



67th IFLA Council and General Conference

August 16-25, 2001

Code Number: 094-152a-G
Division Number: IV
Professional Group: Cataloguing
Joint Meeting with: -
Meeting Number: 152a
Simultaneous Interpretation: -

Eine virtuelle internationale Normdatei

Barbara B. Tillett

Director, Integrated Library System Program Office
Washington, DC USA

Abstract:

Ein Pool von Normdatensätzen für bibliographische Entitäten (Personen, Körperschaften, Werk/Gestalt, Begriff, Objekt, Veranstaltung und Ort) zur Verwendung im Internet ist nicht nur für Bibliotheken und ihre Nutzer interessant, sondern auch für Verleger, Literaturagenten und Verwertungsgesellschaften, Museen und Archive.

Dieser Vortrag möchte untersuchen, wie dies in der Praxis möglich sein könnte. Die Erstellung von Normdaten ist immer noch der teuerste Teil der Katalogisierung, aber durch gemeinsame Anstrengungen wie NACO, SACO sowie die IFLA-Initiativen, können der Erschließungsaufwand und damit die Kosten für die einzelnen Bibliotheken auf internationaler Ebene reduziert werden.

Es wurde schon häufig festgestellt, dass das World Wide Web in seiner momentanen Ausprägung zur Recherche zu chaotisch ist. Eine Unterstützung ist notwendig und wir können sie anbieten. Die Einführung und Nutzung von Normdaten in der Netzumgebung kann helfen, folgende Ziele zu erreichen:

- die Reduzierung der Katalogisierungskosten durch Verteilung der Arbeitslast, insbesondere da sich unsere Nutzergemeinde ausgedehnt hat und vor allem in Europa heute Bibliotheken zusammen mit Archiven, Museen und Copyright-Agenturen als „Gedächtnis-Institutionen“ angesehen werden. Die kooperative Erstellung und Nutzung von Normdaten wird insgesamt die Kosten reduzieren.
- die Vereinfachung der Erstellung und Pflege von Normdaten auf internationaler Ebene und
- die Zugriffsmöglichkeit für Nutzer auf Informationen in der bevorzugten Sprache, Schrift oder Form, oder die ihre lokale Bibliothek für sie bereithält.

Seit Jahrzehnten werden die Vorteile der Erstellung und Nutzung von Normdaten debattiert und immer wieder herausgestellt. Wenn wir Normdaten in der Netzumgebung einsetzen, wird deutlich werden, wie präzise die Recherche wird, wie die Verknüpfungsstrukturen sich auf die Navigation auswirken und Erklärungen für Abweichungen und Widersprüche bieten, und wie die standardisierten Formen von Namen, Titeln und Schlagwörtern dazu beitragen, Werke anzuzeigen, wie die normierten Ansetzungen von Namen, Titeln und Begriffen mit verschiedenen Hilfsmitteln wie Nachschlagewerke, Lexika, Biographien und Contentdiensten verknüpft werden können. Wir können die Verknüpfungsmöglichkeit dazu nutzen, Bibliothekskataloge in die verschiedenen im Netz verfügbaren Hilfsmittel zu integrieren. Normierte Ansetzungen, die für Zugriff und Anzeige eingesetzt werden, vermitteln dem Nutzer Konsistenz.

Wir alle kennen schlecht ausgestattete OPACs ohne Querverweise oder Verknüpfungen mit Normdateien, die, ehrlich gesagt, ohne diese Fähigkeiten eigentlich gar keine Kataloge sind.

Während der letzten Jahre gab es verschiedene Projekte, die uns dem Ziel der normierten Erstellung und Anwendung von Normdaten auf globaler Ebene näher brachten. In meinem Vortrag auf der LC Bicentennial Conference on Bibliographic Control in the New Millenium im letzten November habe ich das AUTHOR Projekt beschrieben, den Report der IFLA MLAR (Minimal Level Authority Records), das die für normierten Ansetzungsformen wesentlichen Datenelemente (heute nennen wir sie Metadaten) kennzeichnet; IFLA FRANAR (Functional Requirements for Authority Numbers and Records (Sie haben darüber gerade von Francoise Bourdon gehört); Dublin-Core-Sets; Diskussionen über CORC-Normdaten als eine Möglichkeit der Erstellung einer internationalen Normdatei; die Verbesserung der Möglichkeiten durch ein größeres Schriftenangebot anhand Unicode in Windows; und die Erweiterung von NACO und SACO nach den Prinzipien von AACR2 und LCSH (Library of Congress Subject Heading) für die Nutzer weltweit. Ich werde heute nicht in die Einzelheiten gehen, aber in meinem Vortrag habe ich auf die Notwendigkeit der Kompatibilität von Daten hingewiesen und wie wir jetzt verschiedene Datenaustauschformate über Z39.50 protokollieren können. (Das LEAF-Projekt erforscht momentan dieses Modell.)

