

## **Infrastrukturen zur Verwaltung von Metadaten für Offene Bildungsressourcen**

Marco Kalz, Roland Klemke, Stefaan Ternier, Marcus Specht  
Center for Learning Sciences and Technologies  
Open University of the Netherlands

### **Abstract**

In den letzten Jahren sind weltweit zahlreiche Initiativen gestartet worden, um offene Bildungsressourcen zu speichern und der Öffentlichkeit zugänglich zu machen. Zu diesem Zweck sind an vielen Orten Lernobjektorepositorien eingerichtet worden, die die persistente Speicherung von offenen Bildungsressourcen erlauben. Die dezentrale Einrichtung dieser Repositorien machte es für Endbenutzer schwierig, geeignete Ressourcen in ihrer Domäne zu finden und wiederzuverwenden. Darüberhinaus waren in vielen Fällen keine oder wenig standardisierte Metadaten mit den Lerninhalten gespeichert, was das Finden von passenden Lernressourcen noch mehr erschwerte. Um die Sichtbarkeit und Zugänglichkeit von Offenen Bildungsressourcen zu erhöhen, hat die Europäische Kommission im Rahmen des eContentplus-Programmes Projekte gefördert, die diese Situation verbessern sollten. In diesem Beitrag sollen Erfahrungen aus einigen Projekten vorgestellt und diskutiert werden. Dazu werden in diesem Artikel das verteilte Management von offenen Bildungsressourcen erläutert, die Erstellung eines Applikationsprofils diskutiert sowie die Rolle von verschiedenen Metadaten und deren Erstellung vorgestellt. Der Beitrag richtet sich an Projektleiter und Entscheidungsträger von Projekten, die sich mit Lernobjekten beschäftigen sowie an Forscher, die die Zugänglichkeit von Lernobjekten verbessern wollen.

### **Einleitung**

Die Diskussion über den Einsatz von Lernobjekten im technologiegestützten Lernen hat mittlerweile eine längere Tradition und in den letzten Jahren sind in vielen Ländern Sammlungen von Lernobjekten unter dem Stichwort der "Lernobjektorepositorien" (LOR) eingerichtet worden (Wiley, 2000). Dieser Trend, der u.a. durch die zunehmende Bedeutung der Open-Content- und Open-Access-Bewegung aber auch durch ökonomische Gründe zu erklären ist (vgl. Baumgartner & Kalz, 2005), hat zu einem regelrechten Wildwuchs von Lernobjektorepositorien geführt. Diese wurden auf verschiedenen institutionellen Ebenen (Schulen, Hochschulen, Hochschulverbände),

regionalen Netzwerken, Bundesländern, nach thematischer Ausrichtung oder aber auch auf Länderebene eingerichtet. Diese aus der Sicht der Open-Access-Bewegung sehr begrüßenswerte Trend brachte jedoch auch einige Nachteile mit sich. Die sehr heterogene Landschaft von Lernobjektrepositories hat zu zwei problematischen Konsequenzen geführt: Zum einen hat eine fehlende Standardisierung bei den LOR dazu geführt, dass es auf technologischer Ebene Interoperabilitätsprobleme gibt und zum anderen war es für potenzielle Nutzerinnen und Nutzer der Lernobjekte schwierig, geeignete Inhalte zu identifizieren, da unterschiedliche Klassifikationssysteme und Beschreibungsmethoden entwickelt wurden.

Besonders das Problem der fehlenden "Sichtbarkeit" und Auffindbarkeit von Lernobjekten hat die Europäische Kommission motiviert, ein eigenes Förderprogramm "eContentplus" aufzulegen (European Commission, 2005), dessen Ziel die bessere Zugänglichkeit von digitalen Inhalten in Europa war. Im folgenden Beitrag wollen wir unsere Erfahrungen aus einigen ausgewählten Projekten berichten sowie Probleme und Lösungsansätze diskutieren:

MACE (Metadata for Architectural Contents in Europe, <http://portal.mace-project.eu/>) verbindet Lernobjektrepositories aus dem Bereich der Architektur und reichert diese mit weiteren Metadaten an.

Share.TEC (Sharing Digital Resources in the Teaching Education Community, <http://www.share-tec.eu/>) fördert das Teilen von Lernressourcen aus verschiedenen Wissensdomänen.

MELT (A metadata ecology for learning and Teaching, <http://www.melt-project.eu/>) fokussierte auf die Verbesserung der Qualität von Metadaten.

iCOPER (Interoperable Content in a Competency-driven Society, <http://www.icoper.org/>) wird über 12500 Stunden von Lerninhalten zu verschiedenen Themen verfügbar machen.

OpenScout (Skill based scouting of open user-generated and community-improved content for management education and training, <http://www.openscout.net/>) konzentriert sich auf den benutzerfreundlichen Zugang zu Inhalten im Bereich Management und Business Education mit einem speziellen Fokus auf Bedürfnisse von KMU.

Im Rahmen des Beitrages werden verschiedene Fragen zur Entwicklung von Infrastrukturen für die Verwaltung von Metadaten für Offene Bildungsressourcen diskutiert, die alle aus den oben beschriebenen Projekten stammen. Der Beitrag gliedert sich wie folgt: Zu Beginn diskutieren wir das Problem des verteilten Managements von offenen Bildungsressourcen sowie die kritische Aufgabe der Erstellung eines Applikationsprofils. Darauf folgend stellen wir verschiedene Typen von Metadaten vor und diskutieren Ansätze, um Metadaten zu produzieren. Zum Schluss diskutieren wir einige aktuelle Probleme und Forschungsperspektiven.

