



World Library and Information Congress: 69th IFLA General Conference and Council

1-9 August 2003, Berlin

Code Number: 010-G
Meeting: 126. Classification and Indexing
Simultaneous -
Interpretation:

FAST (Facettierte Anwendung der Schlagwort-Terminologie - Faceted Application of Subject Terminology): Ein vereinfachtes Vokabular auf der Basis der LCSH

Edward T. O'Neill

OCLC, Dublin
USA

and

Lois Mai Chan

University of Kentucky, Lexington
USA

Zusammenfassung

Für eine allgemeine Anwendung ist das Schlagwortsystem der Library of Congress das bei Weitem meistgenutzte und anerkannte Erschließungsvokabular. In der Tat handelt es sich um ein universeller, kontrolliertes Vokabular, das in der ganzen Welt als Vorbild für die Entwicklung neuer Sacherschließungssysteme gedient hat. Wie auch immer, die komplexe Syntax der LCSH sowie deren Regeln für die Bildung von Sacherschließungsergebnissen beschränken die Anwendung, da sie hochspezialisiertes Personal benötigen, und sie beschränken weiterhin die Effektivität einer automatisierten Kontrolle der Normdaten.

Neue Trends, bedingt durch die weite Verbreitung und durch das schnelle Wachstum des WEB, drängen auf eine Änderung der Systeme der bibliographischen Kontrolle mit dem Ziel, diese Systeme einfacher nutzbar, verständlich und anwendbar zu machen. Dies gilt auch für die Sacherschließung. Das Schema von FAST ist mit den LCSH kompatibel, und jeder gültige Datensatz der LCSH kann in einen FAST-Satz umgewandelt werden.

FAST besteht aus acht unterschiedlichen Facetten. Bis auf die Facette der Zeit sind für alle bestehenden Schlagwörter Normdatensätze gebildet worden. Die Startversion der FAST-Normdatei wird etwa zwei Millionen Normdatensätze enthalten.

Einführung

Die außergewöhnliche Zunahme elektronischer Ressourcen und die Ausbildung zahlreicher Metadaten-Erfassungsschemata zu deren Beschreibung erfordert insbesondere für den sachlichen Zugriff Methoden, die eine große Anzahl an Dokumenten bewältigen ohne dass Aufwand und Kosten in gleichem Umfang anfallen wie bei der Bearbeitung von traditionellen Bibliotheksmaterialien. Im Rahmen der Suche nach einem sachlichen Zugangssystem, das die Verfahren des Dublin Core optimal nutzt, beschritt OCLC 1998 einen neuen Weg zu einem Sacherschließungs-Vokabular. Unter den Gegebenheiten des Dublin Core musste dieses für das WEB geeignete Erschließungsvokabular folgenden Anforderungen genügen:

- Es sollte eine einfache Struktur haben (d.h. leicht zu bestimmen und zu nutzen) und leicht zu pflegen sein;
- Es sollte einen optimalen Zugriff bieten;
- Es sollte flexibel, interdisziplinär und in allen Wissensgebieten anwendbar sein, und es sollte einen Zugang zu Online-Umgebungen herstellen einschließlich der Online-Kataloge (OPAC).

Die Entwicklung eines Erschließungsverfahrens, welches den genannten Bedingungen genügt, bedarf zweier grundsätzlicher Vorentscheidungen: (1) Festlegung der Semantik (Wahl des Wortschatzes) und (2) Gestaltung der Syntax (Prä- versus Postkoordination). Bezüglich der Semantik entschloss sich OCLC zu LCSH. Mit der Übernahme der LCSH war zugleich die Kompatibilität mit diesen gewährleistet. Als Erschließungs- bzw. Dokumentationssprache bietet LCSH einige Vorteile:

- Das umfangreiche Vokabular deckt alle Sachgebiete ab;
- Es ist die umfangreichste universelle Dokumentationssprache in englischer Sprache;
- Synonymie- und Homonymiekontrolle ist gesichert;
- Zwischen den Deskriptoren bestehen umfangreiche Verweisungen;
- Es handelt sich faktisch um ein universelles kontrollierte Vokabular, welches in weltweit in vielen Ländern übersetzt oder als Modell für die Entwicklung eigener Sacherschließungssysteme adaptiert wurde.
- Es ist mit MARC-Aufnahmen kompatibel;
- Eine automatische Umsetzung in das neue Schema ist bei Allgemeinbegriffen möglich;
- Die Unterhaltungskosten des neuen Verfahrens sind nach unten gedrückt seit die meisten Änderungen hinsichtlich LCSH automatisch in das neue Schema überführt werden können.

