



Die Problematik der Verschlagwortung von Mediendaten

1. Allgemeines

Die Datenbank Bildungsmedien ist eine „Metadatenbank“, also eine Datei, die Informationen *über* Medien, in der Regel aber *nicht die Medien selbst* enthält. Ein Hauptproblem solcher Datenbanken ist es, die Informationen so zu strukturieren, dass sie von den Nutzern leicht gefunden werden können. In der Regel bedient man sich dazu eines „Thesaurus“. Das ist eine komplexe, mehrstufige hierarchische Struktur von Ober- und Unterbegriffen, Deskriptoren und Nicht-Deskriptoren, Querverweisen auf andere Begriffe usw., denen jeweils eine unterschiedliche Anzahl von Datensätzen mit weiteren Informationen zugeordnet ist. Es gibt aber auch eindimensionale Schlagwortlisten, die möglichst den Suchwortschatz der Nutzer abdecken und so die eigentlichen Informationen leichter auffindbar machen sollen.

Soll eine Datenbank mit anderen korrespondieren, benötigen alle eine gleiche oder ähnliche Struktur, vor allem aber einen einheitlichen Suchwortschatz. So läge es eigentlich nahe, für die Datenbank Bildungsmedien die „Deutsche Schlagwort Normdatei“ der öffentlichen Bibliotheken zu verwenden. Dies wurde auch in Erwägung gezogen. Die Schlagwort Normdatei erwies sich aber als zu umfangreich und komplex für den begrenzten Bereich der Metadatenerfassung von Bildungsmedien.

2. Sachgebietssystematik und Schlagwortliste: das Konzept der Datenbank Bildungsmedien

Die Datenbank Bildungsmedien hat einen sehr spezifischen Inhalt (Informationen über Medien aller Art für die schulische und außerschulische Bildungsarbeit) und eine ebenso spezifische Zielgruppe (LehrerInnen und andere in der Jugend- und Erwachsenenbildung tätige Personen). Benötigt werden also Suchbegriffe, die sowohl den Inhalt der Medien beschreiben als auch die Bedürfnisse der Zielgruppe abdecken.

Bereits in den siebziger Jahren entwickelte man im Rahmen des Modellversuchs „Aufbau einer Datenbank AV-Medien“ (siehe [Artikel „Rückblick“](#)) eine fünfstufige Sachgebietssystematik und ein eindimensionales Schlagwortverzeichnis. Diese wurden im Laufe der Jahre weiter entwickelt und bilden heute die Grundlage der entsprechenden Felder der EAF-Datenstruktur¹.

Die Gründe für die Einführung von zwei unterschiedlichen Klassifizierungssystemen werden deutlich, wenn man die Vor- und Nachteile von Systematik und Schlagwortliste gegenüberstellt:

¹ EAF = Erweitertes Austauschformat für Mediendaten. Einzelheiten siehe dort.

Die [Sachgebietssystematik](#) der Datenbank Bildungsmedien orientiert sich an den Schulfächern und Schwerpunkten der außerschulischen Bildungsarbeit. Sie ermöglicht gezielte Recherchen innerhalb eines Fachgebiets und die sofortige Erstellung von Listen, Katalogen usw. So ist z.B. eine Abfrage wie „Suche Videokassetten über Säugetiere“ sofort und ohne Kenntnisse einer Datenbank-Abfragesprache durchzuführen.

Sehr häufig haben Nutzer aber wesentlich spezifischere Fragen. Sie wollen z.B. Medien zur Relativitätstheorie. Solche spezifischen Begriffe lassen sich in einer Systematik nicht mehr darstellen. Hier helfen die Volltextrecherche, die über alle Felder läuft, die Text enthalten, oder/und die Schlagwortsuche, die auch dann Treffer erzeugt, wenn das Suchwort nicht in Titel oder Inhaltsangabe vorkommt.

3. Welche Vor- und Nachteile haben Schlagworte in einer Medien-Datenbank?

Die Meinung, eine Verschlagwortung der Medieninformationen sei angesichts der heute vorhandenen schnellen Computer und Volltext-Suchmaschinen überflüssig, ist weit verbreitet und nicht unbegründet. Dennoch spricht nach wie vor Vieles für die Pflege von Schlagworten, wie diese Gegenüberstellung zeigt:

Vorteile	Nachteile
Die Freitext-Suche ist oft zu unspezifisch.	Es ist zusätzlicher Erschließungsaufwand notwendig.
Wichtige Begriffe sind nicht unbedingt in der Inhaltsangabe oder den Titelfeldern vorhanden.	Die Zuordnung der Schlagworte zu den Datensätzen muss stringent sein, d.h. der Nutzer muss sich darauf verlassen können, dass er wirklich alle Medien zur „Relativitätstheorie“ bekommt, wenn er diesen Begriff eingibt.
Die Sachgebietssystematik hat nur einen sehr begrenzten Wortschatz.	Schlagworte sind nur hilfreich, wenn sie zum aktiven Wortschatz des Nutzers gehören.
Die Sachgebietssystematik bringt zu viele Treffer.	
Die Sachgebietssystematik ist nicht intuitiv zu handhaben. Man muss sich erst mit der Struktur vertraut machen.	
Schlagworte ermöglichen Recherchen und das Erzeugen von Listen und Verzeichnissen, die „quer“ zur Systematik liegen. (Z.B. „Klimaveränderung“ berührt die Biologie, die Geographie, die Chemie, die Physik und die Umwelterziehung.)	
Schlagworte ermöglichen bzw. erleichtern komplexe Abfragen.	

Der grau hinterlegte Nachteil ist das Hauptproblem bei eindimensionalen, intuitiven Schlagwortsystemen. Es ist in der Praxis unmöglich, die gesamte Datenbank nach relevanten Medien zu durchsuchen und diese zuzuordnen, wenn man ein neues Schlagwort eingibt. Somit produziert eine eindimensionale Schlagwortliste nahezu ebenso viele Zufallstreffer wie die Volltextrecherche. Ein voll durchstrukturierter Thesaurus hingegen, wie

ihn die öffentlichen Bibliotheken verwenden, erfordert einen erheblichen Zeitaufwand und Fachpersonal für die Medienschließung, was bei den Medienzentren in der Regel nicht vorhanden ist.

4. Der Kompromiss: ein „Kontrolliertes Schlagwortverzeichnis“

Als Kompromiss wurde von den Landesbildstellen Baden und Württemberg (jetzt Landesmedienzentrum Baden-Württemberg) ein zweistufiges „Kontrolliertes Schlagwortverzeichnis“ entwickelt, das darüber hinaus in Sachsen, im FWU und z. T. in anderen Landesmedienzentren verwendet wird. Es soll auch in die Datenbank Bildungsmedien eingebaut werden.

Dieses „Kontrollierte Schlagwortverzeichnis“ wird [in einem eigenen Artikel](#) genauer vorgestellt.