### **Verteiltes Management von Offenen Bildungsressourcen**

Die Sammlung von offenen Bildungsressourcen findet in den meisten Fällen auf der Ebene von einzelnen Institutionen oder Institutionsnetzwerken statt. Dies führt zu heterogenen und verteilten Ressourcen, deren Nutzergruppen oft isoliert von einander agieren, ohne die Potenziale der gemeinschaftlichen Bewertung, Tagging etc. zu nutzen (Kalz, Drachsler, Van Bruggen, Hummel, & Koper, 2008). Im Rahmen des E-Learning Framework (Wilson, Blicno, & Rehak, 2004) wurden verschiedene Klassen von E-Learning Diensten vorgeschlagen, u.a. auch Dienste für die Verwaltung von Lerninhalten (Archivierung, Digitales Rechte management, Verwaltung von Metadaten, verteilte Suche etc.). Aus Benutzersicht stehen im Kern des verteilten Managements von offenen Bildungsressourcen die Aktivitäten des Suchens, Betrachtens und Publizierens von Lernobjekten. Zur Unterstützung dieser Aktivitäten stehen derzeit verschiedene Ansätze und Werkzeuge zur Verfügung wie z.B. die folgenden:

Suchen: Protokolle und Programmierschnittstellen (sog. APIs) zur Suche von Lerninhalten (Simon et al, 2005) wie z.B. SQI (Simple Query Interface).

Betrachten: Verteilung von Metadaten über Spezifikationen wie z.B. RSS (Really Simple Syndication) (RSS, ohne Jahr) oder aber OAI-PMH (Open Archive Initiative's Protocol for Metadata Harvesting) (Lagoze and van de Sompel, 2001).

Publizieren: Hier existieren verschiedene Lösungen wie z.B. SPI (Simple Publishing Interface) (Ternier et al, 2009), SWORD (Simple Webservice Offering Repository Deposit) oder auch PENS (Package Exchange Notification Services).

Für die gleichzeitige Suche in verteilten Lernobjektrepositories bieten sich verschiedene Optionen. Ternier et al. (2009) beschreiben verschiedene Implementationsmöglichkeiten einer verteilten Suche. Dabei gibt es zwei wichtige Hauptinitiativen, wie die gleichzeitige Suche in verteilten Repositorien implementiert werden kann. Dies ist zum einen die "Federated Search" Strategie und zum anderen die "Metadata Harvesting" Strategie. Bei der Implementation des "Federated Search Patterns" werden Suchanfragen über ein einheitliches Interface gleichzeitig in verschiedenen Repositorien durchgeführt. Dies hat jedoch den Nachteil, dass die Zuverlässigkeit der Suchergebnisse und auch die Geschwindigkeit nur so hoch sein kann wie das langsamste Repository im Verbund. Im Rahmen des "Search on Harvest Patterns" werden Metadaten der Repositorien an einer zentralen Stelle gesammelt und Suchanfragen werden in diesem durchgeführt, ohne jedoch den Ursprungsrepositorien die Suchanfragen weiterzuleiten. Hier spricht man auch vom Aufbau eines "Referatory", da nicht die digitalen Objekte sondern Links und Metadaten zu diesen gesammelt werden.

Im Rahmen der oben genannten Projekte standen und stehen wir oft vor der Herausforderung, eine einheitliche Plattform für die Suche und Nutzung von Lernobjekten zu gestalten, ohne jedoch die grundlegende Architektur der Ursprungsrepositorien verändern zu können. Daher zeigte sich die Lösung des Metadata-Harvesting als die gangbarste, die gleichzeitig auch noch ausreichend Flexibilität bietet, um Mehrwertdienste für Endnutzer auf dieser Basis zu entwickeln. Dieser Ansatz lässt sich über das sog. OAI-PMH (Open Archives Initiative's Protocol for Metadata Harvesting) realisieren. OAI-PMH ist vergleichbar mit dem RSS oder dem Atom-Protokoll, wobei es den Vorteil hat, dass zum einen nur in einem bestimmten Zeitraum veränderte Metadaten übertragen werden und zum anderen diese Veränderungen effizienter verwaltet werden können. Diese drei Protokolle kommen i.d. Regel auch für das Betrachten von Lernobjekten in externen Community-Plattformen in Frage.

Für das Publizieren von Lernobjekten mit externen Werkzeugen in ein entferntes Repository ist vor allem das Simple Publishing Interface (SPI) von Bedeutung. Ternier et al. (2008) argumentieren, dass das entfernte Publizieren vor allem wichtig ist, um Lernobjektautoren zu ermöglichen, gleichzeitig in ein Lernmanagementsystem ihrer Wahl und ein Repository zu publizieren. Ein weiterer interessanter Anwendungsfall ist die Publikation aus klassischen Autorenwerkzeugen wie Textverarbeitungsprogrammen oder Präsentationssoftware. Die dabei übermittelten Metadaten spielen im Publikationsprozess aber

auch zu einem späteren Zeitpunkt eine wichtige Rolle für die Auffindbarkeit der Lernressourcen. Im folgenden Teil wollen wir die Entwicklung eines Applikationsprofils sowie verschiedene Arten von Metadaten diskutieren.

