



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

1-9 August 2003, Berlin

Code Number: 054-G
Meeting: 157. Bibliographic Control - **Workshop**
Simultaneous Interpretation: -

Die Beziehung zwischen Subject Gateways und Nationalbibliografien im internationalen Kontext

Bohdana Stoklasová

Deputy for Library Management
National Library of the Czech Republic
Prague, Czech Republic

Marie Balíková

Subject Cataloguing Department Head
National Library of the Czech Republic
Prague, Czech Republic

Ludmila Celbová

Remote Electronic Resources Head
National Library of the Czech Republic
Prague, Czech Republic

Zusammenfassung

In diesem Papier wird die Beziehung zwischen Subject Gateways und Nationalbibliografien zusammen mit allgemeinen Grundsätzen der universalen bibliografischen Erschließung und im weiteren Zusammenhang mit der Integration von heterogenen Informationsquellen untersucht. Die Beispiele mit den Erfahrungen der Tschechischen Republik werden zur Darstellung allgemeiner Probleme genutzt, die bei der Integration heterogener Informationsquellen aus verschiedenen Ländern und zu verschiedenen Fachgebieten entstehen können. Die wichtigsten Hindernisse bei der Integration von Informationsquellen werden beschrieben und zum Abschluss Empfehlungen für die Verbesserung der bibliografischen Erschließung gegeben.

Einführung

Der aktuelle IFLA Bibliography Section Strategic Plan <http://www.ifla.org/VII/s12/annual/sp12.htm> legt großen Wert auf die Entwicklung von Richtlinien, die mit Beispielen und Empfehlungen Nationalbibliografischen Agenturen helfen sollen, bibliografische Dienste zu beginnen oder zu verbessern. Dieses Papier soll Diskussionen auslösen, die zu Richtlinien sowohl für Subject Gateways als auch für Nationalbibliografien in einem internationalen Kontext führen können.

Um allgemeine Fragen und Probleme darzustellen werden wir Beispiele aus unserem eigenen Land verwenden. Die Tschechische Republik ist ein Land, das seine Nationalbibliografie im letzten Jahrzehnt grundlegend ändern und verbessern musste, nachdem die Revolution von 1989 Freiheit und die Möglichkeit gemeinsamer Datennutzung und des Datenaustauschs gebracht hat. Dort, wo es keine oder nur unklare Richtlinien gab, konnten wir Sackgassen nicht vermeiden und verloren viel Zeit und Ressourcen, sowohl personell wie finanziell. Heute ist unsere Nationalbibliografie vergleichbar mit denjenigen anderer Länder, die uns als Beispiele und Empfehlungen gedient haben. Jetzt, an der Startlinie zur Bildung von Subject Gateways, wollen wir Sackgassen vermeiden, Antworten auf unsere Fragen erhalten und diese Antworten mit anderen als Richtlinien oder mutmaßliche Trends nutzen.

Die Czech National Bibliography: eine kurze Beschreibung und Geschichte

Die Czech National Bibliography (CNB) wird seit 1922 regelmäßig veröffentlicht. Das ermöglichte eine relativ einfache retrospektive Konversion aller Daten für die im 20. Jahrhundert im Land erschienenen Publikationen. Die aus der Konversion entstandenen MARC-Datensätze sind zugänglich in der CNB http://aip.nkp.cz/tornado/webtor.cgi?MainPage=%2E%2E/cnb_en/cnben_main%2Ehtm, der Datenbank der Nationalbibliothek http://sigma.nkp.cz:4505/F/4GVLVLI3BEUJ8YSF5TH12PU2R17SUMGBFBA6ADJFLYSMRRKF36-04203?func=file&file_name=find-b&CON_LNG=ENG und auch in WorldCat.

Für die Nationalbibliothek der Tschechischen Republik war es relativ leicht, den Berichtsumfang für die Nationalbibliografie zu bestimmen: Wir sprechen eine Sprache – tschechisch, in der die meisten Veröffentlichungen erscheinen und verwenden eine Schrift, die lateinische Schrift. Das Gebiet der Nation hat sich im Zeitablauf geändert, aber die Bestimmung des Berichtsumfangs für bibliografische Zwecke ist relativ gut möglich. In Bezug auf Standards haben wir jedoch eine vollständig andere und weit schwierigere Situation. Während der fünfzig Jahre dauernden totalitären Herrschaft war kein freier Datenaustausch in unserem Land möglich und bis zur Revolution im Jahr 1989 hatten wir so gut wie keine praktische Erfahrung mit internationalen Standards. Um kooperieren zu können, mussten wir die Standards übersetzen und nicht nur deren Sprache verstehen, sondern auch deren für uns häufig neue Konzepte.

