

# *Metadata Registries*

THOMAS FISCHER

## *Metadata Registries: Aktueller Stand und zukünftige Entwicklungen<sup>1</sup>*

Registries were established as tools of reference in association with Dublin Core metadata. In the future, complex and diversified metadata systems will be employed like the »application profiles« that are being developed at an increasing rate. At the same time, the visions expressed in the term »semantic web« are leading to a growing need for automated analysis and the translation of metadata procedures. These are to be accomplished through metadata registries. This will place higher demands on registries as far as implementing and administrating the metadata schemes is concerned, on the one hand, and the implementation of machine-readable structures on the other hand. This task can build upon existing registries and will lead to divided structures, which will ease the evaluation of metadata and thus generally contribute to raising the acceptance of the concept of »metadata.«

Registries haben sich im Umfeld des Dublin-Core-Metadatensatzes als Referenzwerkzeug etabliert. In der Zukunft werden komplexe und vielfältige Metadatensysteme benutzt werden, wie sie als »Anwendungsprofile« zunehmend entwickelt werden. Andererseits erwachsen aus den Visionen, die unter dem Stichwort »Semantic Web« formuliert werden, Anforderungen an automatisierte Prozesse der Analyse und Übersetzung von Metadaten.

Diese Aufgaben sollten die Metadaten-Registries bearbeiten. Das stellt höhere Ansprüche an Registries, was die Aufnahme und Verwaltung der Metadatenschemata einerseits und die technische Implementierung maschinenlesbarer Strukturen andererseits betrifft. Auf den bestehenden Registries aufbauend wird das zu verteilten Strukturen führen, die die Auswertung von Metadaten erleichtern und damit die Akzeptanz gegenüber dem Konzept »Metadaten« allgemein erhöhen werden.

### **WAS IST EIN »REGISTRY«?**

Der Begriff des *Registry*, der eigentlich Verzeichnis oder Register meint, wird hier in seiner englischen Form genutzt, um damit spezifische Register zu bezeichnen, wie sie im Internet seit einiger Zeit angeboten werden. Spezifischer noch soll *Registry* im Folgenden bezeichnen

- ein Verzeichnis von Metadatenelementen,
- das über das Internet erreichbar ist,
- und Informationen über die Definition und den Gebrauch der Metadatenelemente liefert.

Im Rahmen dieser Definition können sowohl normative als auch deskriptive Registries auftreten, solche, die auf ein Metadatensystem spezialisiert sind und solche, die verschiedene Systeme miteinander in Verbindung setzen. Der Schwerpunkt liegt auf Registries, die sich auf den Dublin-Core-Standard der Dublin Core Metadata Initiative (DCMI)<sup>2</sup> beziehen. Dass hier die Entwicklung noch im Fluss ist, auch verschiedene Ansprüche und Bedürfnisse unter dem einen Begriff subsummiert werden, soll zunächst an einigen Beispielen demonstriert werden. Registries sind im

Zusammenhang mit der intensiveren Auseinandersetzung mit Metadaten für das Auffinden von Informationen im Internet seit etwa 1998 entstanden. Eine genaue Chronologie soll und kann hier nicht vorgelegt werden, und die Auswahl der vorgestellten Registries ist sicher subjektiv gefärbt.

### **BEISPIELE VON REGISTRIES**

Das Projekt DESIRE (Development of a European Service for Information on Research and Education, URL: <http://www.desire.org/>) wurde von Juli 1998 bis Juni 2000 von der EU gefördert und von einem Konsortium aus Großbritannien, Schweden, Dänemark, Norwegen und den Niederlanden durchgeführt. Es hatte zum Ziel, die europäische Infrastruktur für Forschung und Bildung im Internetbereich zu verbessern. Dazu wurde im Wesentlichen im Bereich der Zwischenspeicherung (»Caching«), der Katalogisierung (»Resource discovery«) und der Strukturierung (»Directory Services«) gearbeitet. In diesem Kontext entstand das **DESIRE Metadata Registry** (URL: <http://desire.ukoln.ac.uk/registry/>), das von UKOLN (Bath, Großbritannien) entwickelt wurde, um Informationen über die Struktur und Semantik von Metadatensätzen bereitzustellen:

