Verantwortlicher der Lernobjekte ist, kann von einem Verschmelzen der beiden aus Bild 1 bekannten Rollen Experte und Autor gesprochen werden. Zusätzlich nimmt der Lehrende die Rolle eines Beobachters bzw. eines Begleiters des Lernprozesses ein: Ggfs. können Hinweise gegeben oder bestehendes Material modifiziert und ergänzt werden. Benutzer eines Lernobjektes sind allgemein als Lernende zu bezeichnen; dies können Studierende wie auch sich fortbildende Dozenten sein. Der kollaborative Ansatz ist hier eher als Spezialfall zu sehen: Insbesondere unter dem Aspekt der Motivation, die für die Community-getriebene

Alternative später anhand eines realen Szenarios detailliert begründet wird, sind Nachteile zu identifizieren. Während der kollaborative Ansatz in erster Linie externen Druck, z. B. durch den Lehrenden, der eine gute Note vergibt, voraussetzt (extrinsische Motivation), basiert der Community-Ansatz auf innerer, dem eigenem Antrieb begründeter Motivation (intrinsische Motivation): Multimediale Inhalte werden nach individuellen Vorstellungen kategorisiert und auffindbar gemacht.

Desweiteren ist zu beantworten, wie eine nun eine detaillierte Zuordnung von einzelnen Metadatenattributen zu den identifizierten Erstellungswegen aussehen kann. Eine Sicht auf LOM, die sich losgelöst vom hierarchischen Modell und den neun Unterkategorien an einer klassischen Kategorisierung von Metadaten orientiert, soll erste Aufschlüsse auf Gruppierungen von Metadaten, die auf gleichem Wege erstellt werden können, liefern.

Greenberg (2001) untersucht Metadaten-Schemata hinsichtlich einer Zuordnung ihrer Attribute in unterschiedliche Metadatenklassen. Unter anderem werden die Attribute des Dublin Core-Standards (DC) zugeordnet. DC ist sehr viel allgemeiner gehalten als LOM und erlaubt die Beschreibung beliebiger digitaler Objekte. Da in der Praxis auch Lernobjekte mit DC ausgezeichnet werden, liegt es nahe, das verwendete Klassifizierungsschema zu übernehmen und auf LOM anzuwenden. Im Detail bedeutet dies, dass die einzelnen LOM-Attribute untersucht werden und eine Einordnung in

- Discovery Metadata
- Use Metadata
- Authentication Metadata
- Administration Metadata

vorgenommen wird. Discovery Metadata umfasst alle Attribute, die der Auffindbarkeit des Lernobjektes bei einer Suche dienen. Zwecks Einschränkung werden im Folgenden Attribute als Discovery Metadata bezeichnet, sofern diese allein – ohne Kombination mit weiteren Daten – ein Auffinden (z. B. über eine Volltextsuche) ermöglichen. Als Use Metadata werden alle Attribute bezeichnet, die während des Einsatzes des Lernobjektes von Bedeutung sind. Dies umfasst technische Informationen wie das Format oder die Systemvoraussetzungen ebenso wie intellektuelle Ausprägungen wie die

Eigentumsrechte oder Restriktion bezüglich der Nutzung. Unter Authentication Metadata werden Attribute verstanden, welche die Integrität und das allgemeine Vertrauen in ein Lernobjekt sicherstellen. Administration Metadata umfasst Metadaten, die das Management eines Objektes unterstützen.

Um eine feiner granulare Unterteilung zu erwirken, wird zusätzlich jede der vier Metadatenklassen in subjektive und objektive Metadaten (Duval et al. 2002) unterteilt. Unter objektiven Metadaten werden Fakten wie z. B. das Erstellungsdatum oder die aktuelle Versionsnummer verstanden. Subjektive Metadaten können dagegen variieren und hängen stark von der Sicht des Metadatenerstellers ab. Zusammenfassungen oder Stichwortlisten können sich z. B. wenig bis sehr stark unterscheiden.