Während das umfangreiche Vokabular und die semantischen Relationen in den LCSH einen sachlichen Zugang über die Möglichkeiten von einfachen Suchwörtern

hinaus eröffnen, ist ihre komplexe Syntax häufig ein Hindernis und ein Widerspruch der grundsätzlichen Voraussetzung des Dublin Core bezüglich Einfachheit und semantischer Durchlässigkeit. OCLC entschied sich für eine mit dem LCSH Vokabular zu verwendende vereinfachte Syntax. Als Ergebnis liegt ein kontrolliertes Vokabular vor, welches auf der Terminologie und deren bereits in den LCSH enthaltenen Relationen aufbaut, dies allerdings mit einer veränderten Syntax strukturiert und mit veränderten Richtlinien und Verfahren angewendet, die insgesamt mehr in die Richtung der Postkoordination weisen. Durch die Trennung von Syntax und Semantik kann der Anwendungsvorgang vereinfacht und dennoch die Fülle des LCSH-Vokabulars erhalten werden, was Pflege und Anwendung dieses Verfahrens vereinfacht. Mit der vereinfachten Syntax und den vereinfachten Anwendungsregeln kann die Informationstechnologie vorteilhafter eingesetzt werden und zwar sowohl für die Bildung und Pflege der Schlagwörter als auch für die Kontrolle der sachlichen Normdaten.

Während die LCSH über ein Jahrhundert lang Bibliotheken und ihren Nutzern gedient haben, lässt ihre Komplexität die Anwendung über den Bereich der traditionellen Kataloge hinaus kaum zu. Die LCSH waren für Zettelkataloge bestimmt und in diesem Umfeld sicher bedeutend. Da nun aber der Raum auf einer 3 mal 5 Karte begrenzt ist und jedes neue gedruckte Schlagwort eine neue Katalogkarte benötigt, war die Anzahl der pro Einheit zugelassenen Schlagwörter streng begrenzt. Da außerdem der Zettelkatalog mit einer Postkoordination unvereinbar war, blieb in dieser Umgebung die Präkoordination die einzig gültige Möglichkeit.

Im Sinne einer umfassenden Liste aller gültigen Deskriptoren sind die LCSH kein wirklicher Thesaurus. Vielmehr kombiniert LCSH Deskriptoren, in nunmehr fünf Bänden in gedruckter Form enthalten. Ein vierbändiges Handbuch enthält die Regeln mit genauen Erläuterungen der Anforderungen für die Bildung neuer, nicht in der Normdatei enthaltener Schlagwörter, und für die weitere Unterteilung zahlreicher vorhandener Schlagwörter.

Die Regeln für den Gebrauch frei verfügbarer Unterschlagwörter, die über Musterschlagwörter geregelt sind, erhellen einige der vorhandenen Komplexitäten. Unter ganz bestimmten Voraussetzungen können diese freien Unterschlagwörter vorhandenen Schlagwörtern beigeordnet werden. Der Anwendungsbereich ist häufig durch bestimmte Arten (Muster) von Schlagwörtern begrenzt. So ist zum Beispiel **Burns and scalds—Patients—Family relationships** ein zugelassener Schlagworteintrag, der Anhängen zweier freier Unterschlagwörter an das vorhandene Hauptschlagwort **Burns and scalds**. Das Unterschlagwort **Patients** ist eines von etlichen hundert, die in Verbindung mit Schlagwörtern für Krankheiten und anderen medizinischen Sachverhalten verwendet werden können. Es kann also als Unterschlagwort an das Hauptschlagwort **Burns and scalds** angehängt werden. Wie auch immer, die Hinzufügung von **Patients** verschiebt die Bedeutung der Beschlagwortung von einem medizinischen Sachverhalt zu einer Personengruppe. Da nun **Family relationships** in Verbindung mit Personengruppen zugelassen ist, kann es auch zur Vervollständigung der Beschlagwortung benutzt werden.

