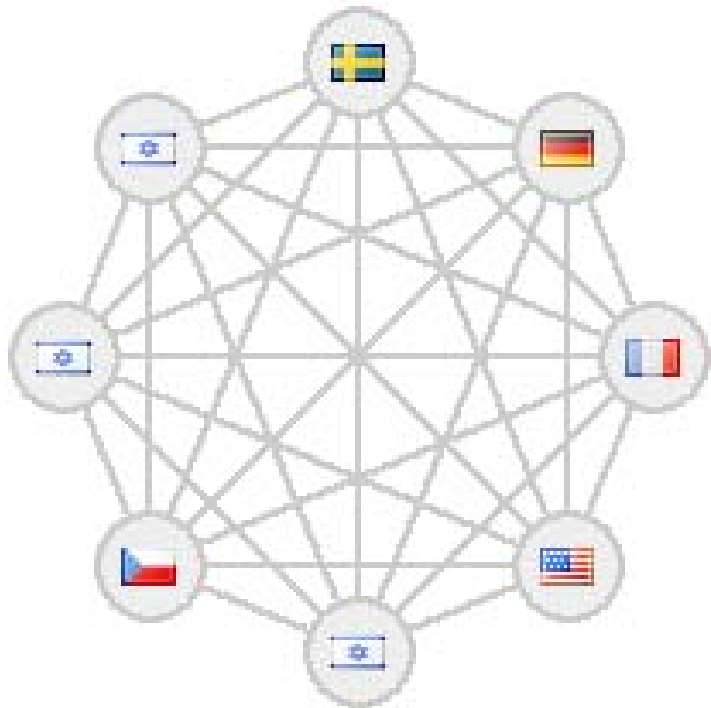


The Virtual International Authority File



Edward O'Neill
Thomas Hickey

2009 RLG Partnership Annual
Meeting
Boston, MA
June 1-2, 2009

Goals of the Virtual International Authority File

- Link national-level authority records
- Expand the concept of universal bibliographic control
 - Allow national or regional variations in authorized form to co-exist
 - Support needs for variations in preferred language, script and spelling
- Play a role in the emerging semantic web

VIAF participants

- Full partners:
 - Bibliothèque nationale de France
 - Deutsche Nationalbibliothek
 - Library of Congress
 - OCLC
- Other participants
 - Sweden
 - Czech Republic
 - Israel (4 files!)
 - Egypt (Bibliotheca Alexandrina)
 - Vatican