## **Entwicklung eines Applikationsprofils**

Bei der Umsetzung des "Metadaten Harvesting" Ansatzes ist die Erstellung eines Metadaten-Applikationsprofils kritisch für den Erfolg der Zusammenführung verschiedener Lernobjektorepositorien. Dieses meist auf dem LOM-Standard (IEEE LTSC, 2002) basierende Profil definiert die Metadattypen, die für den Erfolg der Community-Plattform von Bedeutung sind. Dabei geht es zum einen um die Entscheidung, welche Metadaten als zwingend benötigt angesehen werden und welche optional sind. Welche Metadaten in welche Kategorie fallen, hängt von den Zielen und Endnutzern der Plattformen ab. Zu den zwingend benötigten Metadaten gehören oft z.B. Informationen wie Titel, Autoren, eindeutige Identifier, Format etc.. Zu den optionalen Metadaten können z.B. Informationen zur Lerndauer, dem didaktischen Ansatz der Lernressource oder aber auch die Zielgruppe gehören. Zu den weiteren Metadaten gehören in der Regel Metadaten, die nicht aus den ursprünglichen Repositorien stammen, sondern erst im Rahmen der Community-Plattform erstellt werden. Im Rahmen des MACE-Projektes wurde z.B. die Möglichkeit geschaffen, Bauwerke über mobile Endgeräte zur MACE-Plattform hinzuzufügen. Diese neuen Objekte haben in diesem Fall z.B. Metadaten wie die Geoposition der Objekte, die bei vielen Objekten aus den Ursprungsrepositorien nicht vorhanden waren. Im Rahmen von OpenScout werden verschiedene Möglichkeiten erforscht, Nutzer bei der Anreicherung der Lernressourcen mit Metadaten zu unterstützen, die Informationen zur Zielkompetenz einer Lernressource geben. Weitere projektübergreifende Funktionalitäten sind z.B. Content Tagging, Kommentare oder Ratings.

Die Erstellung eines Applikationsprofils umfasst zwei unterschiedliche Prozesse: die technische Definition des Applikationsprofils und den inhaltlichen Einigungsprozess zur Festlegung der semantischen Informationen. Die technische Definition legt Metadatenfelder, Wertebereiche, Austauschformate, Schnittstellen und Services fest. Der technische Prozess ist gut verstanden und wiederholbar. Auch wenn die Erstellung eines Applikationsprofils aus technischer Sicht keine sehr komplizierte Aufgabe ist, stellte der Einigungsprozess über

dieses Profil und die benötigten Metadaten die Projekte vor besondere Herausforderungen. Besonders bei der Nutzung von Domänen-Ontologien oder auch Kompetenzprofilen gilt es hier, verschiedene Fachvertreter und deren Sicht auf Ihr Fach so zusammen zu bringen, dass sich einerseits diese Perspektiven in der Endimplementation wiederfinden, andererseits der Endbenutzer jedoch trotzdem eine klare Orientierungshilfe durch die Nutzung dieser Kompetenzmaps und Kompetenzprofile haben sollen.

Der Prozess der Entwicklung des Applikationsprofils ist kritisch für die gesamte Plattform und muss daher besonders gründlich durchgeführt werden. Insbesondere können spätere Änderungen des Applikationsprofils weitreichende Auswirkungen haben, wenn bereits erhebliche Mengen Metadaten gesammelt wurden oder viele Repositorien verbunden sind. Im folgenden gehen wir auf verschiedene Arten von Metadaten ein, die im Rahmen des Applikationsprofils zum Einsatz kommen können.

## **Verschiedene Arten von Metadaten**

Metadaten werden informell als „Daten über Daten“ definiert. Sie beschreiben jedoch nicht nur Daten: Hunter (2003) definiert Metadaten als mehrwerthaltige Information, die Charakteristika von Ressourcen dokumentiert. Metadaten sind demnach die Grundlage für schnellen, präzisen und automatisierten Ressourcenzugriff über Organisations- und Netzwerkgrenzen hinweg. Metadaten bilden somit eine Schnittstelle zu Daten (Schaffner, 2009). Im Folgenden diskutieren wir verschiedene Klassen von Metadaten: soziale, kontextuelle, kompetenzbezogene oder domänenbezogene Metadaten.

Die Hinwendung zu nutzergenerierten Inhalten und sozialen Web-Communities führt derzeit zur massiven Verfügbarkeit von sozialen Metadaten: Content- und Bookmark-Sharing Sites oder Weblogs verwenden Tags, Annotationen, Bewertungen, Kommentare und Diskussionen. Diese Mechanismen lassen sich auch für Bildungsressourcen anwenden, wobei verschiedene Gründe für die Erzeugung sozialer Metadaten unterschieden werden (Skågeby, 2009): teilen, empfehlen, filtern, kommunizieren oder die persönliche Organisation. Soziale Metadaten verwenden in der Regel einfache Strukturen: fünf-stufige Skalen oder Eingabefelder für Schlüsselwörter sind alles, was benötigt wird. Die Qualität sozialer Metadaten hängt stark von individuellen Benutzern bzw. der Anzahl beitragender Nutzer

ab. Aktuelle Standardisierungsbemühungen zur Übertragbarkeit von Nutzerprofilen (Attention Profile Markup Language, APMML) und die Entstehung von Metriken zur Messung der Teilnahme von Nutzern in Communities (User Labour Markup Language, ULML) unterstreichen die Bedeutung von sozialen Metadaten.

Kontextuelle Metadaten haben in den letzten Jahren, insbesondere im Zusammenhang mit der mobilen Nutzung von Informationen, erheblich an Bedeutung gewonnen. Durch Kontextinformationen können Nutzer beispielsweise Informationen ortsbezogen filtern und durchsuchen wie es z.B. bei "location based services" geschieht. Als Kontext im weiteren Sinne werden in der Literatur jedoch die "Bedingungen und Umstände welche für ein Ereignis relevant sind" (Collins 1999) oder "in Wechselbeziehung stehende Umstände zu einem Ereignis" (Webster's 1996) verstanden. Jedoch ist das, was als Kontext verstanden wird, abhängig vom Betrachtungsgegenstand (Klemke, 2000). In der Forschung zu kontextsensitiven Systemen wird von einer Vielzahl von Kontextparametern ausgegangen: "Kontext summiert jegliche Information die nützlich zur Beschreibung der Situation einer Entität ist. Entitäten können hierbei sowohl physikalischer wie auch digitaler Natur sein" (vgl. Gross und Specht, 2001).