- Die Notwendigkeit herauszufinden wie Metadaten benutzt werden, betrifft gleichermaßen
- Menschen, die Metadaten entsprechend der gegebenen Standards erstellen wollen, die herausfinden wollen, ob für ihre Zwecke schon ein passender Metadatensatz existiert oder die ihre Metadatensätze anpassen wollen an andere, die schon für andere Zwecke existieren.
- Software, die Metadaten bearbeiten soll und daher die Struktur und Bedeutung erkennen muss, Metadaten-Erstellungsprogramme, die die Korrektheit der Eingabe überprüfen und Hilfestellung beim Ausfüllen der Formulare geben sowie Konvertierungswerzeuge, die Verweise auf Konkordanztabellen benötigen.« (a.a.O., Übers. d. A.)

Das Registry (URL: <http://desire.ukoln.ac.uk/registry/index.php3>) verzeichnet Metadatenschemata und ihre Beziehung zueinander. Auf der Basis einer »semantischen Schicht« können Konkordanzen zwischen verschiedenen Schemata erzeugt werden.

Das **DC Registry** (<http://dublincore.org/dcregistry/>) wird von der Dublin Core Metadata Initiative entwi-



Thomas Fischer

**Development of a European Service for Information on Research and Education**

ckelt und ist speziell der Definition der Elemente des Dublin-Core-Metadatensatzes und den damit zusammenhängenden Begriffen gewidmet. Es liefert in verschiedenen Sprachen und Schriften die gültige Referenz für die Nutzung der Dublin-Core-Metadaten.

Name	ID	Version	Status	Registration Authority	
ADL SCORM Content Metadata	adl-scorm-content/1.1	1.1	recommendation	ADL	<a href="#">Detail</a>
ADL SCORM Course Metadata	adl-scorm-course/1.1	1.1	recommendation	ADL	<a href="#">Detail</a>
ADL SCORM Raw Media Metadata	adl-scorm-raw/1.1	1.1	recommendation	ADL	<a href="#">Detail</a>
Australian Government Locator Service	aglis-ap	v 1.2	Australian Recommended	NAA	<a href="#">Detail</a>
BIBLINK Core eLib Collection Level Description	biblinkcore CLD	1999-08-02 draft	deployed	BIBLINK eLibCDWG	<a href="#">Detail</a>
DC Education Working Group proposal	dc-education	2000-10-05	draft proposal	DCEDWG	<a href="#">Detail</a>
Dublin Core Qualified	dcq/2000-07-11	2000-07-11	DCMI Recommendation	DC	<a href="#">Detail</a>
EASEL DCMES	easel-dcmes	2000-11	draft	EASEL	<a href="#">Detail</a>

Abb.1: DESIRE Registry: »Application Profiles«  
(URL: <http://desire.ukoln.ac.uk/registry/appprofile.php3>)

The screenshot shows the DC Registry's search interface. At the top, there are links for 'ABOUT THE INITIATIVE', 'DOCUMENTS', 'GROUPS', 'RESOURCES', 'DCMI NEWS', 'TOOLS AND SOFTWARE', 'MEETINGS AND PRESENTATIONS', and 'PROJECTS'. Below this is a search bar with fields for 'Search for:', 'Case sensitive?', 'Display results that match:', and 'Display:'. The 'Display:' dropdown is set to 'Summary'. On the left, there is a sidebar titled 'Summary of DCMI Element Refinements' listing terms like 'abstract', 'alternative', 'available', etc. On the right, the results are displayed with 'Items Found: 29'.