Insgesamt ergeben sich also 8 Kategorien, denen LOM-Attribute zugeordnet werden. Betrachtet werden dabei die 45 Elemente unterhalb der LOM Hauptkategorien, d. h. es wird z. B. nicht mehr zwischen Elementen wie LOM:Metametadata:Contribute:Role und LOM:Metametadata:Contribute:Date unterschieden, sondern zusammenfassend von LOM:Metametadata:Contribute gesprochen. Jedes der Attribute wird mindestens einer und maximal vier Kategorien zugeordnet (eine gleichzeitige Zuordnung in subjektiv und objektiv wird ausgeschlossen).

Tabelle 1 zeigt die Zuordnung der LOM-Attribute. Die Einteilung basiert zum einen auf unterschiedlichen Quellen wie z. B. (Greenberg 2001) oder expliziten Vorschlägen von (Duval et al. 2002) oder (Cardinaels, Meire und Duval 2005), zum anderen auf einer eigenständigen Analyse der Autoren; eine Diskussion jeder einzelnen Zuordnung würde allerdings den Rahmen dieses Beitrags sprengen. Es sei bemerkt, dass dies lediglich eine mögliche Zuordnung darstellt; sicherlich lässt sich über die Einteilung einzelner Attribute eine umfassende Diskussion führen. Ein Ziel war hier, einen Mittelweg zwischen der Entlastung von Lehrenden und Lernenden und der Zuverlässigkeit der automatisch erstellten Daten zu finden.

Zusammenfassend lassen sich erste Rückschlüsse ziehen, welche Gruppen von Attributen von wem erstellt werden. Objektive, der Discovery dienende Metadaten bieten sich für die Erstellung durch

Tabelle 1: Alternative Sicht auf LOM

	Discovery	Use	Authentication	Administration
objektiv	General:Title General:Identifier General:Coverage Lifecycle:Contribute	Rights:Copyright Rights:Costs Edu:Res.Type Technical:Duration Technical:Size Technical:Platf.Req. Technical:Format General:Aggr.Level General:Language General:Structure	Rights:Description Edu:Inter.Type Edu:Inter.Level Technical:Location Technical:Req. Technical:In.Remarks	Lifecycle:Status Lifecycle:Version Lifecycle:Contribute Annotation:Entity Annotation:Date Lifecycle:Status Lifecycle:Version Lifecycle:Contribute Rights:Costs Rights:Description Rights:Copyright Metameta:Contribute Metameta:Language Metameta:Identifier
subjektiv	General:Description General:Keyword Edu:Description Class.TaxonPath Class.Keyword Class.Purpose Class.Description Annotation:Description Relation:Kind Relation:Resource	Edu:AgeRange Edu:LearningTime Edu:Language Edu:EndUserRole Edu:Context Edu:Sem.Density Edu:Difficulty Edu:Description	Relation:Kind Relation:Resource Annotation:Description	

einen Lehrenden an, da diese eindeutig (bezogen auf den Inhalt) gefüllt werden können (z. B. Titel oder Coverage). Im Gegensatz dazu lassen die subjektiven Discovery-Metadaten auf individuelle Inhalte schließen: Von Person zu Person kann es variieren, in welcher Rangfolge Hierarchien aufgebaut sind, welche Schlagwörter das Lernobjekt auszeichnen oder welche verwandten Objekte in Bezug gesetzt werden. Entsprechend bieten sich diese Attribute für die Erstellung durch den Lernenden an. Die große Gruppe der Use-Metadaten scheint für die automatische Erstellung sehr gut geeignet. Während objektive Use-Metadaten eher durch Inhalts- Kontext- und Struktur-Analysen generierbar scheinen, ist für die subjektiven Attribute eine Usage-Analyse angebracht. Objektive Authentication- und Administration-Metadaten können teilweise mit hoher Wahrscheinlichkeit korrekt automatisch generiert werden (z. B. alle MetaMetadata-Attribute), zum Teil ist jedoch die Interaktion mit dem Lehrenden unabdingbar. So können die Kosten eines Lernobjekts oder das Copyright als Vorschlag generiert werden, indem in Bezug stehende Lernobjekte (z. B. gleicher Autor, Ober-Unterbeziehungen der Lernobjekte) analysiert werden. Hier lässt sich natürlich der Bezug zu verwandten Lernobjekten auswerten. Eine abschließende Beurteilung sollte jedoch im semi-automatischen Sinne durch den Lehrenden erfolgen, da nur dieser über das explizite Wissen ver-

