

Multilingual Subject Retrieval

Bibliotheca Alexandrina Subject Authority File and Linked Subject Data

European Conference on Data Analysis
July 10, 2013

Magda El-Sherbini
Head, Cataloging Department
The Ohio State University Library

Background

- When did it begin?
- Why is BA's effort unique?
- Tri-lingual

Outline

- Workflow
- Linking process
- Similarity and difference between the BA and the LC/OCLC authority files
- Advantages and disadvantages of creating local AF
- Next steps and conclusion

I. Workflows: Cataloging Process For Latin Scripts

- Search OCLC for copy
 - No verification of access point
- Create original records only in their OPAC
 - Search OCLC Authority file and download
 - Create access points based on AACR2
- Daily report and post cataloging authority processing

Daily List of Access Points



VTLS UTILITIES

[Home Page](#) | [Send Feedback](#) | [Administration](#)

Cataloging

Cataloger Productivity

Authority Messages

Date From :

To :

Select Cataloger Name:

All

Select Message Category:

- Possible duplicate author heading
- Possible duplicate title heading
- Possible duplicate user heading
- No subject permanent authority for heading
- No author permanent authority for heading
- No title permanent authority for heading
- Authority heading updated
- New heading created[Subject 6xx]**
- New heading created[Author 1xx or 7xx]
- New heading created[Title 2xx or 4xx]

Select All>>

UnSelect All>>

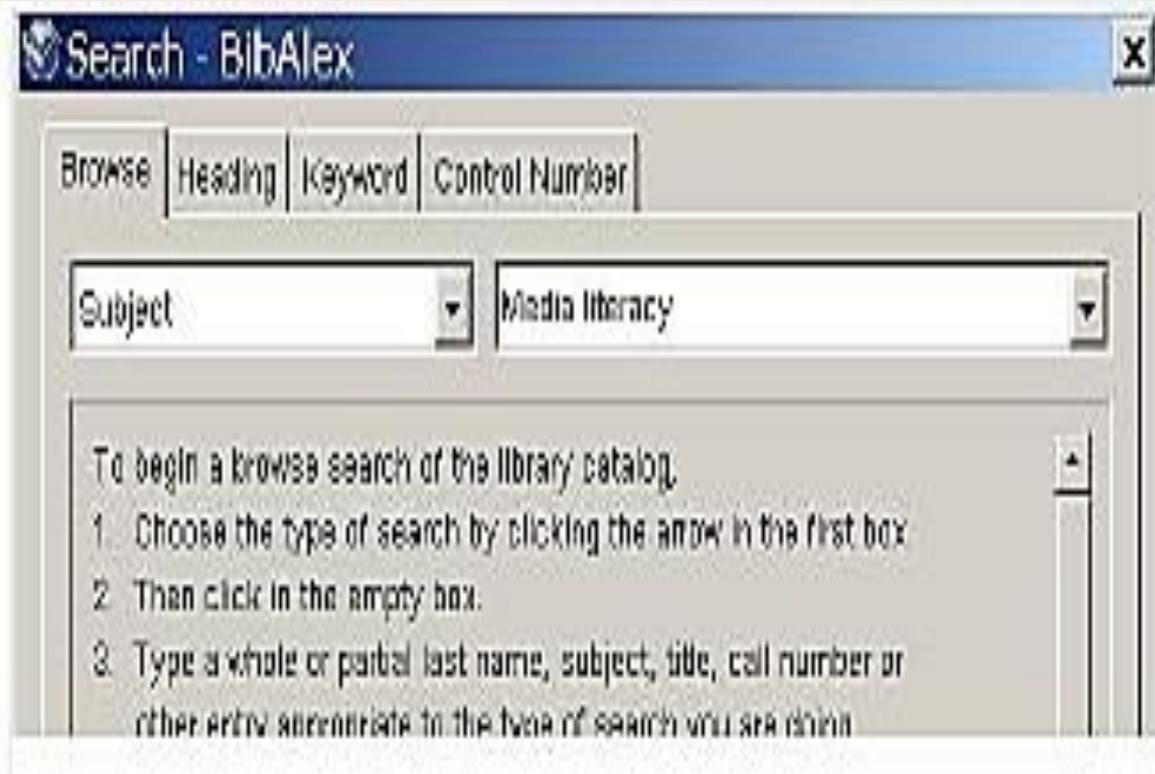
ID	User Name	Time Stamp	Diagnostic Text
392353	nesrine.abdelmeguid	1/13/2009 10:40:13 AM	651 \$a Croatia \$x Antiquities.
392353	nesrine.abdelmeguid	1/13/2009 10:40:13 AM	650 \$a Excavations (Archaeology) \$z Croatia.
392382	nesrine.abdelmeguid	1/13/2009 11:17:58 AM	650 \$a Mass media \$x Social aspects.
392364	may.abdelshaffy	1/13/2009 11:22:05 AM	650 \$a Bubbles \$v Juvenile literature.
392365	nesrine.abdelmeguid	1/13/2009 11:28:45 AM	650 \$a Mass media \$x Study and teaching.
392365	nesrine.abdelmeguid	1/13/2009 11:28:45 AM	650 \$a Media literacy.
392375	may.hamdy	1/13/2009 11:47:41 AM	650 \$a الطب.
392382	may.hamdy	1/13/2009 12:13:25 PM	650 \$a الطب.