Außerdem wurden MARC-Konkordanzen einschließlich XML, ONIX und anderen erstellt. Diese Konkordanzen sind hilfreich bei der Recherche von Bibliotheksressourcen mithilfe von Contentangeboten und anderen Quellen im Netz. Die Gesamtheit der technologischen Ressourcen kann nun zusammengefasst werden und bietet die Möglichkeit, eine virtuelle internationale Normdatei in die Realität umzusetzen.

Unsere Einstellung gegenüber der Normdatenkontrolle hat sich im Lauf der Jahre verändert. Allerdings sind die IFLA-UBC-Richtlinien für Normdatenerstellung und -kontrolle identisch mit denen für die Erstellung bibliographischer Einheiten, nämlich

1. jedes Land ist für die autorisierten Ansetzungen von Personennamen und körperschaftlichen Verfassern verantwortlich (Einheitssachtitel, Serien, Schlagwörter wurden nicht erwähnt), und
2. die von nationalbibliographischen Agenturen hergestellten Normdateien stehen allen Nutzern in anderen Staaten, die Normdaten benötigen, zur Verfügung.

In den 60er und 70er Jahren, als sich die Verwendung von Normdaten wirklich durchsetzte, war die Technik noch nicht in der Lage, eine kooperative Arbeitsweise auf internationaler Ebene umzusetzen. Auch fehlende finanzielle Ressourcen für eine internationale Geschäftsstelle verhinderten, dass das visionäre Konzept in die Realität umgesetzt wurde.

In den letzten beiden Jahren hat sich innerhalb einiger IFLA-Arbeitsgruppen eine andere Einstellung gegenüber Universal Bibliographic Control herausgebildet.

Diese neue Sichtweise bestärkt zwar die Wichtigkeit der Verwendung von Normdaten, die weltweit nach denselben Prinzipien erstellt werden, das Augenmerk richtet sich aber primär auf den Nutzer.

Dies ist ein praktischer Ansatz, der berücksichtigt, dass ein Nutzer in China keinen Wert darauf legt, die Ansetzung für Konfuzius in lateinischer Schrift (oder in seltsamen romanisierten Formen) zu finden, sondern in seiner eigenen Schrift.

Um dennoch die Vorteile der kooperativen Standardisierungsarbeit und der Erstellung von bibliographischen Einträgen, die weltweit genutzt werden können, zu erhalten, können wir die normierten Ansetzungen von Namen, Titeln oder sogar Schlagwörtern mit den Normdateien der nationalbibliographischen Zentren und anderen regionalen Agenturen verknüpfen, um eine virtuelle internationale Normdatei zu erstellen.

Es gibt verschiedene Modelle dafür, wie das funktionieren könnte und wir benötigen zur Entwicklung, zum Test und zur Umsetzung von Prototypen weitere Pilotprojekte.

Um den größtmöglichen Nutzen für die Nutzer der jeweiligen Länder zu erzielen, sollten die Beschreibungen in den Schriften sein, die sie lesen können! Welch eine neue Idee! Transliteration kann für einige Nutzer eine Möglichkeit sein, Einträge zu entziffern, aber die Authentizität ist größer beim Benutzen der Originalschrift. Wir sollten schon jetzt wenigstens Verweise auf abweichende Ansetzungsformen in verschiedenen Schriften anbieten. In den Vereinigten Staaten hat MARBI angefangen, diese Möglichkeit zu erforschen und es ist nötig, weitere Arbeit zu leisten. Wir sollten später in der Lage sein, die Schrift und Form einer Ansetzung so anzuzeigen, wie der Nutzer es erwartet und wünscht.