Tschechische Katalogisierer hatten daher seit 1989 eine anstrengende Zeit. Sie mussten drei Formate lernen, 1989 das tschechische Austauschformat, 1994 UNIMARC und jetzt MARC21 für die 2004 geplante Migration. Weiterhin wurden 1994 neue Katalogisierungsregeln, AACR, übersetzt und eingeführt und neue Schlagwortansetzungen auf der Basis der Library of Congress Subject Headings (die CNB-Datensätze haben auch ihre englischen Entsprechungen, hoch geschätzt von ausländischen Nutzern und von Katalogisierern, die unsere Daten aus dem WorldCat übernehmen) vergeben. Außerdem wurde die neue Version der UDC (der neue Master Reference File – MRF) und die obere Ebene der DDC mit einer Konkordanz DDC – UDC für Subject Gateways und die WLN Conspectus Anwendung einbezogen. All diese neuen Standards mussten im Vergleich zu anderen Ländern in sehr kurzer Zeit eingeführt werden, um sowohl bibliografische Daten als auch Normdaten weltweit nutzen zu können. Das neue Konzept für nationale Normdateien für Personennamen und Schlagwörter erfordert ebenfalls erheblichen zeitlichen und personellen Aufwand. Inzwischen hat sich die Zahl der gedruckten Veröffentlichungen in der Tschechischen Republik verdoppelt (in den letzten fünf Jahren sind jährlich ca. 15.000 Titel erschienen) und es erscheinen Online-Publikationen, die nicht nur bibliografisch erschlossen sondern auch für unsere Bibliotheksdienste umgesetzt werden müssen. Wie andere Bibliotheken auch müssen wir eine Antwort auf die schwierige Frage finden, wie alle diese neuen Aufgaben und die wachsende Zahl von Veröffentlichungen einschließlich Online-Publikationen ohne zusätzliche Haushaltsmittel und mit ständig weniger werdendem Personal bewältigt werden können.

Die Tschechische Nationalbibliografie und Online-Publikationen

Mit dem Blick auf diese Probleme haben wir 1998 an der International Conference on National Bibliographic Services (ICNBS) in Copenhagen teilgenommen und danach alle dortigen Empfehlungen analysiert und dabei bewertet, wie weit wir diesen Empfehlungen bereits folgen. Aufgrund der Empfehlung Nr. 4:

National bibliographies should include the current national output, and where practicable, they should also provide retrospective coverage. When necessary, selection criteria should be defined and published by the national bibliographic agency, 1)

überlegten wir, ob wir auch Online-Publikationen in die nationalbibliografische Sammlung einbeziehen sollten. 1998 gab es in unserem Land nicht sehr viele Online-Publikationen und deshalb entschieden wir, sie vorerst nicht zu sammeln, sondern die Situation in unserem Land und auch die Entwicklung von Standards und von Kooperationen in anderen Ländern zu beobachten, um in ein oder zwei Jahren wieder darauf zurückzukommen. Um offen zu sein, es war uns bewusst, dass die bereits vorhandenen, aber nicht gesammelten Dokumente wahrscheinlich verloren gehen würden. Wir hatten jedoch nicht genug Erfahrung und Mittel, weder für die Entwicklung von Selektionskriterien noch für die Analyse verschiedener Katalogisierungsstandards oder die Katalogisierung selbst. Mit unseren begrenzten finanziellen und personellen Mitteln scheuten wir uns, ein Projekt zu beginnen, für das niemand zukünftige Anforderungen und Ergebnisse vorhersagen konnte. Unsere Situation kann perfekt mit der ersten Aussage von Michael Gormans Vortrag bei der IFLA-Konferenz 2001 beschrieben werden:

The great irony of our present situation is that we have reached near-perfection in bibliographic control of "traditional" library materials at the same time as the advent of electronic resources is seen by some as threatening the very existence of library services – including bibliographic control. 2)

Neben dieser nicht sehr optimistischen Aussage präsentierte Gorman sein Modell einer Pyramide, mit der eine Klassifizierung des Chaos versucht werden sollte, um die bibliografische Verzeichnung von ausgewählten Online-Publikationen möglich zu machen. Das gleiche Modell wurde 2002 von John Byrum verwendet 3).

