Effectiveness of Name Searching in Web OPAC:
From Authority Control to Access Control

Veerankutty Chelatayakkot   V Jalaja

Abstract

Since the advent of Unicode, electronic representation of regional language script is not as issue. However, transliteration is inevitable in a database designed to meet the needs of people who does not know that language. By transliteration, a unique name in one language or culture may have variant spelling in another language. The cataloguer maintainsname authority file to solve this problem. This article analyse the optionsfor retrieving transliterated Arabic names in major OPACs.

Keywords : OPAC, Web OPAC, Information Retrieval.

0.   Information Retrieval and Name Searching

Information Retrieval is the process of finding some desired information from a store of information or database. It implies the concept of selectivity. Information recovery is not the same as information retrieval unless there has been a selection process. Name searching is one of the major activities in information retrieval. The name searching may be either from free-text or from a database.

Database name matching technology has long been used in criminal investigations, counter-terrorism efforts, and in a wide variety of government processes such as Visa processing. In this technology, a name is compared to names contained in one or more databases to determine whether there is a match. Sometimes this matching operation may be straightforward exact match, but often the process is more complicated. Two names may not match exactly for a wide variety of reasons and yet still refer to the same individual. The challenges of name matching are greatly increased by: a) structure of names: people uses different structures in rendering their names i.e., the surname come first, before the given name, b) variant spelling of a unique name, c) pseudonyms, d) name with suffix / prefixes and e) transliterated names from native script (e.g. Arabic, Chinese, Cyrillic, Hindi).[1] As against manual searching, a slight difference in input may affect the effectiveness of Information Retrieval in the context of database name searching. In library context, maintaining name authority files solves these problems.

1.   Transliterated Names and Authority Control

The name authority file links variant forms of heading to the preferred form of heading to help preparation of cross-references. An author may have different names/pseudonyms or variant spellings may be occurred due to transliteration of name. Transliteration is the process of formulating a representation of word in one language using the alphabet of another language. The aim of transliteration is to represent the script of a source language by using the letters or symbols of another script, usually in accordance with the orthographical conventions of the target language. By the result, a unique name in one language or culture may have variant spelling in another language. Further, owners of these names may take certain liberties with spelling of their names. This is the same condition, where an author uses different names or pseudonyms. The cataloguer should create cross-references from variant spellings (different name) to the heading.

Since the advent of Unicode, electronic representation of non-Roman characters is not an issue. As a universal character code set, Unicode provides a unique number to every character used in modern scripts thought out the world. The Unicode Standard is the universal character-encoding standard used for representation of text for computer processing. Versions of the Unicode Standard are fully compatible and synchronized with the corresponding versions of International Standard ISO/IEC 10646. For example, Unicode 4.0 contains all the same characters and encoding points as ISO/IEC 10646:2003. The Unicode Standard provides additional information about the characters and their use.

The application of Unicode helps to reduce the relevance of transliteration up to an extent. However, transliteration is inevitable in a database, which is designed to meet the needs of people all over the world. In a database having Arabic names entered in Arabic using Unicode can't be of use to such people who have no knowledge of Arabic language. It means that Unicode can solve the problem of electronic representation of regional languages, but such a database can be of use only to a limited people who know that language. So transliteration of personal names from regional languages is essential to meet the need of international users.

2. Information Retrieval and Authority Control

The quality of online databases is very much related to the quality of data in the database. There will be variant spelling, omission, deletion or typing errors for a unique entity. By the acceleration of electronic publishing, there is increase in number of online databases. The challenge is to get correct information from the databases. One aspect of improving data quality is detection of variant name for a unique entity and link them to improve searching. This is similar to authority work. The technique of authority work has vide implication in the effectiveness of Information Retrieval in online databases.[2]

When a user keys in a known form of a heading, the system follows internal linkage and displays the requested item even though the preferred form of heading might be quiet different from the form entered. There should be direct linkage between the form of heading entered and record displayed. The authority control mechanism works as invisible mechanism so far as a user is concerned.

In the case of transliterated names authority file, the cross-references are mostly variant spellings of personal names. While transliterating names in to Roman script, by many reasons, variant spellings denote a unique name. The usage of variant spellings compels the cataloguer to create cross-references to the selected heading.