Ich nehme an, dass vielen von uns der Wert parallel existierender Normdatensätze für eine Entität bewusst ist, deren Existenz uns erlaubt, eine katalogabhängige Verknüpfungsstruktur zwischen Querverweisen und Ansetzungsformen herzustellen, um spezifische Benutzerprofile einschließlich abweichender Schriften anhand von Verweisungen bedienen zu können.

Wenn wir uns die Verknüpfungen ansehen, müssen wir erkennen, dass verschiedenen Katalogisierungsregeln unterschiedliche Auffassungen zu Grunde liegen, was als Entität in Frage kommt – die Festlegungen von AACR sind nicht universell, zum Beispiel, im deutschen Regelwerk ist nicht vorgesehen, dass die Logbücher von Schiffen unter den Schiffsnamen anzusetzen sind, deshalb gibt es dort keine Normdatensätze für Schiffsnamen, wohingegen die AACR solche Einträge kennen. Ebenso wird nach den deutschen Regeln keine Ansetzung unter einer veranstaltenden Körperschaft mit Kongressbegriff gemacht, die AACR2 erstellt eine Ansetzung unter dem Körperschaftsnamen mit einem hierarchisch untergeordnetem Kongressbegriff.

Bei nichtindividualisierten Namensansetzungen wird unterschiedlich vorgegangen - die Deutschen haben vor kurzer Zeit ihre Regeln verändert, um mehr Namen zu individualisieren, aber sie benutzen nichtindividualisierte Ansetzungen für Personennamen, deren Vornamen nur aus Initialen bestehen.

Wie auch immer, wir wissen alle, dass sogar mit den gleichen Katalogisierungsregeln, wie zum Beispiel AACR2, ein neuer separater Datensatz für eine Entität angelegt werden kann, wenn wir mehr Information haben, um diese Person zu individualisieren, so dass der Verweis auf den undifferenzierten Namen verschiedener assoziierter Entitäten erhalten bleibt.

Wie können wir eine virtuelle internationale Normdatei erreichen? Es gibt bereits mehrere größere Normdateien, die entsprechend den ihnen zu Grunde liegenden Katalogisierungsregeln und

Regelinterpretationen aufgebaut sind. Es wäre hilfreich, ein einmaliges Projekt durchzuführen, um die bereits bestehenden Datensätze auf die gleiche Entität zu verknüpfen – ein Retro-Abgleichungsprojekt. Ein Vorschlag war, entsprechende „Matching-Algorithmen“ zu benutzen, wie sie von Ed O’Neill und seinen Kollegen bei OCLC entwickelt wurden, die bibliographische Daten mit einem hohen Genauigkeitsgrad maschinell abgleichen.

Wir hätten immer noch manuelle Abgleiche und Nachprüfung ergänzend durchzuführen, würden aber mit einem maschinellen Abgleich zuvor viel erreichen. Des weiteren könnten Textstrings, die aus Verknüpfungen resultieren und Kontrollnummern maschinell hinzugefügt werden, um spätere Verknüpfungen und Pfade zu bevorzugten Anzeigeformen zu erleichtern.

Einige lokale Systeme bieten bereits computer-gestützte Mechanismen für einen Dublettenabgleich für Normdaten. Diese Verfahren könnten der Ausgangspunkt für die Suche in einer virtuellen internationalen Normdatei sein, wenn lokal kein Treffer gefunden wurde.

Wir können uns auch vorstellen, dass es möglich wäre, die gefundenen Treffer in der virtuellen Datei dem Katalogisierer anzuzeigen, damit dieser die Informationen entweder bearbeitet oder in die lokale Normdatei übernimmt, falls gewünscht, direkt in den lokalen Normdatensatz einschließlich der Verknüpfungen.

Einige Systeme stellen zielgruppenspezifische Retrievalmöglichkeiten zur Verfügung, in dem sie Ressourcen für die Online-Suche auswählen, andere Systeme wie „my library“ oder „my opac“ gehen sogar darüber hinaus. Diese können Normdaten-Präferenzen hinsichtlich der vom Nutzer gewünschten Schrift und Anzeige des standardisierten Vokabulars einbauen.

Wir wünschen uns die von der Bibliothek bevorzugte Standardform, die dem Nutzer durch Voreinstellung angeboten wird, aber wir können uns auch vorstellen, eine vom Benutzer ausgewählte Suchmöglichkeit durch Client Software oder Cookies anzubieten, die es dem Nutzer erlaubt, seine gewünschte Sprache, Schrift oder kulturelle Präferenz zu spezifizieren – z.B. bei Rechtschreibeunterschieden durch kulturelle Abweichungen wie beim amerikanischen Englisch oder Rechtschreibungspräferenzen in Großbritannien z. B. labor und labour.