Abb. 2: DC Registry: Summary of DCMI Element Refinements  
(URL: <http://dublincore.org/dcregistry/navigateservlet?reqtype=refinements>)

Der **Metadatenserver** (<http://www2.sub.uni-goettingen.de/>) an der Staats- und Universitätsbibliothek Göttingen entstand im Rahmen des Projektes MetaLib (Metadata Initiative of German Libraries), das seit 1997 an der SUB Göttingen durchgeführt wird. Ziel war hier unter anderem: Die

- »Analyse der Metadatenanwendungen in den deutschen wissenschaftlichen Fachgesellschaften und Berücksichtigung der Anwendung in anderen nationalen und internationalen Fachgesellschaften,
- Vorbereitung eines Vergleichs dieser Modelle mit den bibliothekarischen Regelwerken,
- Unterstützung einer möglichst einheitlichen Anwendung des Dublin-Core Metadatenschemas bei potentiellen Anwendern (Bibliotheken, Verbundsystemen und Fachgesellschaften).«

(URL: [http://www2.sub.uni-goettingen.de/sub\\_de.html](http://www2.sub.uni-goettingen.de/sub_de.html))

Konkret wurden zwei Datenbanken entwickelt: **MetaForm** setzt Metadatenschemata verschiedener Projekte in Beziehung zu Dublin Core (<http://www2.sub.uni-goettingen.de/metaform/>), während **MetaGuide** (<http://www2.sub.uni-goettingen.de/metaguide/>) Informationen über Artikel und Projekte nachweist, die sich mit Metadaten beschäftigen.

Das Registry des Schemas-Projektes (<http://www.schemas-forum.org/registry/desire/index.php3>) wird von dem Projekt CORES weitergeführt. Es hat als Ziel die gemeinsame Nutzung von Metadaten-Definitionen und Metadatenprofilen.

## ERFAHRUNGEN

Die Nutzung von Metadaten hat im Internet noch eine vergleichsweise kurze Geschichte. Während diverse Suchmaschinen wie AltaVista schon früh die elementaren Metatags wie »Author« und »Description« nutzten, wurde erst 1997 mit der Gründung der »Dublin Core Metadata Initiative« (DCMI) eine systematische Nutzung von Metadaten zum Auffinden von Dokumenten vorgeschlagen. Die jährlichen Treffen der DCMI trugen dazu bei, den einfachen und später auch den komplexen (»qualified«) Standard in die internationale Diskussion einzubringen und zunehmend zu stabilisieren. Die etwa gleichzeitig entstehenden Registries (z.B. DESIRE und MetaGuide) halfen, in der Zeit der ersten Unübersichtlichkeit den Überblick über die Entwicklungen zu bewahren und den eigenen Metadaten-Einsatz den sich entwickelnden Standards anzupassen. Entwickler von Internetportalen fühlten sich zu jener Zeit oft in der Situation der Geschichte vom Hasen und Igel: hatte man mit Mühe die eigene Darstellung dem neuesten Standard angepasst, war dieser schon wieder ganz woanders angekommen.

Das hat sich seit einiger Zeit geändert: Der einfache Dublin-Core-Datensatz ist in der Version 1.1 stabil und sowohl ISO-Standard als auch Inhalt einer Übereinkunft des Europäischen Normungskomitees CEN. Der komplexe Datensatz ist auch weit gehend kanonisiert und seine Darstellung in verschiedenen »Transportverpackungen« (XML, RDF) ist gerade in die endgültige Festlegungsphase getreten.<sup>3</sup> Das hat auch zu einer gewissen Standardisierung der Beschreibung im Rahmen des XML-Modells des W3C geführt: Den Metadatenelementen sind universelle Identifikatoren (URIs) zugeordnet, die in verschiedenen Namensräumen existieren und so eindeutige Verweise erlauben.

Und während noch andere Metadatensätze (z.B. »Learning Object Metadata« des IEEE) für spezifische Bereiche große Bedeutung haben, hat sich als allgemeines Beschreibungssystem Dublin Core im Internetbereich weit gehend etabliert. Das zeigt sich zum Beispiel auch daran, dass das »Metadata Harvesting Protocol« (MHP) der »Open Archives Initiative« (OAI) den einfachen Metadatensatz nach Dublin Core verbindlich vorschreibt. So wird das »Dublin Core Metadata Element Set« (DCMES) nicht nur ein Werkzeug zum Auffinden von Daten, sondern zunehmend das generische Vehikel zum Austausch und Transport von Metadaten – zur Lingua Franca der Metadatenwelt.<sup>4</sup>

Das heißt, dass die Registries einige ihrer ursprünglichen Funktionen verloren haben, keinesfalls jedoch, dass sie jetzt überflüssig werden. Im Gegenteil wird jetzt erst zunehmend klar, welche die zukünftige Bedeutung der Registries sein wird. Es zeigt sich nämlich, dass einerseits die beschränkten Möglichkeiten des DCMES von vielen Initiativen und Projekten als unzureichend empfunden werden, um ihre je spezifischen Objekte sinnvoll zu beschreiben, andererseits auch gefordert wird, verschiedene Metadatensysteme möglichst automatisch miteinander in Beziehung zu setzen.