fügt. Die subjektiven Authentication-Attribute scheinen ebenso wie subjektive Discovery-Elemente für die Erstellung durch Lernende prädestiniert.

4. Metadaten-Erstellung in einer Lerner-zentrierten e-Learning Plattform

Die Erkenntnisse aus Abschnitt 3 sollen nun in einem praktischen Szenario zusammengeführt werden. Die an der Universität Münster in der Entwicklung befindliche Plattform Learnr (<http://learnr.uni-muenster.de>) ermöglicht Lernenden neben einem zentralen Zugriff auf Lernobjekte die Annotation von Inhalten im Stile von Social Bookmarking oder Tagging nach dem Delicious- oder Flickr-Vorbild, wie es im Kontext des Web 2.0 unter dem Stichwort „Sozialisierung von Inhalten“ sowie „Benutzer-generierte Inhalte“ diskutiert wird. Aus der realen Welt bekannte Vorgänge etwa während einer Klausurvorbereitung werden in Learnr digital unterstützt. So können Lernende digitale Anmerkungen in Form von Postits an Lernobjekten positionieren, Tags zwecks Auffindbarkeit und Kategorisierung anbringen, Beziehungen zu anderen Inhalten herstellen oder Zusammenfassungen und elektronische Karteikarten für das Wiederholen wichtiger Aspekte erstellen.

Bild 2 zeigt in Form einer ereignis-orientierten Prozesskette, wie im Kontext

der beschriebenen Lernplattform Metadaten durch eine Zusammenführung von Top-Down- und Bottom-Up-Ansätzen entstehen. Initial werden in einer ersten Phase gezielt Metadaten durch den Lehrenden über einen auf relevante Attribute reduzierten LOM-Editor (z. B. ähnlich <http://shareloc.uni-muenster.de>) angegeben. Mit dem Einstellen des Lernobjektes können zudem Algorithmen im Kontext der Plattform durch eine Inhalts-, Kontext- und Strukturanalyse eine beachtliche Menge an Attributen automatisch berechnen.

Beginnend mit der eigentlichen Nutzung des Lernobjektes können zum Einen mit Hilfe von Algorithmen automatisiert Metadaten durch eine Analyse des Nutzungsverhaltens gewonnen werden. Zum Anderen beginnen die Lernenden, die in der Plattform angebotenen Lernobjekte individuell zu behandeln: Analog zu Flickr (<http://www.flickr.com>) werden Tags angebracht, die einerseits die Auffindbarkeit vereinfachen und andererseits eine flache Kategorisierung ermöglichen. Während in Flickr Stichwörter wie Feuer-ventura Bilder auszeichnen, sind für Lerninhalte einordnende Schlüsselwörter sinnvoll, die z. B. beschreibend (SQL, Anfragesprachen, XQuery, XPath,...) oder bewertend (wichtig, nicht verstanden, klausurrelevant,...) sein können. Strukturen werden analog zu RawSugar (<http://www.rawsugar.com>) definiert, indem Tags in Beziehung zueinander gebracht werden. So kann das Schlüsselwort Da-