Die Relevanz von Kontextinformation für Lernunterstützung wird insbesondere im mobilen Lernen und "ubiquitous learning" hervorgehoben. Die Nutzung von Kontextinformationen kann eine Vielzahl von Perspektiven auf Lerninhalte und deren mögliche Kontexte oder Nutzungssituationen eröffnen (Specht, 2008). Hierbei wird auch auf die Relevanz von Kontextinformationen für die Nutzung von Lerninhalten und Services im Blended Learning hingewiesen: der Kontext beschreibt die jeweilige Lernsituation genauer beschreibt und Metadaten ermöglichen eine Anpassung der Lernunterstützung an diese Lehr-/Lernsituation.

Im Projekt MACE (Metadata for Architectural Contents in Europe) (Wolpers et al, 2009b) wurden verschiedene Arten von Metadaten für bestehende Lernobjekte angereichert. Hierbei wurde Kontextinformation zum Einen genutzt um den Bezug von digitalen Lernobjekten mit Gebäuden und anderen real existierenden Objekten zu spezifizieren, zum Anderen wurden auch reale Objekte als Lernobjekte in Lernszenarien verwendet. Darüberhinaus wurde auch die individuelle Nutzung von digitalen und realen Lernobjekten über Nutzungsdaten protokolliert. Die Verbindung von Kontextmetadaten, Nutzungsdaten oder sozialen Metadaten ist ein typisches Beispiel, in dem die Verbindung von verschiedenen Formen von Metadaten zu mehrwertigen Nutzungsszenarien führt. Hierbei können Nutzer

beispielsweise Inhalte oder Lernunterstützung im allgemeinen Sinn nicht nur über ihren aktuellen Standort suchen sondern diese Suche auch mit den Nutzungsgewohnheiten anderer Nutzer verbinden.

Kompetenzen sind in der Lehr/Lernforschung ein seit langem diskutiertes Thema. Auch in neueren Szenarien zum Kompetenzmanagement werden ausgehend von Kompetenzprofilen für bestimmte Positionen in Unternehmen und individuell diagnostizierten Kompetenzprofilen Empfehlungen für berufliche Weiterbildung oder Karrierepfade beschrieben (Erpenbeck & Heyse 1999, Probst, Deussen, Eppler & Raub, 2000).

In den meisten Ansätzen wird hierbei von einer Taxonomie von relevanten Kompetenzen und diagnostischen Verfahren für einen Domäne ausgegangen, welche die Beschreibung von Zielprofilen und die Diagnose relevanter Kompetenzen sowie individuellen Nutzerprofilen ermöglicht. Die Beschreibung von Lerninhalten mit Kompetenzmetadaten stellt sich insofern als problematisch dar, da in Anhängigkeit von der Granularität der Lernobjekte oft keine konkreten Kompetenzen spezifiziert werden können, da fein granulare Lernobjekte in verschiedenen Kontexten und Kompetenzentwicklungsszenarien verwendet werden können. Darüberhinaus ist der Lehr/Lerneffekt eines einzelnen fein granularen Objektes oft nicht zu messen, sondern es bedarf einer höheren Ebene von "Kompetenzentwicklungsprogrammen" in denen Lernobjekte in einem Lernpfad miteinander verbunden sind. Kompetenzmetadaten für wiederverwendbare Lernobjekte können daher auch kritisch gesehen werden: Während Lernobjekte feingranular sein sollten, um eine möglichst gute Wiederverwendung in verschiedenen Kontexten zu ermöglichen, ist Kompetenzentwicklung ein langfristiger Prozess - der Einfluss einer einzelnen Lernressource auf die Kompetenzentwicklung ist daher schwer zu bestimmen.

Zur Beschreibung standardisierter Kompetenztaxonomien, welche den Austausch und die gegenseitige Referenzierung von Kompetenzen ermöglichen, wurden verschiedene Standards und Austauschformate entwickelt. Während HR-XML und IMS-RDCEO es ermöglichen, Kompetenzen zu beschreiben und deren Relationen untereinander zu definieren, fokussiert Learning Object Metadata (LOM) auf die Beschreibung von Lernobjekten mit Lernzielen. Dabei wird das Lernziel in der Praxis oft synonym mit Kompetenz benutzt wird, auch wenn es hier aus theoretischer Sicht Unterschiede gibt. Daher koexistieren Standardisierungsbemühungen in den Bereichen Kompetenzmodellierung und Lernressourcen ohne fundierte



Integration: HR-XML (HR-XML Consortium, 2004) und IMS RDCEO (IMS Global Learning Consortium, 2002) fokussieren auf individuelle Kompetenzprofile, nicht auf Lernprozesse (Sampson et al, 2007). Umgekehrt beschreibt LOM Lernressourcen ohne Kompetenz-bezogene Information. Erst in jüngster Zeit gibt es Ansätze, beides zu verbinden (Sampson, 2009).