Gorman's Practical Model for Bibliographic Control of Remote Resource



Das Konzept dieser Pyramide war außerordentlich hilfreich bei all unseren Diskussionen über die verschiedenen Möglichkeiten, tschechische Online-Publikationen bibliografisch zu erfassen.

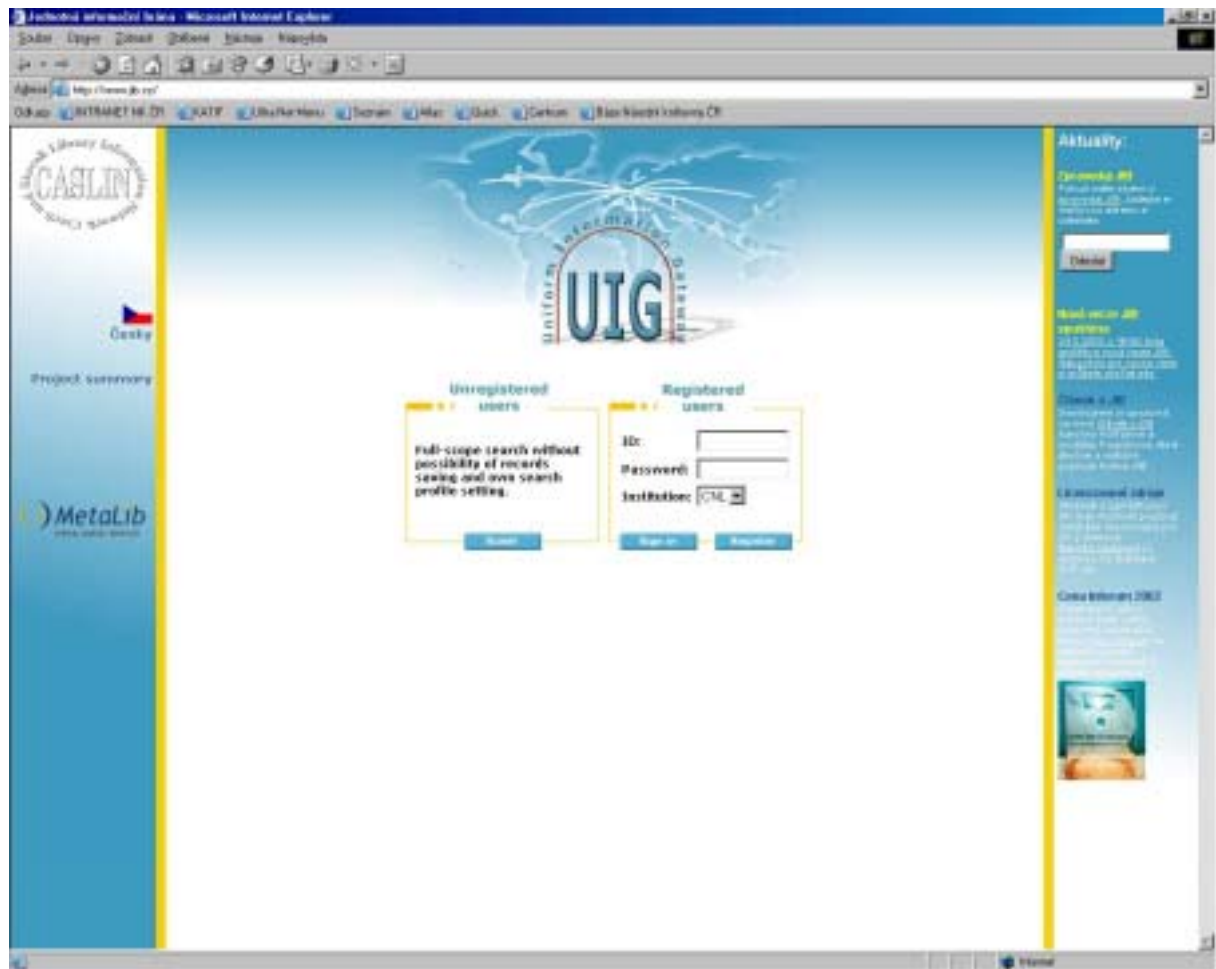
Registrierung, Archivierung und ständige Verfügbarkeit tschechischer Online-Publikationen waren ein wichtiger Teil des Forschungsplans der Nationalbibliothek seit 2000. Das Ziel unseres ersten R&D-Projektes, WeBArchiv, war es, die Leistungen der auf diesem Gebiet fortschrittlichsten Länder zu analysieren und in unserem Land mit solchen Aktivitäten zu beginnen, die auf diesen Leistungen und natürlich unter Beachtung der dort entwickelten Standards aufbauen. Die Nationalbibliothek beteiligt sich aktiv an der Lösung rechtlicher Probleme unter dem Dach internationaler Organisationen wie CDNL und hat eine sehr fruchtbare Kooperation mit verschiedenen ausländischen Bibliotheken hauptsächlich aus den skandinavischen Ländern aufgebaut. Diese Kooperation konzentriert sich vor allem auf die technische Entwicklung, deren Ergebnis Programme sind, die unter dem Dach der Europäischen Union mit unserer Mitwirkung entstanden sind. Die Ergebnisse der verschiedenen ausländischen Projekte waren für uns eine große Hilfe. Wir konnten sehen, auf welchen Wegen versucht wurde, die Pyramide zu „erklimmen“ und welche Anstrengungen für die Koordination aufgewendet wurden. All diese Beispiele dienten uns als Richtlinien und führten dazu, dass wir relativ schnell die Probleme des Harvesting, der Auswahlkriterien und der Metadatenauswahl lösen konnten. Sobald die Richtlinien und Methoden fertig gestellt waren, ergab sich eine andere schwierige Frage: Wer wird die beiden oberen Ebenen der Pyramide erstellen? Die Nationalbibliothek hat eine kleine Abteilung für Online-Publikationen eingerichtet, die für alle Aktivitäten mit diesen Publikationen zuständig ist, einschließlich der Verhandlungen mit den Verlagen. Daher bleibt nicht viel Zeit für das Erstellen von Enriched Dublin Core- und Full Standard-Datensätzen und so ist die Spitze der Pyramide sehr schmal. Wir versuchen Mittel und Kooperationswege zu finden, um diese Situation zu verbessern.