More Workflow

- Search for “media literacy” in the local VIRTUA database to check for duplication
- Search the LC/OCLC Authority File and download the authority record
- If the authority record is not found, the cataloger creates a new record locally.
- For either downloaded or newly created authority records, search the BA database to determine if there are items under term in another language that is equivalent

Let's look at the steps for creating
the term
“Media literacy”

Search the BA

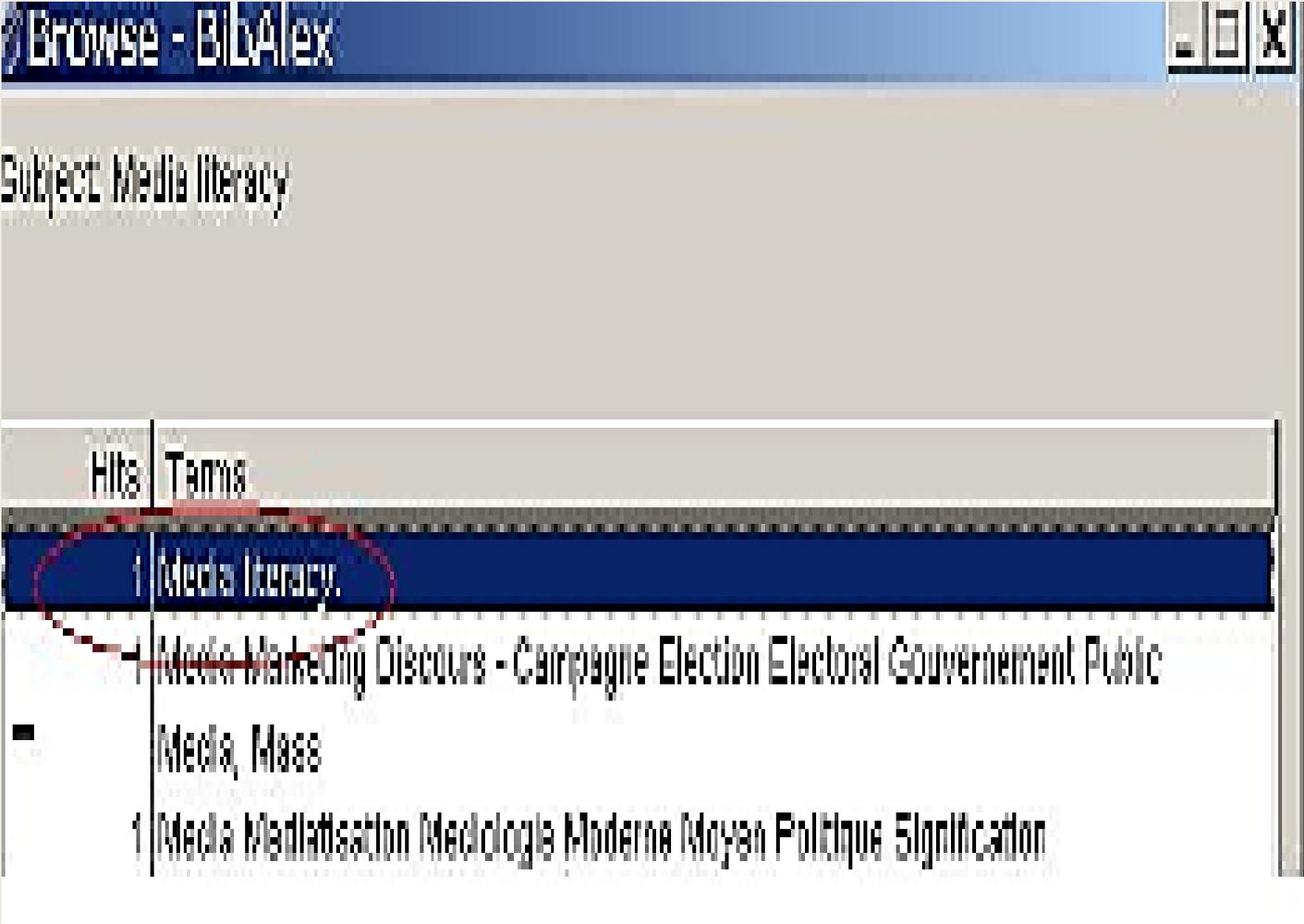


The screenshot shows a web browser window titled "Search - BibAlex". The interface has a navigation bar with four tabs: "Browse", "Heading", "Keyword", and "Control Number". The "Browse" tab is selected. Below the tabs, there are two dropdown menus. The first dropdown is labeled "Subject" and has a downward arrow. The second dropdown contains the text "Media literacy" and also has a downward arrow. Below these dropdowns is a text area containing instructions for a browse search:

To begin a browse search of the library catalog.

1. Choose the type of search by clicking the arrow in the first box
2. Then click in the empty box.
3. Type a whole or partial last name, subject, title, call number or other entry appropriate to the type of search you are doing.

The term is found



Subject: Media literacy

Hits	Terms
1	Media literacy.
1	Media Marketing Discours - Campagne Election Electoral Gouvernement Public
1	Media, Mass
1	Media Mediatization Medecologie Moderne Moyen Politique Signification