Andere Beispiele für einige der komplexen Besonderheiten seien illustriert durch einen Typus von Schlagwörtern, der als „multipel“ bekannt ist. In den LCSH ist eine Gruppe von gleichartigen Schlagwörtern als „multiple Unterschlagwörter“ („multiple subdivision“) eingerichtet, d.h. „ein Unterschlagwort in der Schlagwortnormdatei mit

in eckige Klammern gestellten Bezeichnungen, denen gewöhnlich das Wort **etc.** folgt, um zu verdeutlichen, dass gleichartige Unterschlagwörter gemäß dem vorliegenden Schlagwort gebildet werden können“. Mit anderen Worten, als multipel werden Schlagwörter bezeichnet, die ein Verwendungsmuster darstellen. So zeigt das multiple Unterschlagwort bei **Love—Religious aspects—Buddhism, [Christianity, etc.]**, dass die Bezeichnungen für andere Religionen „Buddhism“ ersetzen können, um so einen neuen Schlagworteintrag zu kreieren. Das derzeit im 1xx Feld des Datensatzes erscheinende multiple Schlagwort sollte nie in seiner multiplen Form in einem bibliographischen Datensatz benutzt werden. All die als multipel möglichen Schlagwörter sind in den LCSH nicht verzeichnet, und es gibt dort auch keine vollständige Liste der Religionen.

Ein dritter Bereich, der den Grad der Komplexität zeigt, ist der Bereich der Musik. Einige dieser Besonderheiten schließen ein: Bestimmung der Zugehörigkeit jedes Soloinstrumentes (z.B. Blasinstrumente), Anordnung der Instrumente innerhalb der jeweiligen Gruppen, und ob ein Schlagwort qualifiziert werden soll (z.B. Concertos). Insgesamt gibt es innerhalb der Schlagwörter für die Musik die größte Anzahl gültiger Schlagwörter, die nicht in den LCSH verzeichnet sind.

Während das umfangreiche Vokabular und die semantischen Relationen in den LCSH einen sachlichen Zugang über die Möglichkeiten von einfachen Suchwörtern hinaus eröffnen, verhindert ihre komplexe Syntax ihre Anwendung über die Welt der traditionellen Bibliothekskataloge hinaus. Nicht nur sind die Regeln für die Musterschlagwörter komplex, auch ihre Verwendung erfordert ein umfangreiches Wissen über das jeweilige Fachgebiet, da es keinerlei expliziten Hinweise dafür gibt, welche frei verfügbaren Unterschlagwörter für bestimmte Hauptschlagwörter passend sind.

Zwar wird FAST Schlagwörter, die gemäß dieser Regeln zugelassen sind, aufnehmen, allerdings werden diese so in eine Normdatei aufgenommen, dass die komplexen Regeln, denen sie ihre Existenz verdanken, nicht erscheinen. Auf diese Weise basiert FAST auf dem Vokabular der LCSH, allerdings mit vereinfachten Syntax, verglichen mit dem, was derzeit von den Bibliotheken entsprechend den Richtlinien der Library of Congress angewendet wird.

SYNTAX

Aus Gründen der Einfachheit und der semantischen Interoperabilität wurde für FAST ein postkoordiniertes Verfahren gewählt. Dies entspricht den grundsätzlichen Voraussetzungen und Charakteristiken automatisierter Retrievalsysteme. Dies entspricht darüber hinaus der ursprünglichen Absicht des Dublin Core, nämlich der „Nutzung durch Nicht-Katalogisierer ebenso wie durch diejenigen, die über Erfahrung mit Modellen formaler Dokumentbeschreibung verfügen“.

In diesem Sinne ist das FAST-Schema

- ein kontrolliertes Vokabular dessen Deskriptoren alle in einer Normdatei verzeichnet sind, mit Ausnahme numerischer Werte,
- basierend auf dem LCSH-Vokabular,
- für eine Online-Umgebung entworfen,
- ein postkoordiniertes, facettiertes Vokabular,
- für Anwender mit geringer Übung und Erfahrung,
- mit automatisierter Normdatenkontrolle vereinbar.