Das Czech National Subject Gateway Project

Als wir mit dem Projekt für ein tschechisches nationales Subject Gateway begannen, war ein folgerichtiger Gedanke, einige Zuständigkeiten für die Bildung der oberen Ebenen der Pyramide entsprechend der Fachgebiete aufzuteilen. Das Hochwasser in Prag im Jahr 2002 führte zu neuen Prioritäten bei der Vergabe der Haushaltsmittel und daher standen für das Projekt eines tschechischen nationalen Gateways nur sehr begrenzte Mittel für die Jahre 2002 und 2003 zur Verfügung. Die Richtlinien zur Gestaltung dieses Subject Gateway, basierend auf internationalen Standards und ausländischer Erfahrung, wurden jedoch fertig gestellt und veröffentlicht. Zahlreiche europäische und australische Subject Gateways und auch solche aus anderen Ländern wurden untersucht, wobei besondere Aufmerksamkeit auf den Inhalt, die Auswahlkriterien und das Management gerichtet wurde. In komplizierten Tabellen wurden Metadaten-, Beschreibungs-, Fachgebiets- und Verwaltungssets verglichen und die am meisten verwendeten in die Empfehlungen übernommen. Wir waren auch sehr daran interessiert, welche Klassifikationen und Schlagwörter in den verschiedenen Projekten genutzt werden und deshalb waren das Renardus-Projekt und seine Ergebnisse äußerst nützlich für uns. Wir haben diese Ergebnisse aber auch mit Lösungen in anderen Ländern und Projekten außerhalb von Renardus verglichen. Das Czech National Subject Gateway Project ist verbunden mit dem Uniform Information Gateway Project. Dies ist ein sehr komplexes Projekt, in dem heterogene Datenquellen (einschließlich Volltexten und verschiedenen digitalen Objekten) aus verschiedenen Ländern eingebunden werden sollen. Es zeigt sowohl Nationalbibliografien wie Subject Gateways in einem breiteren Kontext und daher ist es wichtig, es hier zu erwähnen.

Das Uniform Information Gateway und die Integration heterogener Datenquellen

Das Uniform Information Gateway (UIG) ist ein aktives nationales Projekt und kann unter der folgenden Adresse begutachtet und getestet werden:
<http://www.jib.cz>.