Search OCLC

OCLC Connexion Home |

Cataloging **Authorities** Express General oclc

Search Browse Create Show

Search Options * Authority File Create Options Show Options

Browsing the Authority file Root LCSH in

View Expanded 9 View List

Root Results

R1. Media, Condensed	
R2. Media consultants	
R3. Media coverage of . . .	
R4. Media, Culture (Biology)	
R5. Media, Embedded	
R6. Media (Ill.)	
R7. Media, Inhomogeneous	
R8. Media librarians	1 record
R9. Media literacy	1 record
R10. Media Luna (Granma, Cuba)	1 record
R11. Media, Mass	1 record

Browse

Search History

Authorities

Browse for in

Select root (sca)

Expanded term

Export the record to their local system

OCLC Connexion

→ Cataloging → **Authorities** → Express → General OCLC

Search Browse Create Show

Search Options * Authority File Create Options Show Options

<u>Ser num</u>	<u>Name</u>	<u>Subdiv tp</u>
010	sh 96010160	
040	MnMHCL \$b eng \$c DLC \$d DLC	
150	Media literacy	
450	Mass media literacy	
550	Information literacy \$w g	
670	95042682: DeGaetano, G. Screen smarts : a family guide to media literacy, 1996.	
670	Hennepin \$b (Media literacy)	
670	74077021: Houk, A.S. Media literacy ... 1974.	
670	89130471: Limburg, V.E. Mass media literacy, c1988.	
670	94022004: Silverblatt, A. Media literacy : keys to interpreting media messages, 199:	
670	96035023: Silverblatt, A. Dict. of media literacy, 1997.	
680	\$i Here are entered works on a person's knowledge and ability to use, interpret and	