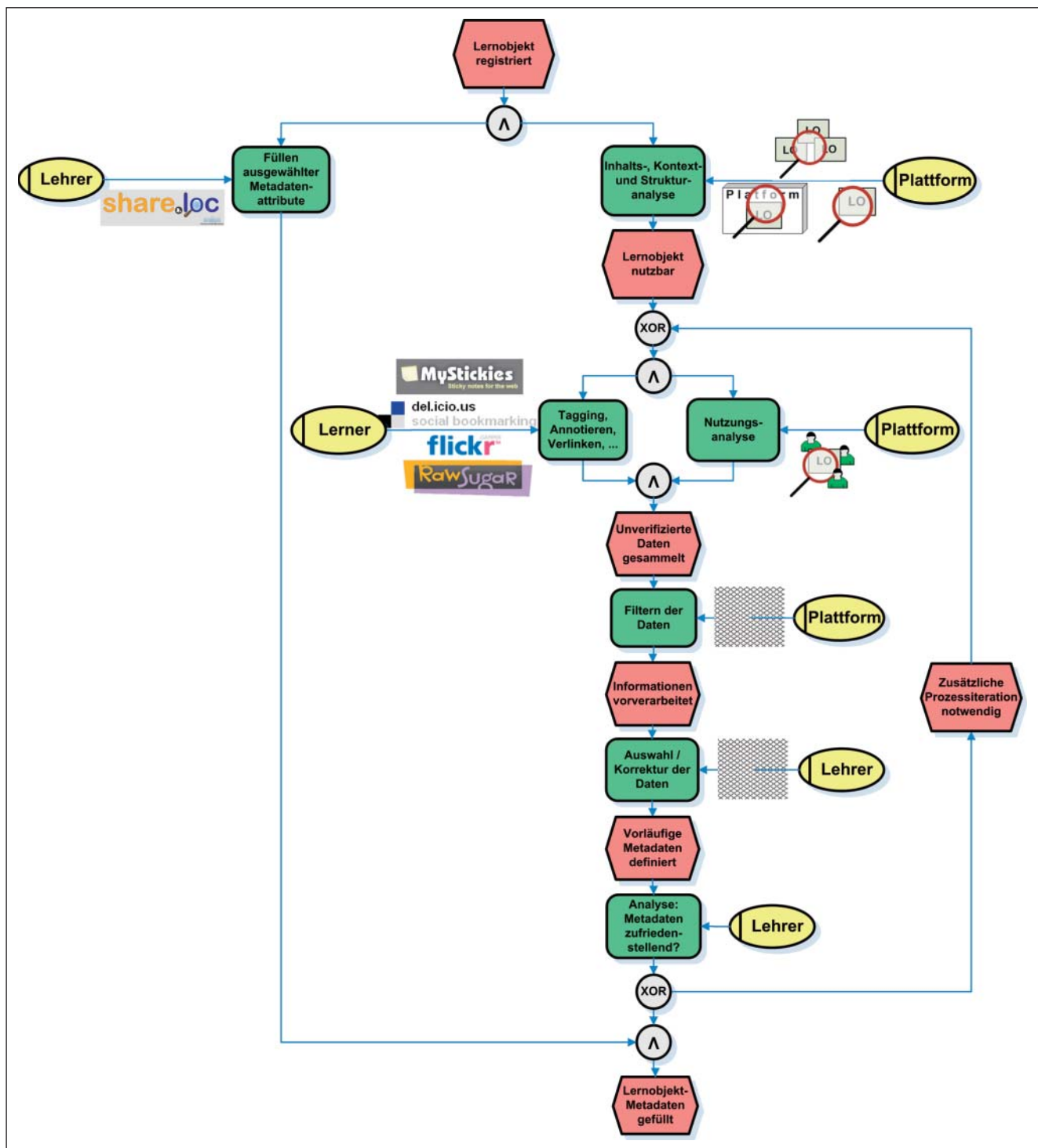


Bild 2: Prozess der Metadatenerstellung im Kontext einer e-Learning Plattform

tenbanken als Oberbegriff für Anfragen definiert werden; darunter finden sich schließlich SQL, XQuery und XPath. In der Folge kann der Lernende die eigens aufgebauten Strukturen intuitiv durchlaufen. Wie von Delicious (<http://www.del.icio.us>) bekannt werden Verweise auf andere Lernobjekte abgelegt und ausgetauscht. Die Verweise können sich auf interne, wie z. B. andere in Learnr