3. Options for Retrieval of Transliterated Arabic Names in Major OPAC

A review of options for retrieval of transliterated Arabic names in major OPACs will help to understand how the problems in name searching are solved in OPACs. The review is based on Library of Congress Name Authority File (LCNAF), which is the largest authority file in the world. The headings available in both libraries i.e., Library of Congress and reviewed library, were taken. Each heading and its see references in LCNAF were searched in reviewed OPAC and total hits were analyzed. Fifteen OPACs were compared with LCNAF. The major options for name searching in OPAC are:

1. *See References* (4xx in MARC). Some OPAC automatically links to variant forms of heading by using See Reference as provided in the LCNAF. For example,

Fig.1 OPAC of Lomia Linda University http://catalog.llu.edu

2.   *Inverted Heading*: Inverted heading are used to solve the problems in name searching occurred by the structure of name. A personal name may consist various elements denoting their family name, house name, place name, honorific titles etc. The structure and rendering of names have no relevance in a automated system, where the system automatically inverts the headings. For example,

Fig.2 OPAC of University of South Africa http://oasis.unisa.ac.za

3. *Nearby Authors*: Some OPACs provide the option to display nearby authors in order to solve the problem of spelling variations in input. For example,
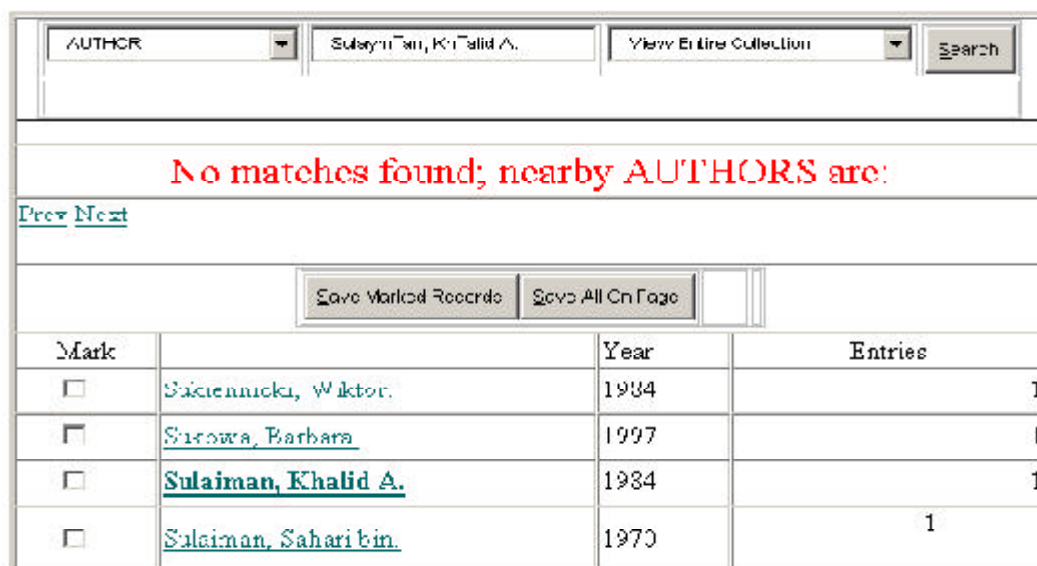
Fig.3 OPAC of Colorado State University http://catalog.library.colostate.edu/

Analysis :

No.	Name of the University/OPAC	Heading as in LCNAF	Inverted Reference	See References	Nearby Author's
1	COPAC	Yes	Yes	No	No
2	Duke	Yes	Yes	No	Yes
3	Edith Cowan	Yes	Yes	No	No
4	Lona Linda	Yes	No	Yes	Yes
5	Ottawa	Yes	No	No	Yes
6	Wales	Yes	No	Yes	No
7	St.Andrew	Yes	No	Yes	Yes
8	South Africa	Yes	Yes	Yes	Yes
9	Colorado	Yes	No	Yes	No
10	Essex	Yes	Yes	No	Yes
11	Ohio	Yes	Yes	Yes	No
12	Oxford	Yes	Yes	Yes	No
13	Trinity	Yes	No	No	No
14	Australia	Yes	Yes	Yes	Yes
15	Nevada	Yes	Yes	Yes	No