Edit the record

MARC Editor - BibAlex			
005			000809 j anannbabn a ana c
010			\a sh 96010160
040			\a MnMHCL \b eng \c DLC \d DLC \d EG-AIBA
150			\a Media literacy
450			\a Mass media literacy
550			\a Information literacy \w g
670			\a 95042682: DeGaetano, G. Screen smarts : a family guide to media literacy, 1996.
670			\a Hennepin \b (Media literacy)
670			\a 74077021: Houk, A.S. Media literacy ... 1974.
670			\a 89130471: Limburg, V.E. Mass media literacy, c1988.
670			\a 94022004: Silverblatt, A. Media literacy : keys to interpreting media messages, 1995.
670			\a 96035023: Silverblatt, A. Dict. of media literacy, 1997.
680			\i Here are entered works on a person's knowledge and ability to use, interpret and evaluate modern means of communication, such as radio, television, newspapers and magazines.
750		0	\a Éducation à l'image \2 ram
898			\a elena.redkina \b V \c 20110731

The authority record is created

Browse - BibAlex

Subject: Media literacy

Hits	Terms
1	Media literacy [SRC=MnMHCL]
-	1 Media Marketing Discours - Campagne Election Electoral Gouvernement Public Media, Mass
-	1 Media Mediatisation Mediologie Moderne Moyen Politique Signification
-	1 Media Mediatisation Moyen Oraison Texte Textuel Media, News
-	Media planning in advertising
-	Media-planning (publicité)
-	Media, Potting
+ 7	Media programs (Education) [SRC=DLC]

Subject: Media literacy

View Page Up Page Down Skip Delete

Result

BA Catalog

« 30 « 20 « 10 « Back | Forward » 10 » 20 » 30 »

You searched Bibliotheca Alexandrina - Subject: media literacy

Hit Count

Scan Term

1

Media literacy

1

Media Marketing Discours - Campagne Election Electoral Gouvernement Public

II. Linked Subject Headings in Arabic, French and English as equivalents

- How it works?
- What is the result?

How it Works

- In the daily report, the cataloger finds the subject “Database management”

Adding the French and Arabic equivalent terms to the English Authority record

MARC Editor - BibAlex

008			060821 j anannbabn a ana
010			\a sh 85035848
035			\a (OCoLC)oca02077602
040			\a DLC \b eng \c DLC \d DLC \d EG-AIBA
053		0	\a 0478.9.D3
150			\a Database management
450			\a Data base management \w n
450			\a DBMS (Computer science)
450			\a Generalized data management systems
450			\a Systems, Database management
450			\a Systems, Generalized database management
550			\a Electronic data processing \w g
670			\a UMI business vocab. \b (DBMS use Data base management systems)
750		7	\a إدارة قواعد البيانات
750		7	\a Bases de données \x Gestion \2 rari
998			\a elena.redkina \b V \c 20060821

(z) Authority data Subject

MARC 21 Authority: Authority OPAC

OPAC

Adding the English and French equivalent terms to Arabic Authority record

The screenshot shows the MARC Editor interface for BibAlex. The main area displays a table of MARC fields for an authority record. Two rows are circled in red: the 040 field and the 750 field. The 040 field contains the code 'EG-AIBA' in both Arabic and English. The 750 field contains the Arabic term 'شعبان خليفة' and its English and French equivalents: 'Database management' and 'Bases de données / Gestion de ram'.