## **ANWENDUNGSPROFILE UND REGISTRIES**

Das erste führt zu dem Wunsch, diesen Datensatz um andere Datenelemente zu erweitern, die entweder ad hoc definiert oder aus anderen Definitionen (IEEE-LOM, MARC) übernommen werden. So entstehen die so genannten »Anwendungsprofile« (»Application Profiles«) als »Mix and Match«<sup>5</sup> der verschiedenen Metadatendefinitionen. Ziel ist damit, nicht für jedes Metadatenprojekt neue Elemente zu erfinden, sondern bestehende Metadatendefinitionen nachzunutzen. Das Vorgehen lässt sich in drei Stufen darstellen:

— Benutze definierte Metadatenelemente. Wenn nötig, spezialisiere die bestehende Definition.

The screenshot shows a web browser window titled 'MetaGuide: Standard Offices and Registries - Netscape'. The URL is <http://www2.sub.uni-goettingen.de/cgi-bin/ssgfi/anzeige.pl?db=meta&sc=LS>. The page displays a list of records under the heading 'Standard Offices and Registries (5 records)'. The first record is '1. MARC Standards : Library of Congress Network development and MARC Standards Office'. It includes fields for Subject Class (Other Initiatives; Standard Offices and Registries; United States; Cross Disciplinary; National/State), Source Type (Organization Infos / Research Projects; Mappings / Crosswalks; Manuals / User Guides), URL (<http://lcweb.loc.gov/marc/marc.html>), Language (English), Metadata present (no), Country (United States), and Last Update (20010802). A 'Metadata' icon is shown next to the URL. The second record is '2. Naming names: metadata registries', the third is '3. SCHEMAS: Forum for Metadata Schema Designers and Implementers', and the fourth is '4. The Diffuse Project Home Page'. Each record has similar fields and a 'Metadata' icon.

Abb. 3: MetaGuide: Standard Offices and Registries  
(URL: <http://www2.sub.uni-goettingen.de/cgi-bin/ssgfi/anzeige.pl?db=meta&sc=LS>)

- Definiere notfalls eigene Datenelemente.
- Fasse die benutzten Datenelemente in einem Anwendungsprofil zusammen.

Auch hier ergeben sich mittlerweile Möglichkeiten zur Kooperation, so werden z.B. für verschiedene Interessengruppen Standardprofile<sup>6</sup> erstellt. Für einzelne Projekte können in diesen Bereichen Standardprofile verwendet werden. Nur wenn dieses nicht ausreicht, muss ein eigenes Profil erstellt werden. Gleichzeitig führt die Diskussion in der Entwicklung dieser Profile schon zur Abstimmung der verschiedenen Interessen und damit zur Vereinheitlichung von Metadatenstandards. Das Ziel bleibt die weitestgehende Interoperabilität der verschiedenen Metadatensysteme. Für die Definition dieser Anwendungsprofile ist bei dem CEN-Workshop »Metadata – Dublin Core« (MMI-DC) eine Anleitung in Vorbereitung, die zur Erleichterung und Vereinheitlichung der Beschreibung dienen soll.<sup>7</sup>

Die Verwaltung dieser Strukturen ist noch nicht umfassend geklärt, aber es ist klar, dass Registries hier eine entscheidende Rolle spielen werden: einerseits als Referenz für die benutzten Datenelemente, andererseits aber auch als Vermittler, um die verschiedenen Systeme miteinander in Beziehung zu setzen mit dem Ziel, auch aus nicht verwendeten Elementen möglichst Informationen für den eigenen Kontext zu erschließen. Während dies bei einigen an Dublin Core angelehnten Konstrukten wie »Contributor.Illustrator« relativ ein-

**Standardprofile  
für verschiedene  
Interessengruppen**

fach ist, lässt sich das Datenelement »SSGFI.Country«, das z.B. im MetaGuide benutzt wird, so nicht erschließen sondern benötigt eine eigene Definition.