The analysis of above 15 OPAC shows following findings:

1.   All headings are similar to Library of Congress headings

2.   Most OPACs have the option of Inverted Reference (i.e. Inverted Heading)

3.   Most   OPACs provide See References

4.   Only 6 OPACs have option of Nearby Authors.

4.   Authority Control Versus Access Control

The access control record is the next generation of the authority record. It may be called as "super authority record" because of the potential it contains for enriched information for indexing.[3] Access control records can be linked both to bibliographic records, to collocate all manifestations of a work, and to other related access control records, to collocate related works. The basic concept behind the access control record is removing both the label and notion of "authority". While authority control record declare a heading as "authorized" form, access control record links all variations without declaring one heading as authorized form. Access control records allow users to choose their preferred form or name, or to have displayed a default heading. It allows for more flexibility in display. The concept of access control entirely contradicts the whole second part of AACR-2, which is devoted to painstaking rules for how to construct authorized forms of names and titles.

The idea behind the access control is that an entity can be known by more than one name. An individual is an entity but may be called by different names by different people at different times in life. In the international realm living persons have name representations in many languages and script. The possible solution is that instead of selecting one form of name as heading, give chance to select the one of the heading as default headings according to their interest (local convenience) and give non-default as reference. We can hope good news from technocrats

5.   References

1.   Patman, Frankie., and Paul Thompson (2003)Names: A Frontier in Text Mining*in Lecture Notes in Computer Science*, pp.27-38,edited by H Chen, Springer-Verlag Heidelberg .

2.   Tillet, Barbara B (2001) " Authority Control on the Web" Conference on Bibliographic Control in the New Millenium USA, November 11-15,2000, organized by Library of Congress URL: http://www.loc.gov/catdir/bibcontrol/tillett-_paper.html (Accessed on 23/11/2004)

3.   Barnhart, Linda (1996)"Access Control Records : Precepts and Challenges" Authority Control in 21st Century :An invitational Conference Ohio, USA, March 31-April1,organized by OCLC Online Computer Library Center Inc. URL http://www.lib.byu.edu/dept/catalog/authority (Accessed on 23/11/2004)

6.   Appendix

Details of Data Analysis

1.   COPAC (http://www.copac.ac.uk/copac)

Entry in LC Authority Record	Nature of Entry in LC	Hits in COPAC	Nature of Entry In COPAC
Haykal, Muhammad Husayn, 1888-1956	Heading	26	Heading
Haikal, Muhammad Husain, 1888-1956	See Ref.	0	No. Ref.
Heikel, Mohammed Husein, 1888-1956	See Ref.	0	No. Ref.
Heikel, Mohamed H 1888-1956	See Ref.	0	No. Ref.
Basheer, Vaikom Muhammad, 1910	Heading	21	Heading
Vaikom Muhammad Basheer, 1910	See Ref.	19	Inverted
Muhammad Bas¯ir, Vaikkam, 1910	See Ref.	0	No. Ref.
Vaikkam Muhammad Bas¯ir, 1910	See Ref.	0	No. Ref.

2.   Duke University , USA (http://catalog.library.duke.edu/)

Entry in LC Authority Record	Nature of Entry in LC	Hits in DUKE Uni.	Nature of Entry in DUKE Uni.
R¯ashid, Rushd¯I	Heading	2	Heading
Rushd¯i R¯ashid	See. Ref	0	No Reference
Rashed, Roshdi	See. Ref	0	No Reference
Husain, '¯Amir Liy¯aqat	Heading	1	Heading
Amir Liy¯aqat Husain	See. Ref	0	Inverted
Hussain, Aamer Liaquat	See. Ref	0	No Reference

3.   Edith Cowan University , Australia (http://library.ecu.edu.au/)

Entry in LC Authority Record	Nature of Entry in LC	Hits in Edith Cowan Uni.	Nature of Entry in Edith Cowan Uni.
Mohamed Yusoff Ismail	Heading	1	Headings
Ismail, Mohamed Yusoff	See. Ref	0	No Reference
Ahmad Mansoor	Heading	1	Heading
Mansoor Ahmad	See. Ref	11	Inverted