FACETTEN

FAST besteht aus acht unterschiedlichen Facetten: Sache, Geographikum (Ort), Personenname, Körperschaftsname, Form (Typus, Art), Zeit (Zeitabschnitt, Zeitalter), Titel, und Veranstaltungsname. Die Facetten Person und Körperschaft sind in ihrem Gebrauch für sachliche Inhalte beschränkt und enthalten keine für Name-Titel-Eintragungen. Titel und Veranstaltungen werden nicht in der Startphase der Entwicklung mit einbezogen. Es besteht jedoch Einigkeit darüber, dass diese Facetten als notwendige Zugriffsmöglichkeiten zu gelten haben, sie werden in der nächsten Phase der FAST-Entwicklung berücksichtigt.

Die Grundlage für die Bestimmung der benötigten Deskriptoren ist der Literaturanfall. Theoretisch kann es zu einer unendlichen Anzahl an gültigen LCSH-Deskriptoren kommen. Gehen wir z.B. davon aus, 175 unterschiedliche Instrumente seien identifiziert. Wenn wir alle denkbaren Kombinationen für drei Instrumente bilden, so führt das theoretisch zu einer Million an einzelnen Schlagwörtern für Trios. Für Nonette sind eine Billion, also 10^{18}

Kombinationen für neun Instrumente möglich. Die meisten dieser Kombinationen werden offenkundig niemals benutzt werden – viele werden auch nicht einmal logisch sinnvoll sein. Deshalb ist es weder sinnvoll noch machbar, für jeden denkbaren Fall ein Schlagwort zu bilden. Die Bildung eines bestimmten FAST-Deskriptors ist abhängig von seiner Verwendung in WorldCat von OCLC, wobei alle von der Library of Congress gebildeten Deskriptoren eingeschlossen sind. Deskriptoren ohne jede Verwendung in WorldCat werden nicht in FAST aufgenommen, auch wenn sie gültig sein sollten.

FAST führt die Nutzung von Unterschlagwörtern fort und bewahrt die hierarchische Struktur der LCSH. Der Hauptunterschied zu LCSH besteht darin, dass bei bestimmten FAST-Schlagwörtern die Unterschlagwörter zur selben Facette wie die Hauptschlagwörter gehören müssen.. Sachs Schlagwörter können also nur durch andere Sachs Schlagwörter, geographische Schlagwörter nur durch geographische Unterschlagwörter unterteilt werden usw. Ein bestimmtes Hauptschlagwort kann also nicht durch Unterschlagwörter anderer Facetten unterteilt werden.

SACHSCHLAGWORT

Die Sachs Schlagwörter in FAST bestehen aus sachlichen Hauptschlagwörtern mit geeigneten sachlichen Unterteilungen, einschließlich jener sachlichen Unterteilungen, die in LCSH als Namen und als geographische Schlagwörter verzeichnet sind. Die Sachs Schlagwörter von FAST sind ganz ähnlich der eingeführten Form der LCSH-Schlagwörter, zum Beispiel:

Education

Natural gas pipelines—Economic aspects

Photoconductivity—Measurement

Travel—Safety measures

Urbanization

Hospitals—Staff—Labor unions—Organizing

Ein Unterschied zwischen LCSH und FAST besteht darin, dass alle frei verfügbaren sachlichen Unterschlagwörter Teil der vorgegebenen Schlagwortform werden und

dass alle Multiplen Unterschlagwörter expandiert werden. Die Schlagwortelemente innerhalb derselben Facette sind im Rahmen der von der Library of Congress vorgegebenen Richtlinien präkombiniert. Es werden allerdings nur diejenigen Schlagwörter übernommen, die aktuell benötigt werden. So werden in die Normdatei Schlagwörter aufgenommen, die auf dem folgenden Schlagwort mit multiplem Unterschlagwort basieren. Beispiel:

Love—Religious aspects—Buddhism, [Christianity, etc.]

In FAST werden keine multiplen Unterschlagwörter benutzt. Jede Kombination aus **Love—Religious aspects** und einer Religion wird, sofern in WorldCat benutzt, wird einzeln aufgenommen wie etwa

Love—Religious aspects—Buddhism
Love—Religious aspects—Christianity
Love—Religious aspects—Islam
Love—Religious aspects—Hinduism
etc.

Auf jeden Fall werden derartige Schlagwörter nicht für alle bekannten Religionen sondern nur für die jeweils benutzten gebildet.