### AUTOMATISIERTER DATENAUSTAUSCH

Das andere Problem berührt die Nutzung der bestehenden und künftigen Registries. Bis jetzt gibt es praktisch keine maschinenlesbaren Informationen in Registries. Das ist aber die Bedingung dafür, dass Überlegungen, wie sie mit dem Konzept des »Semantic Web«<sup>8</sup> verbunden sind, einer Realisierung näher kommen. Um im Internet beliebige vorgefundene Dokumente »verstehen« zu können, müsste ein Programm (z.B. ein Agent im Sinne von Berners-Lee) deren Bedeutung aus strukturierten Informationen extrahieren, d.h., dass hier Metadaten (allgemeiner: semantische Auszeichnungen) eine zentrale Rolle spielen werden. Da ein einziger zentraler Standard hierfür nicht zu erwarten (wohl nicht einmal zu wünschen) ist, müssen verschiedene Standards miteinander in Beziehung gesetzt werden: eine Aufgabe für maschinenlesbare Registries.

In dieser Richtung arbeitet auch eine Arbeitsgruppe der International Organization for Standardization (ISO) und der International Electrotechnical Commission (IEC), der es um einheitliche Identifikation und Beschreibung von Metadaten und die Erleichterung ihrer Nachnutzung und ihres Austauschs geht. Die Arbeit an »ISO/IEC 11179: Information Technology – Metadata Registries« ist aber noch nicht abgeschlossen und nicht (ganz?) öffentlich.<sup>9</sup> Ebenso wurde von der »Organization for the Advancement of Structured Information Standards« (OASIS) im Kontext von E-Busi-

ness und XML ein Standard zur Beschreibung von Registries entwickelt.<sup>10</sup> Diese Entwicklungen finden allerdings in einem etwas anderen Kontext als dem hier beschriebenen statt (vgl. dazu z.B. das »Open Forum on Metadata Registries«, 20.–24.1. 2003, in Santa Fe, New Mexico, USA; URL: <http://metadata-stds.org/OpenForum2003/>).

Für die weitestgehende Ausnutzung bestehender Informationen und eine möglichst umfassende Interoperabilität wird es nicht ausreichen, einfache Konkordanzen aufzubauen (was schon schwierig genug ist), sondern es gilt komplexe kontextabhängige Beziehungen zwischen unterschiedlichen Begriffssystemen zu entwickeln, was derzeit unter dem Schlagwort »Ontologien« diskutiert wird. Das weist aber weit in die Zukunft hinein und wird unter anderem im 6. Rahmenprogramm der EU zu den Forschungsschwerpunkten gehören.<sup>11</sup>

Derzeit sollten Registries etwas bescheidenere Anforderungen erfüllen:

- Sie sollten Datenmodelle (im XML-Sprachgebrauch: Namensraum, »namespace«) speichern und als autoritative Quelle dienen (Beispiel: DC-Registry).
- Sie sollten Anwendungsprofile nachweisen oder speichern, miteinander in Beziehung setzen und ihre Verwaltung unterstützen. Insbesondere wäre es hilfreich, standardisierte Formate von Anwendungsprofilen zu haben, die in solch ein Registry automatisch eingelesen werden könnten.
- Sie sollten bestehende Konkordanzen speichern und die Entwicklung neuer Konkordanzen unterstützen.

In der vom CORES-Projekt initiierten »CORES Resolution«<sup>12</sup> sichern Vertreter von Metadatenschemata zu, für die Metadatenelemente eindeutige Identifikatoren (URIs) bereitzustellen. Dies ist ein wichtiger Schritt auf dem Weg zur Interoperabilität: Nur eindeutige Metadatenelemente können auch eindeutig miteinander in Beziehung gesetzt werden.

Für das DC-Registry ist ein Weg (»roadmap«) zu einem Registry von Harry Wagner, Rachel Heery und Stu Weibel entworfen worden, in dem die einzelnen Bausteine eines solchen Systems benannt werden (vgl. Abb. 4).<sup>13</sup>

### OFFENE FRAGEN

Die damit verbundenen Fragen sind zurzeit Thema intensiver Diskussionen.