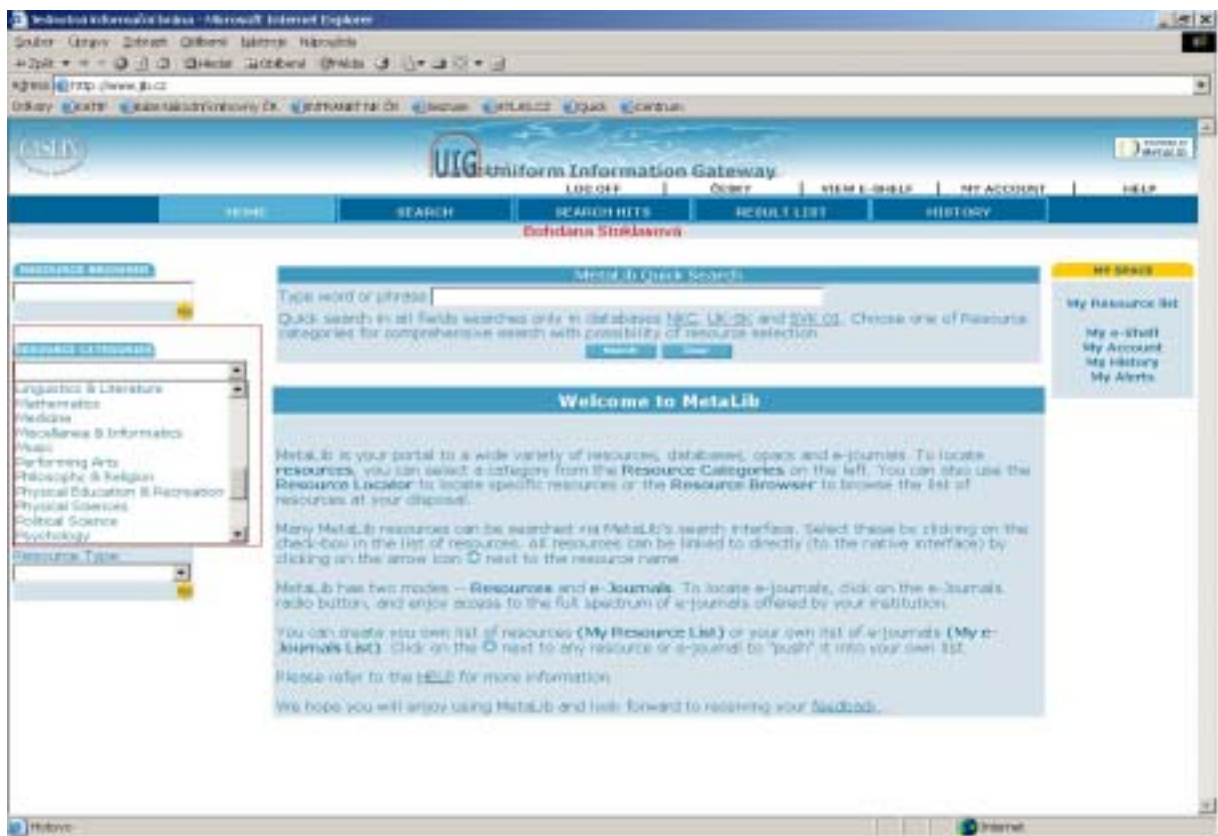
Das UIG-Projekt hat vier Hauptziele:

1. **Integration und Bereitstellung eines qualitativ neuen Zugangs zu heterogenen Datenquellen:** Integrierter Zugang zu Informationen über Dokumente und zu den Dokumenten selbst; einheitliche benutzerfreundliche Oberfläche; einheitliche Suchoberfläche und einheitliche Anzeige der Ergebnisse; Vermischung und Bereinigung von Ergebnissen; Authentifizierung und autorisierungspflichtige Dienstleistungen; Konfiguration persönlicher Oberflächen einschließlich einer Liste von häufig genutzten Datenquellen und eines persönlichen Katalogisierungsbereichs mit der Möglichkeit des Formatwechsels; Speichern von Anfragen und Ergebnissen; wiederholtes Ein- und Ausloggen nicht erforderlich (einmaliges Anmelden im Bereich der Datenquellen); erweiterter Service für Bestandsangabe, Volltext, Abstracts, elektronische Dokumentlieferung und ILL; Informationen zu Verfassern; zugehörige Dokumente entsprechend Schlagwort und Titel.
2. **Bestandsbeschreibung und Bestandsentwicklung:** einheitliche, eindeutige, standardisierte und international anerkannte Beschreibung von Datenquellen aus der Tschechischen Republik (z.B. Thematische Karte); Kooperation bei der Bestandsentwicklung und gemeinsame Nutzung
3. **Erschließungstransformation:** Nationale Normdatei für Schlagwörter; englische Entsprechungen, DDC-UDC-Konkordanzen.
4. **Czech National Subject Gateway:** Erstellung von einheitlichen Richtlinien nach internationalen Standards für den Aufbau von Tschechischen Subject Gateways; Aufteilung der Zuständigkeit für bestimmte Fachgebiete; nationale Koordination.