Auch wenn der Einfluss von Domänen-bezogenen Metadaten auf Genauigkeit und Effizienz von Suchergebnissen gering sein kann (Hawking and Zobel, 2007) und insbesondere allgemeine web-basierte Informationssysteme kaum profitieren (Henshaw and Valauskas, 2001), sind die Anstrengungen zur Entwicklung von Domänen-bezogenen Metadatenmodellen vielfältig: von kontrollierten Wortschätzen über Ontologie-gestützten, semantischen Applikationsprofilen wie z.B. OWL (Koutsomitropoulos et al, 2007) bis hin zu komplexen Domänenmodellen wie das CIDOC Conceptual Reference Model für kulturelles Erbe (Crofts et al, 2003).

Domänen-bezogene Metadatenmodelle wurden auch für MACE und ICOPER angelegt. Während MACE deutlich von einer architektonischen Begriffstaxonomie profitiert, was sich in verbesserten Suchresultaten niederschlägt (Wolpers et al, 2009), ist für den breiter angelegten inhaltlichen Focus von ICOPER auf "Hochschulbildung" ein einheitliches Vokabular nur schwer zu finden. Das ICOPER Domänenmodell bleibt daher vergleichsweise klein.

Wir stellen daher fest, dass Domänenmodelle einen Einfluss auf Genauigkeit und Effizienz von Suchprozessen haben, wenn (a) das Domänenmodell für den Anwendungsfall gut definiert ist und (b) die Metadaten für die zu suchenden Ressourcen in guter Qualität vorliegen. Um jedoch Metadaten in hoher Qualität zu erlangen, bedarf es einer Strategie, die verschiedene Ansätze mit einander vereint. Diese Ansätze werden im nächsten Teil des Artikels diskutiert.

## **Generierung von Metadaten**

In diesem Abschnitt werden Verfahren zur Generierung von Metadaten vorgestellt: Manuelle (durch Experten, Autoren, Anwender), automatische und kombinierte Verfahren.

Klassische Produzenten von Metadaten sind z.B. Bibliothekare (Greenberg, 2002). Im digitalen Zeitalter werden Metadaten für Archive und Repositorien generiert. Wissenschaftliche Verlage organisieren ihre Inhalte anhand standardisierter oder proprietärer

Metadaten schemata. Die professionelle Erstellung von Metadaten ist langsam, teuer und skaliert schlecht. In sehr dynamischen Domänen ist sie daher ungeeignet.

Autoren sind Spezialisten in ihrem Fachgebiet, aber keine Metadatenexperten (Thomas and Griffin, 1999). Dennoch können durch Autorinnen generierte Metadaten hochwertig sein, wenn ein verständliches Metadaten schema wie z.B. Dublin Core eingesetzt wird (Greenberg et al, 2001). Auch die Autoren profitieren von der Metadaterstellung: ihre Ressourcen werden findbar und wiederverwendbar. Allerdings mangelt es an einem konsistenten Verständnis von Metadaten zwischen verschiedenen Autoren wodurch inkonsistente Metadaten entstehen können.

Die Eingabe von Metadaten erfordert hohen Aufwand - auch wenn bereits viel Aufwand in die Erstellung von benutzbaren Eingabeformularen investiert wurde. Die ARIADNE Foundation unterscheidet fünf Metadaten-Kategorien ("General", "Semantics", "Pedagogical", "Technical" and "Indexation data"), die als Basis für den IEEE LOM Standard dienen. LOM umfasst ca. 80 Metadatenfelder. Automatische Verfahren zur Generierung von Metadaten können hier den manuellen Aufwand deutlich verringern.

Metadaten können aus verschiedenen Quellen generiert werden. Das Samgi Framework (Cardinaels et al, 2005) definiert Extraktoren für Objekt- und Kontext-basierte Metadaten. Objekt-basierte Extraktoren generieren Metadaten direkt aus der Ressource selbst: Das Harvard Object Store Validation Environment (JHOVE, 2010) benutzt z.B. Dateigröße, MIME-Typ, Änderungsdatum und andere Dateieigenschaften. Kontextuelle Informationen über Ressourcen helfen ebenfalls bei der Metadatengenerierung: beispielsweise kann aus dem Vorgang, bei dem eine Präsentation in einen Kurs zu "Künstliche Intelligenz" für Bachelor Studenten im Fachbereich Informatik hochgeladen wird, folgende Kontext-Information erschlossen werden: Klassifizierung unter "Informatik", Zielgruppe "Hochschulbildung".

Nutzergenerierte Metadaten können für Suchprozesse in bestimmten Kontexten ebenso nützlich sein, wie professionell erstellte Metadaten (Melenhorst et al, 2008). Allerdings: Nutzer, die Metadaten liefern wissen vielleicht gar nicht, was Metadaten sind. Aus diesem Grund können formale Prozesse und komplexe Taxonomien unangebracht für die Generierung relevanter Metadaten sein. Erfolgreiche Lösungen für

diese Situation existieren in verschiedenen Domänen (e-Shops, Content-Sharing Anwendungen, Community Anwendungen):

- *Tags* liefern Schlüsselwort Annotationen.
- *Bewertungen* reflektieren die Qualität oder Angemessenheit von Inhalten über Skalen.
- *Kommentare* können zur semantischen Annotation verwendet werden.
- *Feedback-Formulare* liefern semi-strukturierte Information durch den Einsatz geschlossener und offener Fragen.