#### Wie können Anwendungsprofile effektiv beschrieben werden?

Sowohl für die Erstellung als auch für die Verarbeitung von Anwendungsprofilen wären Standards hilfreich.

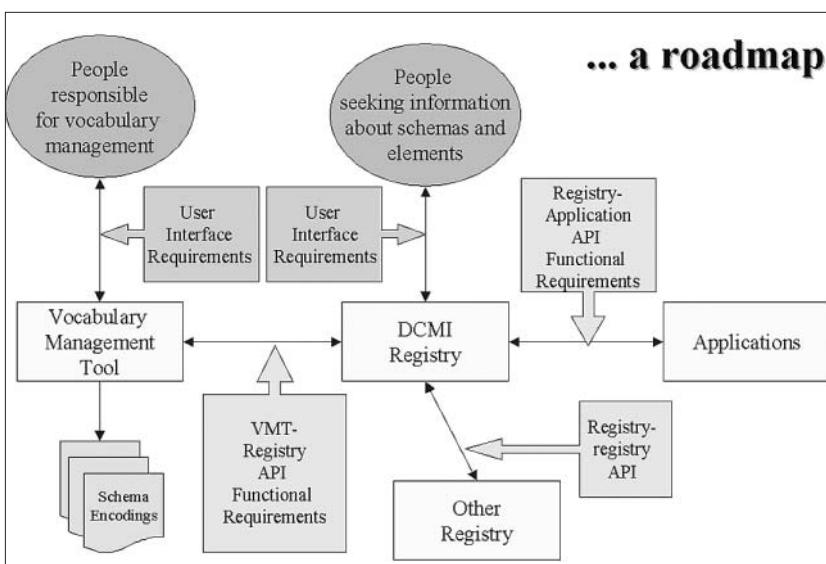


Abb. 4: ... a roadmap: Skizze aus: Harry Wagner, Rachel Heery: DCMI Roadmap for development of Vocabulary Management and Schema Registry Systems.

Besser noch wäre eine so weit reichende Formalisierung, dass das Profil maschinenlesbar wird. Von Seiten des Registries müsste umgekehrt eine allgemeine Schnittstelle von Anwendungsprofilen zu Registries definiert werden.

### **Wie und wo werden Klassifikationen beschrieben?**

Klassifikationen, Kodierungen, Thesauri oder auch Ontologien müssen effektiv (maschinenlesbar) beschrieben werden, damit sie auch automatisch ausgelesen werden können. Es ist derzeit nicht klar, ob dies auch eine Aufgabe für Registries ist oder dafür andere Strukturen nötig sind. Derzeit passen Klassifikations-schemata (wie z.B. die Dewey Decimal Classification) nicht in das XML-orientierte Schema von Namespaces. Es sind auch noch keine eindeutigen Identifikatoren (URIs) festgelegt. Ein zusätzliches Problem entsteht, da nicht alle Klassifikationen frei zugänglich sind.

### **Wie funktionieren Vergröberungsabbildungen?**

Es erscheint sinnvoll, die Vergröberungsabbildung (traditionell im Dublin-Core-Umfeld als »dumb down« bezeichnet), präziser zu fassen und von der Bezeichnung der Metadatenelemente zu trennen. Das würde die Semantik der Bezeichnungen vereinfachen, da dann z.B. DC.Creator.Birthdate eine sinnvolle Konstruktion wäre. Die Vergröberung müsste dann aber auf DC.Date und nicht auf DC.Creator führen. Da also diese Abbildung nicht mehr in der Bezeichnung des Datenelementes »kodiert« ist, muss die Information an anderer Stelle gespeichert werden, sinnvollerweise in einem Anwendungsprofil und dem Registry.