Um dies zu realisieren wurden immer wieder auch Standardnummern vorgeschlagen, z.B. ISAN, INSAN, ISADN usw., ich würde allerdings vorziehen, dass wir die Persistent Record Control Numbers testen und sehen, ob sie funktionieren. Dies würde die Einrichtung einer weiteren internationalen Organisation zur Vergabe und Verwaltung von Nummern unnötig machen. In meinem Vortrag auf der LC Bicentennial Conference habe ich verschiedene Szenarien des Zusammenspiels vorgestellt. Lassen Sie uns einen kurzen Blick auf zwei von ihnen werfen.

Ein Katalogisierer gibt die Information in eine Titelaufnahme ein und das lokale System durchsucht die lokale Normdatei. Das lokale System findet den Eintrag in der lokalen Normdatei und zeigt ihn so an, dass der Katalogisierer bestätigen kann, wenn es sich um die selbe Entität handelt. Ferner hätten wir gern, dass der bibliographische Datensatz anhand der genormten Informationen aus der Normdatei durch das System automatisch aktualisiert wird, nachdem dies vom Katalogisierer bestätigt wurde.

Aber was passiert, wenn in der lokalen Normdatei kein Eintrag gefunden wurde?

Ein Katalogisierer gibt die Information in eine Titelaufnahme ein und das lokale System durchsucht die lokale Normdatei. Es gibt keinen Treffer, das System informiert den Katalogisierer und erlaubt ihm, im internationalen Netz in der virtuellen internationalen Normdatei zu suchen. Wenn das Internet in Zukunft

schneller als jetzt geworden ist, würden wir das Ergebnis sehr schnell finden. Es wird ein Treffer gefunden, beispielsweise ein Normdatensatz, der in der Nationalbibliothek in St. Petersburg erstellt wurde (siehe Graphiken). Unser Katalogisierer sieht sich den Eintrag an und ist vielleicht nicht an allen Informationen, die geboten werden, interessiert, aber er hätte gern eine oder zwei Verweisungen sowie eine Verknüpfung. Das lokale System fragt den Katalogisierer, ob er/sie das System beauftragen möchte, einen Basisnormeintrag aus der gefundenen Datei zu erstellen und ihn zu verknüpfen, was er bejaht.

Unser lokales System erstellt nun einen lokalen Normdatensatz, indem es die Information aus der virtuellen Normdatei zieht, nämlich dem Katalogisat aus St. Petersburg. Der Katalogisierer fügt nun die Kategorie 100 in der entsprechenden Form gemäß den lokalen Katalogisierungsregeln hinzu, in diesem Fall AACR2, und er kann bei Bedarf andere Felder hinzufügen.

Das Lokalsystem fügt das Verknüpfungsfeld 700 hinzu – das MARC-Format mit 7xx Feldern in Normdatensätzen, wo wir die normierte Verknüpfungsform und die Aufnahme-Kontrollnummer und den Quellennachweis für zukünftige Information einsetzen. Dieses Verknüpfen von Normdateien würde primär in den nationalen oder regionalen Normdateien der nationalbibliographischen Zentren vorgenommen werden – in Abhängigkeit von dem jeweils gewählten Modell. Ich werde später darauf zurückkommen.

Wir haben jetzt eine weitere Verknüpfung in der virtuellen internationalen Normdatei zu der autorisierten Form nach AACR2 und dem russischen Datensatz für die gleiche Entität nach den russischen Katalogisierungsregeln in kyrillischer Schrift angelegt. Danach aktualisiert unser lokales System die lokale Titelaufnahme und bringt sie auf den neuesten Stand.

Ein Benutzer kann im Lokalsystem oder über die Cookies im Lokalsystem spezifizieren, ob er die kyrillische Form sehen will – dies können wir ihm anbieten. Es ist auch vorstellbar, dass beliebige Schriften, eine Braille-Tastatur oder auch ein Stimmerkennungssystem angeboten werden und in das spezielle Nutzerprofil oder sein „Cookie“ gespeichert werden können.