GEOGRAPHIKA

Geographische Namen werden in einer indirekten Form angesetzt und genutzt. **Germany—Berlin** ist zum Beispiel die an Stelle der direkten Reihenfolge **Berlin (Germany)** zu verwenden. Ortsnamen werden als Hauptschlagwörter in den LCSH in direkter Folge eingegeben, werden sie aber als Unterschlagwörter benutzt, so erscheinen sie in der umgekehrten Folge. Geographische Namen des ersten Niveaus sind in FAST wesentlich begrenzter als in LCSH und sie sind eingeschränkt auf die Namen der Tafel für die *Geographic Area Codes*. Die Verbindung dieser geographischen Namen mit den Codes für geographische Räume bietet eine zusätzliche Spezifität und gibt den Schlagwörtern eine hierarchische Struktur. So können die geographischen Codes zur Eingrenzung einer Suche benutzt werden.

Einige geographische Schlagwörter sind in ihrer direkten oder indirekten Form deutlich unterschieden. In den LCSH ist so North Carolina als Haupt- oder Unterschlagwort ausgeschrieben, als Namenszusatz ist es jedoch als N.C. abgekürzt. Um sicher zu sein, dass alle Vorkommen des Namens berücksichtigt sind, müssen Benutzer oft mit beiden Namensformen recherchieren. Eine vollständige Suche nach Washington in North Carolina erfordert beispielsweise Suchvorgänge sowohl für **Washington (N.C.)** als auch für **North Carolina—Washington**. Bei der einfachen Frage allein mit dem Namen der Stadt, Washington, North Carolina, wird das bei Weitem meiste Material zu Washington, D.C. nicht aufgefunden. Auf der anderen Seite sind in FAST genaue Recherchen einfacher, da nur mit ein Namensform nötig ist, um das gesamte Material zu erhalten.

Namen des zweiten Niveaus werden als Unterschlagwörter aufgeführt und zwar unter dem Namen des kleinsten Geographikums auf dem ersten Niveau, in welchem sie vollständig enthalten sind. Dies geschieht, um die Anzahl der Namen des ersten Niveaus zu begrenzen, um sicher zu stellen, dass diese Namen allgemein verständlich sind, und um den Namen des zweiten Niveaus in einen umfassenderen Zusammenhang zu setzen. Die Curzon-Linie, die vorgeschlagene Demarkationslinie zwischen Polen und der Sowjetunion während des russisch-polnischen Kriegs 1919-1920 zum Beispiel, wird für FAST als **Europe—Curzon Line** angesetzt. Hingegen in den LCSH wird diese Linie ohne weitere Zusätze schlicht als **Curzon Line** aufgenommen, da sie nationale Grenzen überschreitet.

Geographische Namen sind grundsätzlich auf zwei Ebenen beschränkt, mit Ausnahme der innerstädtischen Ortsnamen. So ist zum Beispiel der Teil Dravlje von Ljubljana als **Slovenia—Ljubljana—Dravlje** angesetzt. Namenszusätze werden benutzt sowohl um die Art der geographischen Benennung zu erläutern (County, Lake, Kingdom, Princely State, usw.) als auch um als identifizierender Zusatz für den Fall, dass es den Namen mehrfach gibt. Für die USA ist der Name des County die gebräuchliche Form, falls dieser Name innerhalb eines Bundesstaates mehrfach vorkommt. So gibt es in Michigan zwei Beaver Islands; größer und bekannter sind die Inseln im Michigansee, doch gibt es auch ein Beaver Island im Oberen See. Um die Insel im Michigansee zu individualisieren, erhält sie die Bezeichnung des Countys als Zusatz: **Michigan—Beaver Island (Charlevoix County)**. Haben unterschiedliche Arten von Geographika gleiche Namen, so drückt der Zusatz die Art des jeweiligen Geographikums aus. Beispielsweise ist Otsego Lake sowohl der Name einer Stadt als auch eines Sees in Michigan. Zur Unterscheidung sind die Namen als **Michigan—Otsego Lake** und **Michigan—Otsego Lake (Lake)** aufgenommen. Folgt man der Praxis der LCSH so müssen bewohnte geographische Räume nicht mit einem Zusatz versehen werden.