Doch trotz optimierter Tools und Verfahren, bedeutet Metadaten-Erstellung immer zusätzlichen Aufwand. Wodurch werden Nutzer motiviert? Persönlicher Nutzen ist wichtig (z.B. das Bedürfnis eine Ressource später selbst wiederzufinden), aber es gibt Gründe, die darüber hinausgehen: die Steigerung der Relevanz einer Ressource in einer Gemeinschaft (z.B. bei Online-Bookmark Anwendungen) oder sozialer Nutzen (z.B. die erhöhte persönliche Sichtbarkeit innerhalb einer Community). Hasan and Jameson (2008) nennen positive Erfahrungen, erfolgreiche Aufgabenbewältigung und sichtbare Verbesserung von Content-Sammlungen als weitere motivierende Faktoren. Allerdings unterscheiden sich Metadaten, die für persönliche Zwecke erstellt wurden erheblich von solchen, die durch soziale Gründe motiviert sind (Bentley and Labelle, 2008). Nutzergenerierte Metadaten verursachen geringe Kosten, sind skalierungsfähig und häufig von brauchbarer Qualität. Allerdings führen unstrukturierte Terminologien häufig zu widersprüchlichen oder unpassenden Klassifizierungen. Cho und Tomkins (2007) benennen zusätzliche Probleme: Einerseits kann die Anfälligkeit für Spam und Missbrauch zu unbrauchbaren Metadaten führen, andererseits können persönliche Interessen zu Metadaten führen, die für andere wenig hilfreich sind. Diese Probleme werden mit größeren Nutzerzahlen kleiner, wenn auch entsprechende Korrekturmechanismen aktiv sind.

Kombinierte Ansätze zur Metadatengenerierung können die Nachteile der einzelnen Verfahren minimieren. Bei Wikipedia, beispielsweise, verschwinden die Grenzen zwischen automatisierten, professionellen oder nutzergenerierten Metadaten. Die Nutzer, die Spezialistinnen, Enthusiasten und interessierte Laien umfassen, entwickeln gemeinsam Inhalte und Metadaten. Dieser Ansatz hat dazu geführt, dass Wikipedia eine sehr bekannte und hochwertige Informationsquelle geworden ist

und somit zum Teil bereits von externen Systemen zur Generierung von Metadaten eingesetzt wird (vgl. Meyer, Rensing & Steinmetz, 2007)

Offensichtlich können manche Metadatenfelder einfach automatisch generiert werden, während andere manuellen Aufwand erfordern (Spaniol et al, 2008). Eine im Rahmen von MACE durchgeführte Evaluation, die Metadaten von Experten und Nutzern kombiniert mit automatisch generierten Metadaten aus Nutzerbeobachtung und Extraktionsverfahren (Mommel et al, 2008), zeigt dass diese Kombination die Performanz von Studenten bei der Suche und Anwendung von Lernressourcen signifikant verbessert (Wolpers et al, 2009a). Allerdings kann die Kombination verschiedener Verfahren auch die Komplexität und Menge der verfügbaren Metadaten erhöhen. Aus diesem Grund werden attraktive, benutzbare und intuitive Benutzerschnittstellen und Metadatenvisualisierungen benötigt, die die Interaktion der Nutzer mit Metadaten und Inhalten vereinfachen (Zambelli et al., 2008).

### **Fazit und Ausblick**

In den vorgestellten Projekten (und vielen weiteren) wurden bereits verfügbare Lernobjektsammlungen miteinander vernetzt und zugänglich gemacht. Die dazu nötigen technischen und fachlichen Prozesse sind mittlerweile gut verstanden und können für weitere Anwendungsfelder reproduziert werden.

Es hat sich gezeigt, dass folgende Aspekte für eine erfolgreiche Vernetzung von Lernobjektsammlungen essentiell sind:

- Das Applikationsprofil ist kritisch für den Erfolg des gesamten Projektes. Es ist mit höchster Sorgfalt zu entwickeln und muss die Verständlichkeit durch Endnutzer, die Wartbarkeit, die Vollständigkeit und die Korrektheit der verwendeten Terminologien und Strukturen berücksichtigen.
- Ein klarer Domänenbezug hilft bei der Erstellung des Applikationsprofils und bei der Akzeptanz der vernetzten Sammlungen durch den Endnutzer.
- Die verwendeten Arten von Metadaten müssen für den Anwendungsfall sorgfältig ausgewählt werden, um zu gewährleisten, dass Metadaten verfügbar, bedeutungsvoll und nutzbar sind.
- Der Erstellungsprozess für Metadaten muss auf die verfügbaren Metadatenquellen abgestimmt sein. Verschiedene Quellen

- erfordern verschiedene Generierungsprozesse. Der individuelle manuelle Aufwand muss so klein wie möglich gehalten werden.
- Eine gute Visualisierung der Metadaten erleichtert den Zugang zu den gesammelten Metadaten.

Probleme bei der Vernetzung von verteilten Lernobjektsammlungen bestehen in der nachhaltigen Pflege der erstellten Strukturen: Für verschiedene Situationen und Ereignisse sind Regeln nötig, die sicherstellen, dass die vernetzten Bildungsressourcen langfristig erreichbar und einsetzbar sind:

- Nach welchen Verfahren werden neue Sammlungen integriert? Welcher Aufwand entsteht beim Betreiber der Lernobjektsammlung, welcher bei der zentralen Metadatensammelstelle?
- Wie wird sichergestellt, dass Sammlungen, die nicht mehr verfügbar sind, entfernt werden können?
- Wie wird mit Änderungen im Applikationsprofil umgegangen?
- Wie wird eine hohe Qualität der vernetzten Bildungsressourcen sichergestellt?