bereitgestellte Vorlesungsfolien, oder externe Inhalte, wie bspw. Einträge in Wikipedia, beziehen. Auf digitalen Post-Its, wie z. B. bei MyStickies (<http://www.mystickies.com>) zu sehen, werden Kommentare oder Zusammenfassungen verfasst. In der Folge dieser Benutzeraktionen fällt im Sinne eines Bottom-Up-Ansatzes eine reichhaltige Menge potentieller Metadaten an; insbesondere, wenn

die durch die einzelnen Nutzer angesammelten Informationen auf geeignete Weise zusammengeführt werden („wisdom of crowds“).

Aufgabe der Lernplattform ist es nun, alle angefallenen Metadaten aufzubereiten. Zuerst werden die von Lernerseite angebrachten Daten mit Data-Mining-Techniken analysiert und selektiert (erster Filter). Relevante Informationen werden

Tabelle 2: Definition der Lernobjekt-Metadatenerstellung

Wer	Wann	Wie	LOM-Attribute	
Lehrer	initial	LOM-Editor	General:Title Lifecycle:Contribute Lifecycle:Status Rights:Description Rights:Copyright	General:Coverage Lifecycle:Version Educational:Description Rights:Costs
Plattform	initial	Strukturanalyse	Relation:Kind Educational:Description Rights:Copyright	Relation:Resource Rights:Costs Rights:Description
		Inhaltsanalyse	General:Language Technical:Size Technical:Requirement Technical:PlatformRequirements Educational:InteractivityType Educational:ResourceType	Technical:Format Technical:Location Technical:InstallationRemarks Technical:Duration Educational:InteractivityLevel
		Kontextanalyse	General:Identifier General:AggregationLevel Lifecycle:Date MetaMetadata:Contribute MetaMetadata:Language	General:Structure Lifecycle:Contribute MetaMetadata:Identifier MetaMetadata:Schema Educational:Context
Lerner	Nutzung	Tags, Annotationen, Links, etc.	General:Keywords General:Description Annotation:Description Classification:TaxonPath Classification:Description	Relation:Resource Educational:Difficulty Classification:Purpose Classification:Keyword
Plattform	Nutzung	Kontextanalyse	Annotation:Entity	Annotation:Date
		Nutzungsanalyse	Educational:SemanticDensity Educational:AgeRange Educational:Language	Educational:EndUserRole Educational:LearningTime

dem Lehrenden in einem Dialogfenster (zusammen mit den automatisch berechneten Daten) präsentiert. Der Lehrende entscheidet nun, welche Daten von Interesse sind und in der Folge in die Lernobjekt-Metadaten einfließen (zweiter Filter).

Beispiel:

20 Lernende zeichnen ein Lernobjekt mit jeweils 5 Schlagwörtern aus. Von diesen 5 Begriffen sind bei 10 Lernern 3 identisch, die restlichen unterscheiden sich in allen Fällen. Die Plattform berechnet die Menge der gemeinsam genutzten Schlagwörter und präsentiert diese dem Lehrenden. Dieser entscheidet nun, dass zwei der drei häufig genutzten Schlagwörter das Lernobjekt sehr genau beschreiben, das dritte jedoch nicht angebracht ist. Folglich selektiert er zwei Tags, welche daraufhin als General:Keyword in den repräsentativen LOM-Datensatz aufgenommen werden.

Da mit der Zeit sowohl durch automatische Analysen (Kontext, Nutzung) wie auch durch die Lernenden weitere wichtige Informationen anfallen, ist das Aufbereiten der Daten durch die Plattform und

das Selektieren durch den Lehrenden keine einmalige Aktion, sondern ein wiederkehrender, iterativer Prozess. Als geeigneter Zeitrahmen bietet sich z. B. im Kontext der Hochschule ein Semester an, in dessen Ablauf eine konsequente Steigerung der Metadatenqualität zu erwarten ist.