Einige Beispiele für geographische Schlagwörter in FAST nebst ihren zugehörigen geographischen Codes sind:

Great Lakes [nl]

Germany [e-gx]

Mars [zma]

Maryland—Worcester County [n-us-md]

Slovenia—Maribor [e-xv]

France—Loir River Valley [e-fr]

England—Chilton (Oxfordshire) [e-uk-en]

India—Limbdī (Princely State) [a-ii]

Californien—San Francisco—Chinatown [n-us-ca]

Personennamen und Namen von Körperschaften

Personennamen einschließlich Familiennamen und Körperschaftsnamen sind beide abgeleitet aus der NACO Normdatei. Die Voraussetzung dafür dass Personen- oder Körperschaftsnamen in FAST aufgenommen werden, bestehen darin, dass (1) der Name mindestens ein Mal in einer WorldCat-Aufnahme verwendet wurde und dass (2) der Name in der Normdatei der Namen enthalten und als Schlagwort zugelassen ist. Bei Personennamen werden zusätzlich zur reinen Namensform

Jahreszahlen, Titel und vollständige Schreibweisen erfasst, um den Namen anzusetzen. Bei Körperschaftsnamen werden die Bezeichnungen aller untergeordneter Einheiten in die Ansetzung eines FAST-Körperschaftsschlagworts aufgenommen. Diese Körperschaftsnamen beinhalten sowohl jene des Rechts als auch sonstige Vereinigungen. Name – Titel – Ansetzungen wie etwa **Smollett, Tobias George, 1721-1771. Expedition of Humphry Clinker** sind weder in der Facette für Personen noch in der für Körperschaften zugelassen. Sie werden in einer zweiten Phase der FAST-Entwicklung berücksichtigt. Einige Beispiele aus FAST für Personenschlagwörter:

Woodward, Bob
Dewey, Melvil, 1851-1931
Kennedy family
Edward II, King of England, 1284-1327
Bush, George W. (George Walker), 1946-

FAST Körperschaftsnamen:

OCLC
Bayerische Motoren Werke
United States. Coast Guard
Bodleian Library

FORMSCHLAGWÖRTER

Angaben für die Form oder das Genre sind in einer eigenen Facette festgehalten. Formschlagwörter werden aus verschiedenen Quellen entnommen. Zahlreiche Formschlagwörter werden aus den Unterschlagwörtern der LCSH-Normdaten und den bei Titelaufnahmen verwendeten Schlagwörtern gewonnen. Einzelheiten dieses Arbeitsvorgangs sind genauer beschrieben bei O'Neil et. al. Einige Beispiele für Formschlagwörter:

Case studies
Abstracts
Census
Rules
Dictionaries—Portuguese
Folklore
Bibliography—Union lists
Periodicals

Zeitschlagwörter

Die Zeitschlagwörter in FAST folgen den Empfehlungen der SAC/ALCTS Unterkommission und der Airlie Konferenz. FAST-Zeitschlagwörter berücksichtigen die auf das jeweilige Dokument bezogenen Zeitabschnitte und sind nicht beschränkt auf besondere Perioden, die mit bestimmten Einzelereignissen einhergehen. Zeitschlagwörter werden entweder als einzelne numerische Daten oder als zeitliche Erstreckung angegeben. Sofern in den LCSH die Zeitangabe ein Jahrhundert betrifft, zum Beispiel **20th century**, so wird in FAST dieses Schlagwort auf die zeitliche Erstreckung hin umgewandelt: 1900-1999. In ähnlicher Weise werden geologische Zeitalter als zeitliche Erstreckung zusätzlich zu der Bezeichnung für

das jeweilige Zeitalter angegeben. Für das Zeitalter des Jura erfolgt die Wiedergabe also sowohl als **Jurassic** als auch in der Form **From 140 to 190 million years ago**.

Die einzige Einschränkung bei der Nutzung von FAST-Zeitangaben für zeitliche Erstreckungen besteht darin, dass die zweite Zahlenangabe größer sein muss als die erste. Es besteht somit keine Notwendigkeit, routinemäßig Normdaten für Zeitangaben zu bilden. So bedarf es z.B. keines Normdatensatzes für die dem 20. Jahrhundert entsprechende Zeitangabe 1900-1999. Für Zeitangaben werden Normdatensätze nur dann gebildet, wenn sie für Verweisungsformen benötigt werden.