Als wir mit dem UIG-Projekt begannen, rechneten wir mit gravierenden Problemen in den Bereichen, die unter Ziel 1 aufgeführt sind. Bei den Zielen 2, 3 und 4 vermuteten wir, dass sie leichter zu erreichen sind. Die Realität war ganz anders als erwartet. Die „technische“ Integration mit all den genannten Aspekten war überraschenderweise leicht. Das UIG-Projekt wird unterstützt von MetaLib und SFX. Diese Tools decken alle unsere Anforderungen ab und bieten manchmal sogar mehr als erwartet. Es gibt auch andere vergleichbare Tools auf dem Markt, die ähnliche Funktionalitäten haben. Das bedeutet, dass „technische“ Aspekte der Integration handhabbar sind. Voraussetzung dafür ist, dass die Informationen über die zu integrierenden Datenquellen in Anlehnung an internationale

bibliothekarische und technische Standards aufbereitet sind und dass man weiß, welche Datenquellen zur Verfügung stehen und für ein Land, Fachgebiet etc. integriert werden können. Wie bereits angeführt, integriert das UIG nicht nur Informationen über Dokumente sondern auch die Dokumente selbst, vornehmlich als Volltexte. Wir gingen davon aus, dass die Integration von Informationen über Dokumente, die seit vielen Jahren von Bibliotheken aufbereitet, erstellt und verbreitet werden, leicht sein würde, während die Integration von Volltexten, die in den vergangenen Jahren nur von verschiedenen Informationsanbietern und Aggregator-Datenbanken angeboten wurden, wesentlich schwieriger sei. Überraschenderweise war die Integration von Diensten der Informationsanbieter häufig wesentlich einfacher als diejenigen der Bibliotheken. Im Vergleich zu Bibliotheken scheinen diese Organisationen wesentlich flexibler und sich der Tatsache bewusst zu sein, dass für mehr und mehr Nutzer Information, die nicht einfach zu integrieren ist, nicht existiert.

Lassen Sie uns auf die Ziele 2, 3 und 4 schauen, zwischen denen eine starke Verbindung besteht. Von Anfang an können die Nutzer von UIG einige der weiten Fachgebietenkategorien wählen. Wir haben uns für die 24 WLN-Conspectus-Abteilungen als "top subject crossroads" entschieden und gliedern unterschiedliche Datenquellen unter dem Dach dieser weiten Kategorien.



In der Theorie scheint das eine leichte Aufgabe zu sein – einfach Bibliotheken zu ermitteln und zu kontaktieren, die gute Sammlungen in bestimmten Fachgebieten mit Subject Gateways zu diesen Fachgebieten und anderen Quellen einschließlich elektronischer Zeitschriften haben. Leider ist die Wirklichkeit vor allem bei Bibliotheken davon weit entfernt. Einige Bibliotheken haben sehr gute Bestandsbeschreibungen und Bestandsentwicklungsstrategien, die online zugänglich und international verständlich sind. Diese Bibliotheken haben teilweise Subject Gateways für ihre wichtigsten Fachgebiete aufgebaut. In unserem Land war das nicht der Fall, daher konnten wir keine geeigneten Verbindungen zu tschechischen Bibliotheken herstellen. Der Aufbau einer einheitlichen, eindeutigen, standardisierten und international anerkannten Erschließung von Datenquellen aus der Tschechischen Republik hat gerade erst auf der Basis der WLN-Conspectus-Methode begonnen. Bibliotheken, die bedeutende Sammlungen in einzelnen Fachgebieten haben, werden für den Aufbau von Subject Gateways für diese Fachgebiete verantwortlich sein. Wir legen Wert auf eine hochwertige