4.   Loma Linda University , USA (http://catalog.llu.edu/)

Entry in LC Authority Record	Nature of Entry in LC	Hits in Loma Linda Uni. OPAC	Nature of Entry in Loma Linda Uni. OPAC
Ahmed, M. Samir	Heading	1	Near by Authors
Ahmed, Samir	See. Ref	1	See Ref.
Ahmed, Mahmoud Samir	See .Ref	1	Heading
Haykal, Muhammad Husayn, 1888-1956	Heading	1	Heading
Haikal, Muhammad Husain, |d 1888-1956	See .Ref	1	See. Ref
Heikel, Mohammed Husein, |d 1888-1956	See .Ref	1	See. Ref

5.   University of Ottawa . (http://www.biblio.uottawa.ca/orbis-e.php)

Entry in LC Authority Record	Nature of Entry in LC	Hits in Uni. Ottawa	Nature of Entry in Ottawa Uni. OPAC
Hassan Ibrahim	Heading	1	Nearby authors
Ibrahim Hassan	See. Ref	1	Heading
Ibrahim, Datuk Haji Hassan	See .Ref	111	Related Heading
Yousef, Yousef A.	Heading	1	Heading
Yīusuf, Yīusuf A.	See. Ref	0	No reference
Yousef A. Yousef	See .Ref	1	Nearby Aauthors

6.   University of Wales., UK (http://library.bangor.ac.uk/)

Entry in LC Authority Record	Nature of Entry in LC	Hits in Uni. of Wales	Nature of Entry in Uni. of Wales OPAC
Jinnah, Mahomed Ali, 1876-1948	Heading	1	Heading
Muhammad 'Alī i Jinnah, 1876-1948	See. Ref	1	See. Ref
Quaid-i-Azam	See .Ref	1	See. Ref
Faruqee, Rashid, 1938-	Heading	1	Heading
Rashid Faruqee	See .Ref	1	See. Ref
Faruqee, R	See. Ref	1	See. Ref

7.   University of St.Andrews (http://138.251.116.3/)

Entry in LC Authority Record	Nature of Entry in LC	Hits in SAULCAT	Nature of Entry in SAULCAT
Ali, Mohamed	Heading	1	See. Ref
Ali, Muhammad	See. Ref	1	Heading
Mohammad Ali, |c Maulana,	See. Ref	0	No. Ref.
Sulaiman, Khalid A.	Heading	1	Heading
Sulaymīan, Khīalid A.	See. Ref	1	Nearby auth.

8.   University of South Africa (http://oasis.unisa.ac.za/)

Entry in LC Authority Record	Nature of Entry in LC	Hits in Nat. Lib. of S. Africa	Nature of Entry in Nat. Lib. of S.Africa
Ishaq Ashfaq	Heading	1	Heading
Ashfaq Ishaq,	See. Ref	1	Inverted heading
Rasheed, Sadig	Heading	1	Heading
Rash¯ id, Sad¯ iq	See. Ref	1	See. Ref
Rasheed, Sadiq	See. Ref	1	See. Ref

9.   Colorado University USA (http://catalog.library.colostate.edu/)

Entry in LC Authority Record	Nature of Entry in LC	Hits in Colorado State Uni. Lib.	Nature of Entry in Colorado State Uni.
Salam, Abdus, 1926	Heading	1	Heading
Muhammad 'Abd al-Sal¯ am, |d 1926-	See. Ref	1	See. Ref
Salam, Muhammad Abdus, |d 1926-	See. Ref	1	See. Ref
Sulaiman, Khalid A.	Heading	1	Heading
Khalid A. Sulaiman	See. Ref	1	See. Ref
Kh¯ alid A. Sulaym¯ an	See. Ref	1	See. Ref