Validierung der FAST-Schlagwörter

FAST-Schlagwörter gebildet, indem entweder vorhandene LCSH Schlagwörter oder aus den MARC-Datensätzen entnommene Schlagwörter facettiert werden. Die Zerlegung in Facetten beispielsweise des folgenden LCSH-Schlagworts

Architecture, Modern \$y 20th century \$z United States \$v Bibliography

ergibt in FAST die folgenden Schlagwörter

Sache: **Architecture, Modern**
Geographikum: **United States**
Zeit: **1900-1999**
Form: **Bibliography**

Wenn ein Schlagwort wie **Heart \$x Diseases \$x Diet therapy \$v Recipes** multiple Unterschlagwörter innerhalb der gleichen Facette besitzt, so wird die hierarchische Wiedergabe erhalten, woraus sich diese FAST-Schlagwörter ergeben:

Sache: **Heart \$x Diseases \$x Diet therapy**
Form: **Recipes**

Für die Validierung wurde eine Datei mit allen sachlichen Einzelschlagwörtern der LCSH und den geographischen Schlagwörtern aus WorldCat von OCLC gebildet. Diese Datei enthält 6 912 980 einzelne Sachschlagwörter und 1 471 023 geographische Schlagwörter die ihrerseits 50 Millionen Beschlagwortungen entsprechen, die MARC-Datensätzen zugeordnet sind. Diese Schlagwortangaben wurden in Facetten zerlegt, um so die Startversion der sachlichen, geographischen, zeitlichen und formalen Facetten von FAST aufzubauen. Zur automatisierten Validierung der Schlagwörter wurden verschiedene Algorithmen entwickelt. Die Menge der Schlagwörter in der Startversion wurde zahlreichen ausführlichen Validierungen unterzogen, um die Anzahl irrtümlicher Eintragungen zu minimieren. Die nach dieser Validierung verbliebenen Eintragungen wurden dann als FAST-Schlagwörter aufgenommen.

Normdaten

Der letzte Schritt hin zur Entwicklung von FAST war die Bildung eines Normdatensatzes für jedes aufgenommene Schlagwort. Für die Normdaten wurde wegen seiner weiten Verbreitung das Format MARC 21 gewählt. Das Format ist sehr ausführlich und entspricht den meisten, wenn nicht allen Anforderungen von

FAST. Allerdings sehen weder die Normdaten noch die bibliographischen Formate Zeitschlagwörter vor - lediglich als Unterschlagwörter. Hierzu erging ein Vorschlag an MARBI (Machine-Readable Bibliographic Information), dem übergreifenden Komitee der American Library Association zur Pflege der Standards für die maschinenlesbare Wiedergabe von bibliographischen Informationen. Dieser Vorschlag wurde im Juni 2002 mit einigen kleineren Verbesserungen angenommen. Im Ergebnis wurden dem MARC 21 zusätzliche Unterfelder hinzu gefügt, um den besonderen Anforderungen von FAST zu genügen. Beispiele für FAST-Normdaten finden sich in Abb. 1 und 2.

Schlussbetrachtung

Die Werkzeuge zum Zugriff auf bibliographische und sachliche Informationen müssen in einer vernetzten elektronischen Umgebung den folgenden Anforderungen entsprechen:

- Effizienz und Kapazität zur Behandlung großer Mengen an Ressourcen
- Fähigkeit zur Messung und zur Erweiterung
- Interoperabilität

Die Vorzüge von FAST, das auf den in Facetten zerlegten Schlagwörtern der LCSH basiert, lassen sich in folgenden Aussagen zusammenfassen:

- Die computerunterstützte Kontrolle der Normdaten wird erleichtert
- Es ist leichter und wirtschaftlicher zu Pflegen als ein in hohem Maße aneinandergereihtes Vokabular
- Es ist offen für eine computerunterstützte Inhaltserschließung
- Es ermöglicht eine in sich gegliederte Methode um auf unterschiedliche Niveaus die Wiedergabe von Schlagwörtern zu gewährleisten
- Es kann sowohl eine prä- als auch eine postkoordinierte verbale Inhaltserschließung und ihr Wiederauffinden aufnehmen
- Es kann unterschiedlichen Retrieval-Modellen angepasst werden
- Es erleichtert die Aufnahme von Sacherschließungsdaten und eine fächerübergreifende Suche.