### **Gibt es eine allgemeine Sprache zur Beschreibung von Beziehungen zwischen Metadaten(systemen)?**

Im Sinne der Interoperabilität geht es letztlich darum, auch aus fremden Metadaten »Sinn zu machen«. Das heißt, dass fremde Daten möglichst in das eigene System übersetzt werden müssen. Bei präziser Übereinstimmung ginge dies mit einer Konkordanz; oft werden jedoch die verschiedenen Datensysteme nicht genau passen. Dann würde eine vorgesetzte Vergröberung (siehe oben) vielleicht zu einem Element führen, das direkt verstanden oder übersetzt werden kann. In jedem Falle würde die Erstellung von Übersetzungsschemata von jedem Anwendungsprofil in jedes andere eine sehr komplizierte und umfangreiche Aufgabe sein. Einfacher wäre es, wenn sich alle Anwendungsprofile auf ein universelles System (wie das »Semantic Layer« im DESIRE Registry) beziehen könnten. Dann müsste ein neu hinzugefügtes Anwendungsprofil nur mit einer Abbildung zu diesem universellen Sys-

**Interoperabilität von Metadaten**

tem ausgestattet werden und nicht mit je einer Abbildung für jedes bereits vorhandene Profil.

## A U S B L I C K

Im Idealfall wird das zukünftige Registry automatisch und klug reagieren. Das Szenario dazu sähe so aus:

Ein Programm (Tim Berners-Lee's »Agent«) ist im Internet auf der Suche nach Informationen. Wenn es auf ein Dokument stößt, so liest es die zugehörigen Metadaten, die im Dokument selbst stehen können oder an anderer Stelle, auf die im Dokument verwiesen wird.<sup>14</sup> Jetzt wertet es die Metadaten aus:

- Bekannte Elemente werden übernommen (z.B. dc.Title).
- Danach wird das Anwendungsprofil ausgewertet, auf das in den Metadaten verwiesen wird. Geprüft wird, ob das Element mit einem im eigenen System bekannten Element in Beziehung steht: Gibt es für das Metadatenelement ein äquivalentes bekanntes Element, so wird dieses genommen, ansonsten wird nach möglichen Vergrößerungen gesucht. Wird eine Vergrößerung auf ein bekanntes Element gefunden, so wird dieses ausgewertet, wird eine Vergrößerung auf ein äquivalentes Element gefunden, so wird das verwendet.
- Wenn das Anwendungsprofil keine dieser Möglichkeiten liefert, wird das Element verworfen.
- Alternativ wird das Registry aufgerufen: Das eigene und das vorgefundene Metadatenprofil sind in einem Registry gespeichert. Die dort vorliegenden Abbildungen werden für die vorgefundenen Metadaten ausgewertet, nicht verwertbare Daten werden verworfen.

Bis dies soweit ist, wird noch einige Zeit vergehen. Die Zwischenzeit kann sinnvoll damit verbracht werden, Metadaten zu erzeugen bzw. die Erzeugung von Metadaten zu erleichtern. Denn wie oben schon zu sehen ist: Ohne Metadaten zu dem gefundenen Dokument bricht das schöne Zukunftsszenario schon auf der ersten Stufe zusammen.

### Links zu Registries

- CORES: <http://www.cores-eu.net/registry/>  
DESIRE: <http://desire.ukoln.ac.uk/registry/index.php3>  
Diffuse: <http://www.diffuse.org/standards.html> (Registry in allgemeinerem Sinn)  
Dublin Core: <http://dublincore.org/dcregistry/>  
MetaLib: <http://www2.sub.uni-goettingen.de/>  
MMI-DC: <http://www.sub.uni-goettingen.de/ssgfi/observatory/>  
ROADS: <http://www.ukoln.ac.uk/metadata/roads/templates/>

SCHEMAS: <http://www.schemas-forum.org/registry/design/index.php3>

<sup>1</sup> Dies ist die erweiterte Fassung eines Vortrags gleichen Titels, der am 22. Oktober 2002 auf dem 1. Metadaten-Workshop der Arbeitsstelle für Standardisierung, gleichzeitig Abschluss-Workshop des Projektes META-LIB, an der Deutschen Bibliothek in Frankfurt gehalten wurde.