Tabelle 2 fasst die bisherigen Erkenntnisse zusammen und beantwortet explizit die eingangs gestellten Fragen. Der Lehrer (vgl. objektive Discovery-Metadaten in Tabelle 1) setzt initial einen auf die relevanten Attribute reduzierten LOM-Editor ein und nutzt sein Expertenwissen, um zunächst den Titel des Lernobjekts und die Abdeckung anzulegen. Ebenso legt er die Attribute Ersteller, Version und Status an. Die pädagogische Beschreibung liegt in dem oben beschriebenen Szenario in der Vorlesungsbegleitung im Sinne des Blended Learning, kann aber durchaus variieren (z. B. Selbstlerneinheiten im Distance Learning). Die Rechte werden ebenfalls vom Lehrenden angegeben.

Natürlich kann zu diesem Zeitpunkt die Plattform bereits unterstützend eingreifen (vgl. objektive Use-Metadaten in

Tabelle 1). Durch eine Strukturanalyse kann z. B. initial beim Einstellen eines Kurses überprüft werden, ob hierarchisch ähnlich platzierte oder vom selben Lehrer stammende Lerneinheiten existieren. Entsprechend werden im LOM-Editor Vorschläge, abgeleitet aus den gewonnenen Informationen, gemacht (Rights: Costs, Educational:Description). Die pädagogische Beschreibung ist z. B. zu einer sehr hohen Wahrscheinlichkeit bei einem Kurs und den einzelnen Untereinheiten identisch. Zugleich wird das Lernobjekt in Beziehung zu bereits existierenden Lerneinheiten gesetzt (z. B. isPartOf-Beziehung). Die Inhaltsanalyse liefert vornehmlich technische Informationen wie z. B. das Format oder die Größe des Lernobjektes. Auch die verwendete Sprache, der Ressourcen-Typ (z. B. Folie, Test), Interaktionstyp (aktiv/expositiv) und -level werden mit großer Wahrscheinlichkeit korrekt erkannt. Über eine Kontextanalyse kann z. B. ein geeigneter Bezeichner generiert oder Informationen aus Nutzerprofilen extrahiert werden. So lassen sich Rückschlüsse auf den Veröffentlichenden und den Publikationszeitpunkt, den pädagogischen Kontext oder die MetaMetadaten im Allgemeinen ziehen.

Wie bereits beschrieben reichern die Lernenden das Lernobjekt mit zusätzlichen Informationen an (vgl. subjektive Discovery-Metadaten in Tabelle 1). Aus den Tags (Schlagwörtern) lassen sich z. B. Stichwortlisten extrahieren, Bookmarks (Verweise) auf interne (andere Lernobjekte) oder externe Quellen (z. B. Wikipedia) liefern Beziehungen, und Annotationen (Notizen) auf virtuellen Post-Its lassen sich als LOM-Annotations ablegen. Durch geeignete Algorithmen und das Zusammenführen von individuellen Tag-Hierarchien lassen sich zudem Kategorisierungen ableiten, die über flache Folksonomies hinausgehen. Auch eine allgemeine Beschreibung des Lernobjektes kann aus den vorhandenen Informationen geniert werden.

Während des Einsatzes des Lernobjekts liefert die Kontextanalyse Informationen über Annotationsersteller (Profilanalysen) und den Annotationszeitpunkt. Eine Usage-Analyse unter Einbeziehung der Userprofile lässt Rückschlüsse auf die Enduser-Rolle (z. B. Lerner oder Lehrer), den Age-Range oder die LearningTime zu: wie lange wurde das Lernobjekt betrachtet, bis es z. B. als „verstanden“ markiert wurde und

somit für den weiteren Lernprozess eine niedrigere Priorität bekommt (vgl. subjektive Use-Metadaten in Tabelle 1).