10.   University of Essex, UK   (http://serlib0.essex.ac.uk/)

Entry in LC Authority Record	Nature of Entry in LC	Hits in Essex Nature	Nature of Entry in Essex OPAC
Basheer, Tahseen	Heading	1	Heading
Bashir, Tahseen	See. Ref	0	Not found
Tahseen Basheer	See. Ref	1	Inverted reference
Ahmad Ibrahim	Heading	1	Nearby Authors
Ahmad Mohamed Ibrahim	See. Ref	0	Not found
Ahmad bin Mohamed Ibrahim	See. Ref	101	Heading

11.   Ohio University, Athens (http://www.library.ohiou.edu/)

Entry in LC Authority Record	Nature of Entry in LC	Hits in Essex Ohio	Nature of Entry in Ohio
Bin Laden, Osama	Heading	1	Heading
Usama bin Laden	See. Ref	1	See. Ref.
Ibn L¯ adin, Us¯ amah, |	See. Ref	1	Inverted reference
Zakaria bin Haji Ahmad	Heading	1	Heading
Ahmad, Zakaria bin Haji	See. Ref	1	See. Ref
Zakaria Haji Ahmad	See. Ref	1	See. Ref

12.   Oxford Library and Information System. (http://www.lib.ox.ac.uk/olis/)

Entry in LC Authority Record	Nature of Entry in LC	Hits in OLIS	Nature of Entry in OLIS
B̄ ar̄ ud̄ i, 'Abd All̄ ah 'Umar	Heading	1	Heading
Abd All̄ ah 'Umar al-B̄ ar̄ ud̄ I	See. Ref	1	See. Ref
Husayn̄ i, 'Abd All̄ ah ibn 'Umar al-B̄ ar̄ ud̄ i	See. Ref	0	Not Found
Bashier, Zakaria	Heading	1	Heading
Zakaria Bashier	See. Ref	1	Inverted Ref.
Bash̄ ir, Zakar̄ iȳ a	See. Ref	0	Not found

13.   Australian National University, Canberra (http://library.anu.edu.au/search~S1/)

Entry in LC Authority Record	Nature of Entry in LC	Hits in ANU	Nature of Entry in ANU
Abdesselem, Mohamed	Heading	1	Heading
Mohamed Abdesselem	See. Ref	1	See. Ref
Muhammad 'Abd al-Sal̄ am	See. Ref	1	See. Ref
Ish̄ aq, Muhammad Qamar, 1961-	Heading	1	Heading
Ishaque, Mohammad Qamar,1961-	See. Ref	1	See. Ref
Mohammad Qamar Ishaque, 1961	See. Ref	1	See. Ref

14.   Trinity Theological College, Australia (http://www.trinity.qld.edu.au/)

Entry in LC Authority Record	Nature of Entry in LC	Hits in Triniti OPAC	Nature of Entry in Triniti Theo. College
Muhammad Abul Quasem	Heading	1	Heading
M. A. Quasem	See. Ref	0	Not Found
Abul Quasem	See. Ref	1	Heading
Ahmad, Bashiruddin Mahmud	Heading	1	Heading
Bashiruddin Mahmud Ahmad	See. Ref	1	Heading
Mahmood Ahmad, Bashir-ud-Din	See. Ref	0	Not found

15.   University of Nevada, Rino(http://www.library.unr.edu/)

Entry in LC Authority Record	Nature of Entry in LC	Hits in i UNLOPAC	Nature of Entry in Uni. of Nevada Lib.
Badaw̄ i, Muhammad Mustafá	Heading	1	Heading
Badaw̄ i, Mustafá	See. Ref	1	See. Ref
Badawi, M. M.	See. Ref	1	See. Ref
Hussain, Asaf	Heading	1	Heading
Asaf Hussain	See. Ref	1	Inverted Heading
Hüseyin, Asaf	See. Ref	0	Not found

About Authors

Veerankutty Chelatayakkot, Junior Librarian, Cochin University of Science and Technology, Kochi-22
E-mail : veerankutty@cusat.ac.in

Dr. V. Jalaja, Presently she is working as Reader, Dept. of Library and Information Science, University of Calicut. She has 15 years experience in teaching. A number of Ph.D has produced under her guideship. She has a number of publications in her credit. Her area of interest is Bibliomining.