<sup>2</sup> Dublin Core Metadata Initiative. URL: <http://dublincore.org/>

<sup>3</sup> Der Status ist derzeit der von »Proposed Recommendations«, siehe: Dublin Core Metadata Initiative: DCMI Recommendations. URL: <http://dublincore.org/documents/#recommendations>

<sup>4</sup> Tom Baker spricht in diesem Zusammenhang gerne von einer »Pidgin«-Sprache für Metadaten (»A Pidgin for Digital Tourists«). Baker, Thomas: A Grammar of Dublin Core. In: D-Lib Magazine Volume 6 Number 10 October [2000]. URL: <http://www.dlib.org/dlib/october00/baker/tobaker.html>, und zieht aus dieser Metapher einige interessante linguistische Konsequenzen. Mir klingt der Begriff der »Lingua Franca« positiver.

<sup>5</sup> Heery, Rachel; Patel, Manjula: Application profiles: mixing and matching metadata schemas. In: Ariadne Issue 25 September (2000). URL: <http://www.ariadne.ac.uk/issue25/app-profiles>

<sup>6</sup> Für Bibliotheken gibt es das DC-Library Application Profile: DC-Lib, URL: <http://dublincore.org/documents/library-application-profile/>, für die öffentliche Verwaltung wird an einem DC-Government Application Profile (DC-Gov) gearbeitet, siehe URL: <http://dublincore.org/groups/government/>. Weitere thematische Anwendungsprofile sind in Vorbereitung.

<sup>7</sup> Baker, Thomas; Dekkers, Makx; Fischer, Thomas; Heery, Rachel: Dublin Core Application Profile Guidelines. (In Vorbereitung)

<sup>8</sup> Berners-Lee, Tim; Hendler, James; Lassila, Ora: The Semantic Web: A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities. In: Scientific American May (2001). URL: [http://www.sciam.com/print\\_version.cfm?articleID=00048144-10D2-1C70-84A9809EC588EF21](http://www.sciam.com/print_version.cfm?articleID=00048144-10D2-1C70-84A9809EC588EF21)

<sup>9</sup> ISO/IEC JTC1 SC32 WG2 Development/Maintenance: ISO/IEC 11179, Information Technology – Metadata Registries (MDR). URL: <http://metadata-stds.org/11179/index.html>. Eher technische Fragen zu diesem Komplex werden im Zusammenhang mit »ISO/IEC 20944, Information Technology – Metadata Registry Interoperability and Bindings« bearbeitet; URL: <http://metadata-stds.org/20944/index.html>

<sup>10</sup> OASIS/ebXML Registry Information Model v2.1 June (2002). URL: [http://www.oasis-open.org/committees/regrep/documents/2.1/specs/ebrim\\_v2.1.pdf](http://www.oasis-open.org/committees/regrep/documents/2.1/specs/ebrim_v2.1.pdf)

<sup>11</sup> Knowledge and interface technologies: 2.3.1.7 Semantic-based knowledge systems. URL: <http://www.cordis.lu/ist/so/knowledge/home.html>

<sup>12</sup> CORES Standards Interoperability Forum (Ed.: Thomas Baker, Makx Dekkers): Resolution on Metadata Element Identifiers. URL: <http://www.cores-eu.net/interoperability/cores-resolution/cores-resolution.pdf>

<sup>13</sup> Wagner, Harry; Heery, Rachel: DCMI Roadmap for development of Vocabulary Management and Schema Registry Systems. URL: <http://dublincore.org/groups/registry/DCMI-reg-roadmapv4.html>

<sup>14</sup> Ob eine automatische Erzeugung von Metadaten funktionieren kann, ist derzeit umstritten, hängt aber auf jeden Fall von der Komplexität des Metadatenschemas ab; einfache Informationen können heute schon automatisch extrahiert werden.

## D E R V E R F A S S E R

**Dr. Thomas Fischer** arbeitet an der SUB Göttingen im Bereich der digitalen Bibliothek. Schwerpunkt seiner Tätigkeit sind Metadaten für digitale Objekte: Katalogisierung von Internetressourcen und elektronischen Dokumenten, Metadatensysteme, Anwendungsprofile etc. Niedersächsische Staats- und Universitätsbibliothek, Historisches Gebäude, Papendiek 14, 37073 Göttingen  
[fischer@mail.sub.uni-goettingen.de](mailto:fischer@mail.sub.uni-goettingen.de)