Offene Fragen ergeben sich im Zusammenhang mit den Einsatzmöglichkeiten der findbaren Lernressourcen. Hier ist es nötig verschiedene Einsatzszenarien zu unterscheiden und technologische Unterstützung für diese Szenarien zukünftig zu entwickeln:

- Einsatz in Lehrprozessen durch Lehrkräfte: wie können Lernressourcen in Vorlesungen/Übungen/Seminare integriert werden?
- Einsatz in selbstgesteuerten Lernprozessen: wie können Lernende auf Lernressourcen zugreifen und sinnvoll komplexe Lernprozesse damit gestalten?
- Einsatz in Erstellungsprozessen: wie können offene Lernressourcen eingebunden werden in die Erstellung neuer Inhalte?
- Einsatz in kontextualisierten Lernszenarien: wie können die Metadaten der Lernressourcen eingesetzt werden, um eine automatische Versorgung von Lernenden mit Lernmaterialien in der richtigen Situation zu gewährleisten?

### **Acknowledgement**

This research has been co-funded by the European Commission within the eContentplus targeted project OpenScout, grant ECP 2008 EDU

428016 (cf. <http://www.openscout.net>), the best practice network iCoper, grant ECP-2007-EDU-417007 (cf. <http://www.icoper.org/> and the targeted project Share.Tec, grant ECP-2007-EDU-427015 (cf. <http://www.sharetecproject.eu/>).

## **Referenzen**

APML, Attention Profile Markup Language, <http://www.apml.org/>

Baumgartner, P., & Kalz, M. (2005). Wiederverwendung von Lernobjekten aus didaktischer Sicht. In D. Tavangarian & K. Nölting, Auf zu neuen Ufern! E-Learning heute und morgen. (pp. 97-106). New York: Waxmann.

Bentley, C. & Labelle, P. (2008). A comparison of social tagging designs and user participation. Metadata for semantic and social applications: Proceedings of the International Conference on Dublin Core and Metadata Applications - DC-2008, Berlin, 22-26 September 2008.

Cardinaels, K., Meire, M., and Duval, E. (2005). Automating metadata generation: the simple indexing interface. Proceedings of the 14th international conference on World Wide Web, pages 548-556. ACM Press.

Cho, J. & Tomkins, A. (2007). Guest Editors' Introduction: Social Media and Search, IEEE INTERNET COMPUTING, Vol. 11, 06, pp. 13-15.

Collins 1999. Collins English Dictionary, Millennium Edition, HarperCollins Publishers, Glasgow, 1999.

Crofts, N., Doerr, M. and Gill, T. (2003). The CIDOC Conceptual Reference Model: A standard for communicating cultural contents, Cultivate Interactive, issue 9, 7 February 2003.

Erpenbeck, J., & Heyse, V. (1999). *Die Kompetenzbiographie. Wege der Kompetenzentwicklung.* (2 ed.). Münster: Waxmann. Erpenbeck, J., & Heyse, V. (1999). *Die Kompetenzbiographie. Wege der Kompetenzentwicklung.* (2 ed.). Münster: Waxmann.

European Commission (2005). DECISION NO 456/2005/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL. Official Journal of the European Union.

Greenberg, J., Pattuelli, M. C., Parsia B., Davenport Robertson, W. (2001). Author-generated Dublin Core Metadata for Web Resources: A Baseline Study in an Organization. Proceedings of the International Conference on Dublin Core and Metadata Applications DC-2001, October 24-26, 2001, NII, Tokyo, Japan.

Greenberg, J. (2002) Metadata generation: Processes, people and tools. *Bulletin of the American Society for Information Science and Technology*, December 1, 2002.

Gross, T. and Specht, M. (2002). Aspekte und Komponenten der Kontextmodellierung. *i-com – Zeitschrift fuer interaktive und kooperative Medien* 3 (2002). pp. 12-16. (ISSN: 1618-162)

Hasan, T., Jameson, A. (2008). Bridging the Motivation Gap for Individual Annotators: What Can We Learn From Photo Annotation Systems? Proceedings of INSEMTIVE 2008, Workshop on Incentives for the Semantic Web, Karlsruhe, Germany, October 26th, 2008.

Henshaw, R. & Valauskas, E.J. (2001). Metadata as a catalyst: experiments with metadata and search engines in the Internet journal, *First Monday*. Libri. 51.2 (2001): 86-101.

HR XML Consortium (2004). Competencies (measurable characteristics). Retrieved 1st of July, 2006 from [http://ns.hr-xml.org/2\\_3/HR-XML-2\\_3/CPO/Competencies.html](http://ns.hr-xml.org/2_3/HR-XML-2_3/CPO/Competencies.html)

Hunter, J. (2003) Working towards MetaUtopia - a survey of current metadata research. Special issue: Organizing the Internet. *Library Trends*, 52(2).

IMS Global Learning Consortium (2002). IMS reusable definition of competency or educational objective specification. Retrieved 1st of July, 2006 from <http://www.imsglobal.org/competencies/>

IEEE LTSC (2002). IEEE Standard for Learning Object Metadata. The Institute of Electrical and Electronics Engineers, Inc., 3 Park Avenue, New York, NY 10016-5997, USA.

JHOVE (2010). JHove JSTOR/Harvard Object Validation Environment, Zuletzt aufgerufen am 3. März 2010 unter <http://hul.harvard.edu/jhove>

Kalz, M., Drachsler, H., Van Bruggen, J., Hummel, H., & Koper, R. (2008). Wayfinding Services for Open Educational Practices. *International Journal of Emerging Technologies in Learning*, 3(2), 1-6.

Klemke, R. (2000). Context Framework - an Open Approach to Enhance Organisational Memory Systems with Context Modelling Techniques, PAKM2000: Third International Conference on Practical Aspects of Knowledge Management, 2000, Basel, Switzerland.