Einige der möglichen Modelle für ein virtuelles internationales Normdateiensystem sind: ein verteiltes System, das die unabhängigen Normdateien der nationalen bibliographischen Zentren über eine Z39.50-Schnittstelle durchsucht; ein System mit verknüpften Normdatensätzen, in dem über eine Z39.50-Schnittstelle gesucht werden kann (dies erhöht die Präzision der Recherche); ein zentralisiertes System, wie zum Beispiel das Modell Open Archive Initiative (OAI), welches eine Verbund-Normdatei von Metadaten auf Servern, die mit nationalen Normdateien verknüpft sind, schafft; wiederum ein anderes Modell ist ein zentrales Verknüpfungssystem, in dem eine Normdatei mit allen nationalbibliographischen Zentren verknüpft ist.

Mit den verteilten bzw. verknüpften Z39.50-Modellen kann ein Nutzer ein Suchwort eingeben, das vielleicht nur als Verweis in einer Datei vorhanden ist, aber auch mit einem anderen Datensatz in einer anderen Datei verknüpft ist. Er kann also auch mit einem anderen Sucheinstieg die gleiche Entität recherchieren. Möglicherweise wird sich dieses Modell der Datensatzverwaltung als beste Methode herausstellen.

Das Open-Archives-Initiative-Modell schafft einen Server mit Metadaten aus den nationalen Normdateien, und diese Information wird immer dann aktualisiert, wenn Änderungen in den nationalen Dateien durchgeführt werden. Das bedeutet, dass die tägliche Datenpflege weiterhin so vorgenommen wird, wie sie jetzt von den nationalbibliographischen (oder auch regionalen) Zentren durchgeführt wird. Wenn keine Verknüpfung eingebaut wäre, würden wir möglicherweise an Präzision verlieren, wenn wir in dieser Umgebung recherchieren; aber es gibt Wege, auch die Verknüpfung der Entitäten für dieses Modell einzubauen. Es sind viele Varianten dieses Modells vorstellbar.

Ein anderes Modell wäre, eine zentrale Normdatei zu benutzen und alle anderen mit ihr zu verlinken, so dass nicht jedes nationalbibliographische Zentrum mit allen anderen Teilnehmern diese Arbeit erledigen muss. Ein Katalogisierer hätte mit einer einzigen Recherche in der Zentraldatei weltweit Zugang zu allen anderen Normdaten für diese Entität. Wenn kein Treffer in dieser Datei gefunden würde, könnte über Z39.50 eine Suche in den anderen Dateien durchgeführt werden.

Ich bin sicher, dass Sie sich weitere Varianten dieses Modells vorstellen können.
Wir müssen sie testen, um zu sehen, welche am besten in der heutigen Internetumgebung funktioniert.

Ich habe folgende Empfehlungen in meinem Papier gegeben:

1. Testen Sie das Konzept der verknüpften Normdateien.
2. Erstellen Sie ein Z39.50-Profil für Normdateien (das geschieht bereits als Erweiterung des Bath Profile).
3. Arbeiten Sie mit lokalen Systemanbietern zusammen, um die Fähigkeit lokaler Systeme zur Erstellung von Normdatensätzen sowie die Suche nach Normdatensätzen im Netz zu verbessern, ebenso wie die Möglichkeit für die Nutzer, die Anzeige in der gewünschten Schrift oder Sprache zu realisieren. Und
4. arbeiten Sie mit den Systementwicklern von Suchmaschinen zusammen, um dieses Konzept auf Datenbanken und deren Retrieval zu erweitern, und um etwas Ordnung in das Chaos zu bringen.

Dies gibt Ihnen eine Vorstellung davon wie Katalogisierer im Netz Normdatensätze erstellen können. Wenn sich die Normdatenstruktur weltweit durchgesetzt hat, können auch andere Interessengruppen (Verleger, Verwertungsgesellschaften, Archive, Museen und andere Bibliotheken) diese Information nutzen und ihre Kosten reduzieren. Kontrollierte Erstellung und Anwendung von Normdaten wird Nutzern helfen, von Normierung und präzisiertem Retrieval zu profitieren, und das in der von den Nutzern gewünschten Sprache und Schrift.

Übersetzt von Renate Gömpel in Zusammenarbeit mit Margret Dick, Barbara Pfeifer, Renate Weber (Die Deutsche Bibliothek)