Field Number	Subfield	Value
008		061227L aarababab n ana d
040		\a EG-AIBA \b ara \c EG-AIBA
083	0 4	\a 005.74 \2 21 \5 EG-AIBA
150		\a إدارة قواعد البيانات
450		\a إدارة قواعد المعلومات
450		\a نظم إدارة قواعد البيانات
550		\w g \a معالجة البيانات إلكترونيا
667		\a [مقترح]
670		\a مع العمل المفهرس: إسماعيل، محمود. إنشاء وإدارة قواعد البيانات باستخدام office 2000, 2003. دعم لمستخدمي
670		\a الموسع \b (إدارة قواعد المعلومات) = Data base management
675		\a شعبان خليفة
750	0	\a Database management
750	7	\a Bases de données \x Gestion de ram
998		\a iman.khairi \b V \c 20061226

At the bottom of the window, there is a navigation bar with the following elements:

- Dropdown menu: (z) Authority data
- Dropdown menu: (y) Subject
- Navigation buttons: < >
- Text: MARC 21 Authority: Authority
- Buttons: OPAC, OPAC2

Adding the English and Arabic equivalent terms to the French Authority record

Field	Subfield	Value
008		110719ij acz bab n a a d
040		\a EG AIBA \b fr \c EG AIBA \f ram
150		\a Bases de données \x Gestion
450		\a Dictionnaires de bases de données
450		\a SGBD
450		\a Systèmes de gestion de bases de données
550		\w g \a Informatique documentaire
550		\a Sites Web \x Gestion
680		\a vedette matière nom commun. S'emploie en tête de vedette.
750	0	\a Database management
750	7	\a إدارة قواعد البيانات
998		\a elena.redkina \b \c \d 20110719

(z) Authority data (y) Subject

< > MARC 21 Authority: Authority OPAC

OPAC1 / OPAC2 / OPAC3

Result

- User's ability to search materials by subject in three languages and different scripts.

Searching by Arabic Term

فهرس مكتبة الاسكندرية

ابحث هنا

« 30 » « 20 » « 10 » « الى الامام » | « الى الخلف » « 10 » « 20 » « 30 »

مليت البحث عن - مكتبة الاسكندرية - الموضوع: اداره عامه

نتائج البحث	عدد
إدارة شؤون الموظفين	1
إدارة قواعد البيانات	15
إدارة قواعد البيانات - برامج حسابات	9
إدارة قواعد المعلومات	9
إدارة موارد المعلومات	9
إدارة نظم المعلومات	9
أطب الأطفال	432
أطب الأطفال - بيلوجرافيات	6
أطب الأطفال - تاريخ و نقد	21
أطب الأطفال - تاريخ و نقد - مراجعات الكتب	1
انطب الاطفال - تراجم	2
انطب الاطفال - مجموعات	2
انطب الاطفال - مقالات و محاضرات	1
أطب البيكاريسك	1
أطب الخيال العلمي	12

« 30 » « 20 » « 10 » « الى الامام » | « الى الخلف » « 10 » « 20 » « 30 »

بحث سريع

مصطلحات البحث:

في أي مكان

كتب

صفح

التترات النورية

بحث

دخول المستبر

بحث بسيط

بحث بالداخل

بحث متقدم

بحث خارجي

حفظ/تحميل

اتصال جيد

ملاحظات

مساعدة

تاريخ البحث

الموضوع: اداره عامه

عن المكتبة

نظرة عامة

رسالة المكتبة و أهدافها

رسالة مدير المكتبة

الهيكل التنظيمي

المكتبة و المكتبات المتخصصة

المراكز البحثية و الثقافية

مركز الفية السماوية العلمي

المتاحف

المعارض الدائمة

مركز المؤتمرات

منتدى الحوار

الياتورااما الحضارية

قيستا

أرشيف الإنترنت

Searching by the English Term



The screenshot shows the BA Catalog search results for the term "Database management". The page includes a search bar, navigation links, a search results table, and a quick search sidebar.

Search here... 