5. Zusammenfassung

Der in diesem Papier vorgestellte Weg der Lernobjekt-Metadatenerstellung stellt ein hohes Maß an Metadatenquantität und -qualität sicher, indem die kritische Frage

Wer erstellt wann wie (auf welche Art) was (welche Metadaten)?

beantwortet wird. Dazu wurden zunächst mögliche Wege der Metadatenerstellung aufgezeigt. Ein Aufbrechen des neunelementigen LOM-Standards und die Zuordnung zu klassischen Metadatenklassen liefern erste Hinweise auf Gruppierungen von Metadatenattributen, die den möglichen Wegen des Erstellungsprozesses zugeordnet werden. Das Beispiel einer lerner-zentrierten e-Learning Plattform stellt den zeitlichen Ablauf und die beteiligten Akteure sowie deren Handlungen heraus. Hervorzuheben ist die Rolle des Lehrenden, der durch die endgültige Entscheidung, welche Informationen in die Beschreibung des Lernobjektes einfließen, im Sinne einer Qualitätssicherung agiert.

In zukünftigen Arbeiten wird es uns darum gehen, das hier vorgestellte Modell im Detail zu realisieren. Es gilt zunächst, automatische Lernobjektanalysen etwa über Web-Services lose gekoppelt anzubinden und die in der Lern-Community anfallenden Metadaten geeignet aufzubereiten. In der Folge wird die Frage zu beantworten sein, ob die Metadaten in LOM abgelegt werden können oder ob der Standard aufgrund der Fülle an unterschiedlichen Metadaten erweitert wird. Es stellt sich zudem die Frage, ob neben einem repräsentativen LOM-Datensatz zusätzlich eine Art Personalized LOM sinnvoll ist; dies könnte für jeden Systembenutzer ein individueller LOM Datensatz sein, der sich aus automatisiert berechneten und eigens angegebenen Metadaten (Tags, Notizen, etc.) zusammensetzt. Für den einzelnen würde dies u. a. ein hohes Maß an (individueller) Auffindbarkeit garantieren. In der Folge muss zudem definiert werden, in welcher Form externe Systeme und Anwendungen auf die hochwertigen Metadaten zugreifen können. Denkbar ist eine Harvesting-Varian-

te, in der alle erstellten Metadaten in ein Repository überführt werden. Ebenso könnte die in diesem Papier angesprochene e-Learning Plattform eine geeignete Schnittstelle, z. B. SQL zur Kommunikation mit anderen Repositories, zur Verfügung stellen. Meta-Repositories könnten so dynamisch und on-demand nach passenden Lernobjekten suchen.