Koutsomitropoulos, D. A., Paloukis, G. E., & Papatheodorou, T. S. (2007). Semantic Application Profiles: A Means to Enhance Knowledge Discovery in Domain Metadata Models. Miguel-Angel Sicilia and Miltiadis D. Lytras (Eds.). *Metadata and Semantics*, Springer, ISBN 978-0-387-77744-3.

Lagoze, C. and Van de Sompel, H. (2001). The Open Archives Initiative: Building a Low-Barrier Interoperability Framework. *Proceedings of the 1st ACM/IEEE-CS Joint Conference on Digital Libraries*, ACM Press, pp. 54-62.

Memmel, M., Schirru, R., Wolpers, M. & Tomadaki, E. (2008). Towards the Combined Use of Metadata to Improve the Learning Experience, *IEEE International Conference on Advanced Learning Technologies*, pp. 930-932.

Meyer, M., Rensing, C., & Steinmetz, R. (2007). Categorizing Learning Objects Based On Wikipedia as Substitute Corpus. In D. Massart, J. Colin, & F. Van Assche, *Proceedings of the First International Workshop on Learning Object Discovery & Exchange (LODE'07)* (pp. 64-71). CEUR Workshop Proceedings. Retrieved from <http://sunsite.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-311/paper09.pdf>.

PENS, Package Exchange Notification Services, see <http://www.aicc.org/docs/AGRs/agr011v1.pdf>

Probst, G., Deussen, A., Eppler, M., & Raub, S. (2000). *Kompetenz-Management. Wie Individuen und Organisationen Kompetenz entwickeln..* Wiesbaden: Gabler Verlag.

RSS (no year). RSS 2.0 specification. Retrieved February 2, 2010 from <http://www.rssboard.org/rss-specification>.



Sampson, D., Karampiperis, P. and Fytros, D. (2007). Developing a Common Metadata Model for Competencies Description, *Interactive Learning Environments*, Special Issue on Learning Networks for Lifelong Competence Development, (ISSN 1049-4820), vol. 15(2), pp. 137-150, Routledge , August 2007

Sampson, D. (2009). Competence-related Metadata for Educational Resources that Support Lifelong Competence Development Programmes. *Educational Technology & Society*, 12 (4), 149-159.

Schaffner, J. (2009). The Metadata is the Interface: Better Description for Better Discovery of Archives and Special Collections. Synthesized from User Studies. OCLC Research. Dublin, Ohio: 2009.

Simon, B.; Massart, D.; van Assche, F.; Ternier, S.; Duval, E., Brantner, S., Olmedilla, D., & Miklós, Z. (2005). A Simple Query Interface for Interoperable Learning Repositories. Workshop on Interoperability of Web-Based Educational Systems in conjunction with 14th International World Wide Web Conference (WWW'05). May, Chiba, Japan

Skågeby, J. (2009). Exploring Qualitative Sharing Practices of Social Metadata: Expanding the Attention Economy. *The Information Society*, 25(1), January 2009 , pp. 60-72.

Spaniol, M., Klamma, R., Lux, M., (2008). Imagesemantics: User-Generated Metadata, Content Based Retrieval & Beyond. *Journal of Universal Computer Science*, vol. 14, no. 10 (2008), 1792-1807, 28.05.2008.

Specht, M. (2008). Designing Contextualized Learning. In H. Adelsberger, Kinshuk, J. Pawlowski, & D. Sampson, *Handbook on Information Technologies for Education and Training* (pp. 101-111). Berlin, Heidelberg: Springer Berlin Heidelberg. doi: 10.1007/978-3-540-74155-8.

SWORD, Simple Web-service Offering Repository Deposit (SWORD), <http://swordapp.org/>.

Ternier, S., Massart, D., Van Assche, F., Smith, N., Simon, B. & Duval, E. (2008). A Simple Publishing Interface For Learning Object Repositories. Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2008 (pp. 1840-1845). Chesapeake, VA: AACE.

Ternier, S.; Verbert K.; Parra, G.; Vandeputte, B.; Klerkx, J.; Duval, E.; Ordóñez, V. and Ochoa, X. (2009). The Ariadne Infrastructure for Managing and Storing Metadata. *Emerging E-Learning Technologies. IEEE Internet Computing*, 13 (4), pp. 18-25, July/Aug. 2009.

ULML, User Labor Markup Language, <http://userlabor.org/>

Webster. 1996. Webster's New Encyclopedic Dictionary, 1996.

Wiley, D. A. (2000). The instructional use of learning objects. Retrieved April 12, 2006, from <http://reusability.org/read/>.

Wilson, S., Blinco, K., & Rehak, D. (2004). An e-Learning Framework. A summary. Alt-I-lab 2004 (Advanced Learning Technology Interoperability). JISC:Bristol, UK.

Wolpers, M., Memmel, M., Giretti, A., (2009a). The MACE Experience - first evaluation results. Proceedings of Learning in the Synergy of Multiple Disciplines, 4th European Conference on Technology Enhanced Learning, EC-TEL 2009, Nice, France, September 29 - October 2, 2009, LNCS 5794.

Wolpers, M., Memmel, M., Klerkx, J., Gonzalo Parra, Bram Vandeputte, Erik Duval, Rafael Schirru, Katja Niemann (2009b). Bridging repositories to form the MACE experience. *Journal on New Review of Information Networking (RINN)*, Edited by Leslie Carr, published by Taylor & Francis Group, US, Septemeber 21, 2009.

Zambelli, M., Janowiak, A., Neuckermans, H. (Editors), 2008. Browsing architecture. Metadata and beyond. IRB Stuttgart (Germany), ISBN: 978-3-8167-7770-0, September 2008.