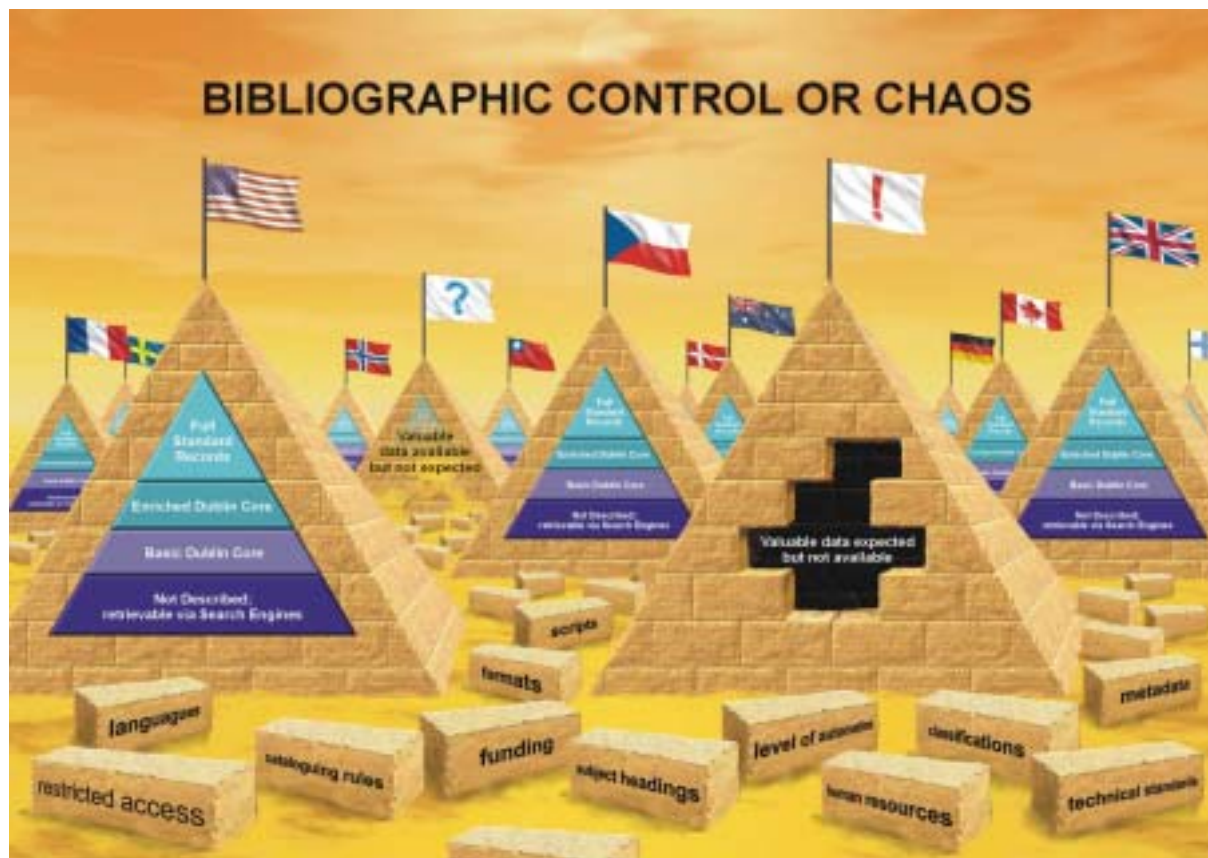
Erschließung der tschechischen Datenquellen, damit andere Länder sie nutzen können, und auf die Verbindung zu ausländischen Katalogen, um nicht nochmals katalogisieren zu müssen. Leider sind wegen der verschiedenen Standards viele Dokumente, gedruckte wie elektronische, mehrfach in verschiedenen Ländern katalogisiert, während andere überhaupt nicht erschlossen sind.

Wir wollten Online-Publikationen genauso behandeln wie die anderen Dokumente in unserer Nationalbibliografie und änderten deshalb die Regeln sowohl für die formale wie die sachliche Erschließung. So konnten wir unsere Datensätze in den WorldCat einbringen und Datensätze für unsere ausländischen Erwerbungen aus dem WorldCat übernehmen. Um die Nutzung unserer Daten zu vereinfachen, entschieden wir uns außerdem, unseren Datensätzen die LCSH in englischer Sprache hinzuzufügen. Es gab viele Diskussionen, ob wir diesen Mehrwert für ausländische Katalogisierer erbringen können. Heute sehen wir, dass es eine gute Entscheidung war, nicht nur für ausländische Katalogisierer, sondern vor allem auch für unsere Nutzer. Wenn sie im UIG der Tschechischen Nationalbibliografie oder im Katalog der Nationalbibliothek auf Tschechisch recherchieren, sehen sie auch die englischen Ansetzungen. Das ist äußerst hilfreich bei der parallelen Suche in ausländischen Datenbanken oder anderen Diensten, die hauptsächlich auf Englisch angeboten werden.