BA Catalog

« 30 « 20 « 10 « Back | Forward » 10 » 20 » 30 »

You searched Bibliotheca Alexandrina - Subject database management

Hit Count	Scan Term
145	 Database management
1	Database management – Auditing – Standards.

Quick Search
Search Terms:

 Anywhere
 Books
 Newspapers
 Periodicals

Cross references

The screenshot shows the BA Catalog interface. On the left is a navigation menu with items like 'About the Library', 'Overview', 'Mission & Objectives', 'Director's Message', 'Organization', 'Library & Specialized Libraries', 'Academic Research Centers', and 'Planetarium Science Center'. The main content area displays search results for 'Database management' with 145 hits. Below this, a section titled 'Related Terms' lists several terms with their respective hit counts, including 'Bases de données -- Gestion' (6 hits), 'إدارة قواعد البيانات' (15 hits), 'Data compression (Computer science)' (8 hits), 'Data warehousing' (15 hits), 'File organization (Computer science)' (12 hits), 'IMS/VS (Computer system)' (1 hit), and 'Text processing (Computer science)' (4 hits). A 'Notes' section provides additional context for the Arabic term. On the right, there is a 'Quick Search' section with a search box, radio buttons for search scope (Anywhere, Books, Newspapers, Periodicals), and a 'Search' button. A 'Patron Access' link is also visible at the bottom right.

	Hits	Original Term
Search on term	145	Database management
	Hits	Related Terms
Search on term	6	Bases de données -- Gestion
Notes.		
Vedette matière nom commun. S'emploie en tête de vedette.		
Search on term	15	إدارة قواعد البيانات
Search on term	8	Data compression (Computer science)
Search on term	15	Data warehousing
Search on term	12	File organization (Computer science)
Search on term	1	IMS/VS (Computer system)
Search on term	4	Text processing (Computer science)

Searching by the French Term

chercher ici...

Catalogue de la BA

Nb. de documents sous:

Rechercher 6 Bases de données -- Gestion

Notes:
Vedette matière nom commun. S'emploie en tête de vedette.

Nb. de documents sous:

Rechercher 145 Database management

Rechercher 1 IDS-II (système informatique)

Notes:
Vedette matière nom commun. S'emploie en tête de vedette.

Rechercher 15 إدارة قواعد البيانات

Quick Search

Search Terms:

Partout

Livres

Journaux

Périodiques

Search

References used to create Subject Headings:

- English subject headings are assigned from (**LCSH**)
- Qaimat ruus al-mawduat al-tibbiyah = Medical subject headings (**MeSH**), National Library of Medicine,
- Arabic subject headings are assigned from (Qa'imat ru'us al-mawdu'at al-'Arabiyah al-qiyasiyah lil-maktabat wa-marakiz al-malumat wa-qawaid al-bayanat (**QRMAK**) and other lists thesaurus as appropriate.
- al-Fayṣal : A List of subject headings used for Islamic studies
- French Subject Headings are assigned by the French Catalogers from Les notices d'autorité de BN-OPALE PLUS (**Rameau**).

III. Similarity and differences between the BA and the LC/OCLC authority files

- The structure of the BA Authority file conforms to the of the LC/OCLC
- BA has established procedures for creating standard forms for subject headings

IV. Advantages

- Possibility to create an authority center that would serve the Arab world through cooperative arrangements.
- Some of the solutions developed at the BA can lead to future developments in resource sharing and cooperation on a global scale
- Their system provides multi-lingual displays in one step.

Disadvantages

- Extensive work has been done to manually map concepts from different knowledge organization systems (LSCH, Rameau, and Arabic Subject headings), which is a tedious and expensive process
- Mapping depends on the cataloger's understanding of the concepts in these three languages
- Sometimes the concepts can be used to mean something else in different languages.

Next steps

- Developing a method to automatically map concepts, using different ontology matching techniques.
- Linking these vocabularies to each other, regardless of their language, is a first step towards searching and browsing across collections
- SKOS and the Arabic alphabet

Questions?

Shukran – Merci - Thank You!