Literatur

- Cardinaels, K.; Meire, M.; Duval, E.: Automating Metadata Generation: the Simple Indexing Interface. Proc. 14th International Conference on the World Wide Web. ACM Press (2005) 548–556.
- Cebeci, Z.; Erdogan, Y.: Tree View Editing Learning Object Metadata. *Interdisciplinary Journal of Knowledge and Learning Objects* 1 (2005) 99–108.
- Dahl, D.; Vossen, G.: share.loc: Transparenz und Integration von Lernobjekten an der Universität Münster. Arbeitsbericht des E-Learning Kompetenzzentrums Münster, 2007.
- Dahl, D.; Vossen, G.; Westerkamp, P.: share.loc – A Multi-Tiered Interoperable E-Learning Metadata Repository. Proc. 6th IEEE International Conference on Advanced Learning Technologies (ICALT). Kerkraide, S. 891–895, 2006.
- Dahl, D.; Vossen, G.: Added Value for e-Learning Repositories through Social Tagging of Learning Objects. Datenbanksysteme in Business, Technology und Web (BTW). Workshop Proceedings, S. 270–279, 2007.
- Duval, E.; Hodgins, W.; Sutton, S.; Weibel, S.L.: Metadata Principles and Practicalities. *D-Lib Magazine* 8/4, April 2002.
- Friesen, N.: International LOM Survey: Final Report, <http://jtc1sc36.org/doc/36N0871.pdf>, 2004 (Letzter Zugriff: 13.04.2007).
- Gill, T.; Gilliland-Swetland, A.; Baca, M.: *Introduction to Metadata: Pathways to Digital Information*, Getty Research Institute, 1998.
- Greenberg, J.: A Quantitative Categorical Analysis of Metadata Elements in Image-Applicable Metadata Schemas, *Journal of the American Society for Information Science and Technology* (2001) 917–924.
- Greenberg, J.; Pattuelli, M.C.; Parsia, B.; Davenport Robertson, W.: Author-generated Dublin Core Metadata for web Resources: A Baseline Study in an Organization. *Journal of Digital Information* 2/2 (2001).
- Lux, M.; Klieber, W.; Granitzer, M.: On the Complexity of High Level Metadata. *Journal of Universal Knowledge Management* 1 (2006) 54–58.
- Mathes, A.: Folksonomies – Cooperative Classification and Communication Through Shared Metadata, <http://www.adammathes.com/academic/computer-mediated-communication/folksonomies.html>, 2004 (Letzter Zugriff: 13.04.2007).
- Motolet, O.; Baloian, N.: Hybrid Systems for Generating Learning Object Metadata. Proc. 6th IEEE International Conference on Advanced Learning Technologies (ICALT), Kerkraide, S. 563–576, 2006.

O'Reilly, T.: What is Web 2.0. <http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html>, 2005 (Letzter Zugriff: 13.04.2007).

Or-Bach, R.: Educational Benefits of Metadata Creation by Students, *inroads – The SIGCSE Bulletin*, Volume 37, Number 4, S. 93–97, 2005.

Vanderwal, T.: Explaining and Showing Broad and Narrow Folksonomies, <http://www.vanderwal.net/random/entrysel.php?blog=1635>, 2005 (Letzter Zugriff: 13.04.2007).

Vossen, G.; Hagemann, St.: *Unleashing Web 2.0 – From Concepts to Creativity*; Morgan Kaufmann Publishers, San Francisco, CA 2007.

Wiley, D.A.: *The Instructional Use of Learning Objects*, Agency for Instructional Technology, 2002.



1 Daniel Dahl, Dipl.-Inf. (FH) M.Sc., studierte den Diplomstudiengang Allgemeine Informatik und den Master-Studiengang Medieninformatik an der Fachhochschule Köln. Er arbeitet zurzeit als Wissenschaftlicher Mitarbeiter am Lehrstuhl für Informatik der Universität Münster. Forschungsschwerpunkte bilden e-Learning an Hochschulen im Allgemeinen und speziell Web 2.0-Technologien und ihre Einflüsse auf die computergestützte Lehre.
E-Mail: dahl@ercis.de

2 Prof. Dr. Gottfried Vossen ist seit 1993 Inhaber der Lehrstühle für Informatik, insbesondere Datenbanken und Informationssysteme, am Institut für Wirtschaftsinformatik der Universität Münster. Er ist europäischer Herausgeber der bei Elsevier erscheinenden Fachzeitschrift *Information Systems*, Associate Editor des *International Journal on Semantic Web and Information Systems*, Editor-in-Chief der Dissertationenreihe „Databases and Information Systems“ (DISDBIS) des Berliner aka-Verlags sowie Mitglied des Herausgebergremiums der Zeitschriften *International Journal of Knowledge and Learning*, *Datenbank-Spektrum*, *Enterprise Information Systems* und *International Journal of Technology Enhanced Learning*. Seine Forschungsinteressen sind konzeptionelle sowie anwendungsnahe Fragestellungen im Umfeld von Datenbanken, Informationssystemen, Prozess-Modellierung, Web 2.0 und E-Learning.
E-Mail: vossen@uni-muenster.de