Der Wechsel von unseren bisherigen Regeln zu den neuen Standards war nicht leicht, aber die Integration in das UIG wäre ohne diesen Wechsel nicht möglich gewesen. Eine nahtlose Verbindung, die auf den gleichen Standards beruht, ist wesentlich besser als die ständige Überbrückung unterschiedlicher Standards. Eine Überbrückung konnten wir nicht vermeiden. Wir haben eine lange UDC-Tradition und Millionen von vorhandenen Datensätzen mit UDC. Daher mussten wir das Problem lösen, die große Zahl der vorhandenen Datensätze mit der ständig wachsenden Zahl von auf DDC basierenden Subject Gateways zu verbinden. Wir entschieden uns, dazu die UDC-DDC-Konkordanzen zu verwenden, die auf den ungefähr 500 WLN-Conspectus-Kategorien basieren. WLN hat einige grundlegende Vorbereitungen übernommen, die von der Nationalbibliothek vervollständigt wurden. Die Konkordanz ist bei der Nationalbibliothek erhältlich und wir testen die Umsetzung gerade.

Fazit

Die Integration heterogener Datenquellen ist kein Luxus, sondern sie sollte ein Standardangebot moderner Bibliotheken sein, wenn sie auf dem mittlerweile weltweiten Informationsmarkt Erfolg haben möchten. Die Integration heterogener Datenquellen ist sehr stark verbunden mit Subject Gateways und der bibliografischen Erschließung im Allgemeinen. Es ist leicht, alles zu integrieren, was geregelt ist, es ist aber nicht möglich, das Chaos zu integrieren. Als wir mit der Integration begannen, waren wir oft in einer ähnlichen Situation wie die Eroberer der ägyptischen Pyramiden. Einige Hindernisse sind die gleichen wie vor Tausenden von Jahren: einerseits irreführende Wege zu Stellen, wo man etwas Wertvolles erwartet, aber nur eine leere Pyramide findet, andererseits sehr verborgene Pyramiden, die Schätze enthalten. Es gibt auch sehr verschiedene Sprachen und Schriften, die Notwendigkeit eines „Passwortes“ und vieles andere. Einige Hindernisse sind typisch für das 21. Jahrhundert: Unterschiedliche Formate, unterschiedliche Katalogisierungsregeln, unterschiedliche Metadaten, unterschiedliche Klassifikationen und Schlagwörter, unterschiedliche technische Standards, Mangel an finanziellen und personellen Ressourcen, die verschwendet werden, um all diese Differenzen zu überwinden.



Kann dieses Bild geändert werden? Wir sind überzeugt davon. Verschiedene Sprachen und Schriften müssen bewahrt und erschlossen werden, für einige Dokumente wird es immer einen eingeschränkten und begrenzten Zugang geben und der Grad der Automatisierung wird für lange Zeit von Land zu Land unterschiedlich bleiben. Dennoch kann mehr Kooperation und Koordination im Bereich sowohl der bibliothekarischen wie der technischen Standardisierung zu einer verstärkten praktischen Anwendung der UBC-Richtlinien führen und personelle und finanzielle Ressourcen sichern, die so dringend für die Reduzierung des Chaos gebraucht werden. Vorausgesetzt, wir sind bereit, uns auf gemeinsame Standards zu verständigen und die verschiedenen Länder verzeichnen tatsächlich ihre nationale Produktion in strikter Übereinstimmung mit diesen Standards, die es anderen ermöglicht, die Ergebnisse zu nutzen, könnte das Bild bald ganz anders aussehen und das Wort Chaos aus dem Titel verschwinden!

- 1) The final recommendation of the International Conference on National Bibliographic Services : Copenhagen, 25-27 November 1998. Available at <<http://www.ifla.org/IV/3/icnbs/fin.htm>> Recommendation no. 4.
- 2) Gorman, Michael, "Bibliographic control or chaos: an agenda for National Bibliographic Services in the 21st century". IFLA Journal 27 5/6: pp. 312.
- 3) Byrum, John. D., "Challenges of providing bibliographic access to remote electronic resources in national bibliographies : problems and solutions – an overview [Paper presented at the 68th IFLA Council and General Conference, Glasgow, 2002]

Deutsche Übersetzung: Claudia Werner, Die Deutsche Bibliothek