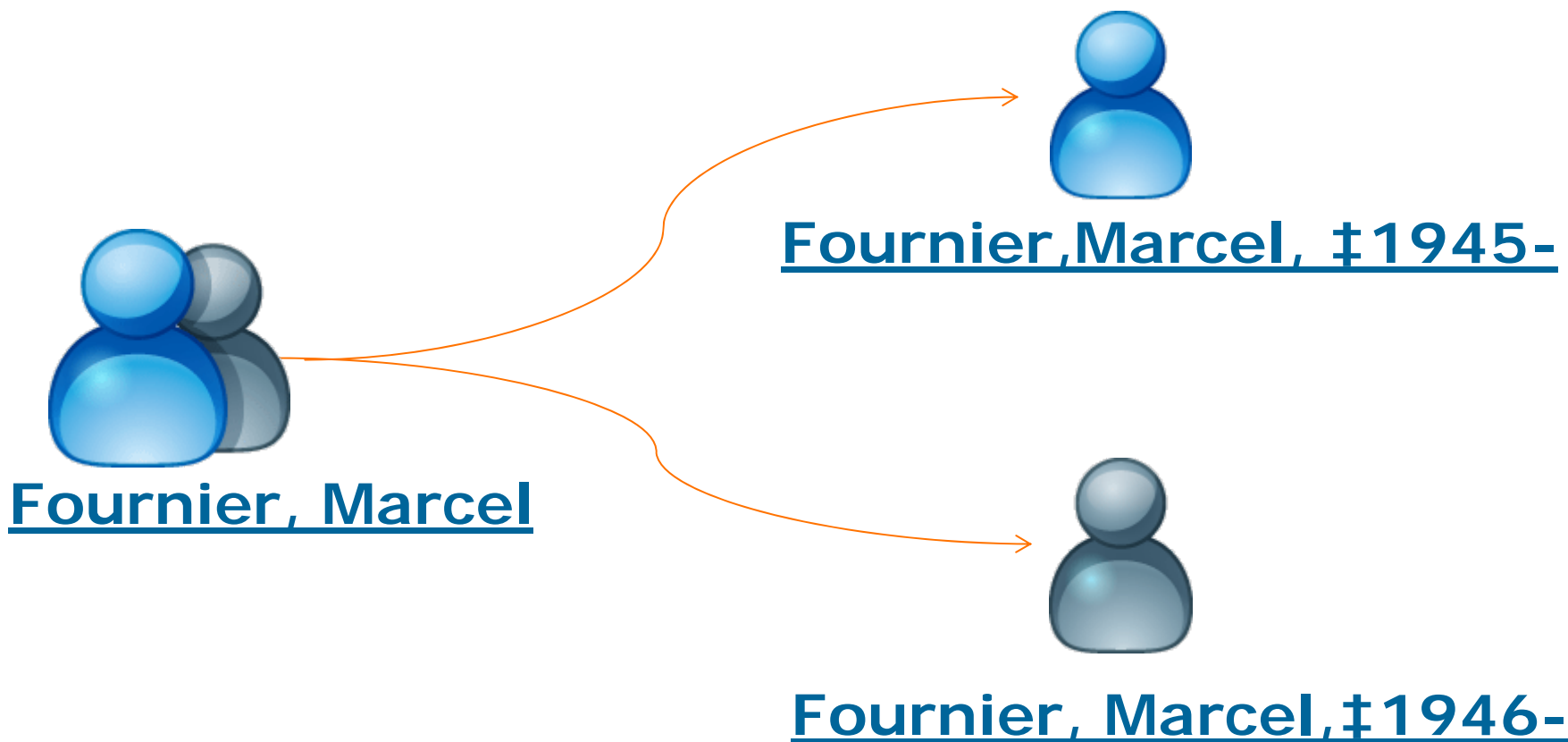
Also invited

- Australia
- Italy
- Spain
- Portugal

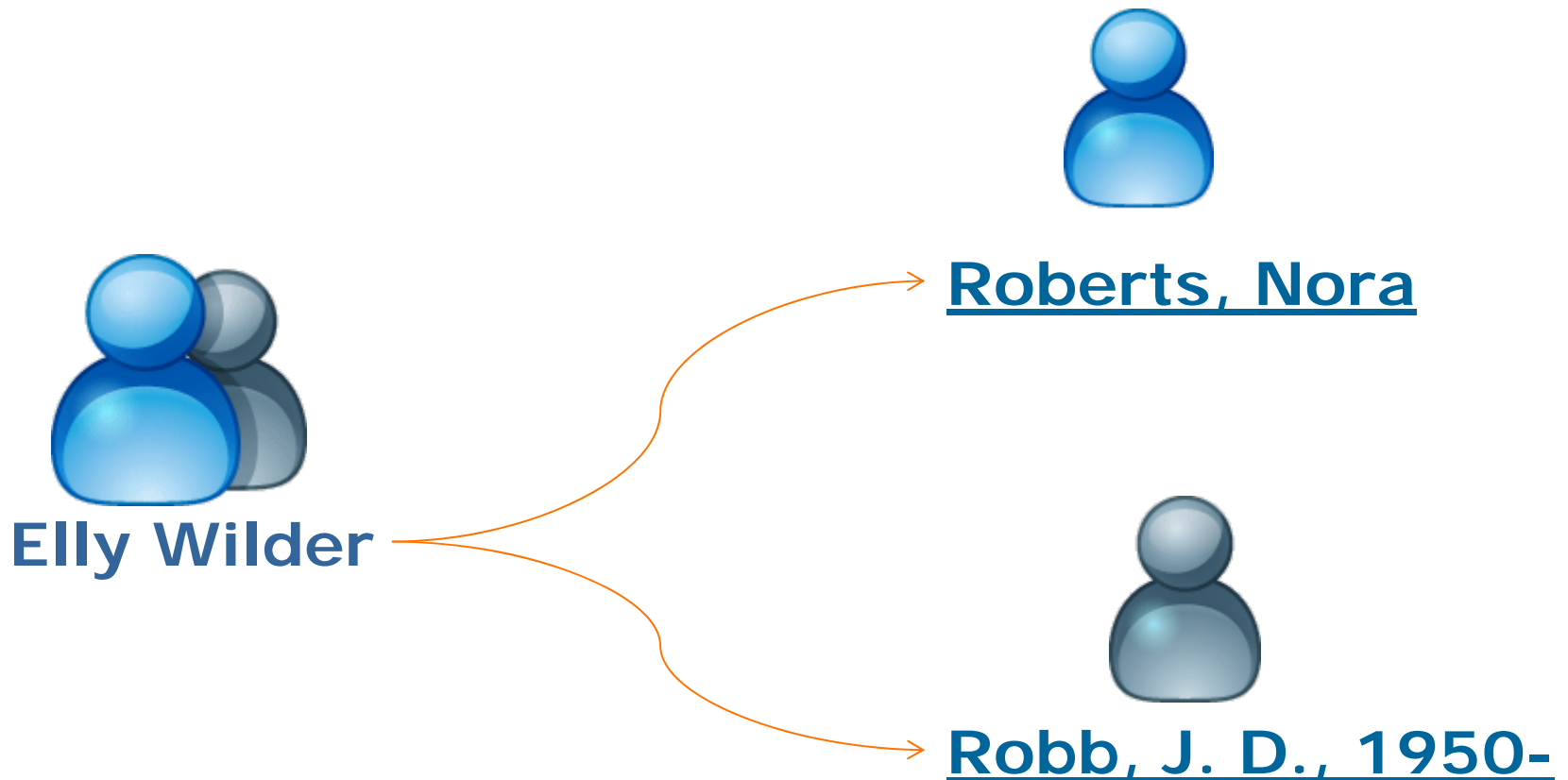
Scope of VIAF

- Personal names
 - Geographic
 - Corporate
 - Title
 - Family
 - Events
-
- Everything but concepts are considered in scope
 - National level, but willing to consider other sources

A standard problem:
One name, multiple people



Another standard problem: One person, multiple personas



Another problem: One persona, many representations

ЧЕХОВ, АНТОН ПАВЛОВИЧ, 1860-1904 

1860-1904 , צ'כוב, אנטון פבלוביץ, 

Čechov, Anton Pavlovič, 1860-1904  

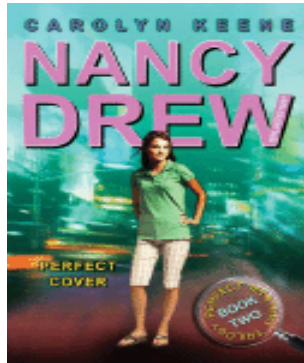
Čehov Anton Pavlovič 1860-1904 

Chekhov, Anton Pavlovich, 1860-1904  

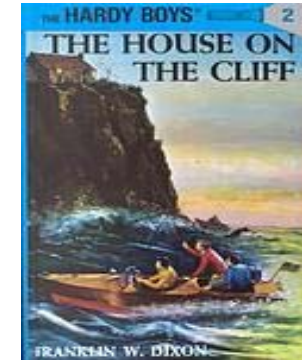
Čechov, Anton P. 1860-1904 

viaf.org/viaf/29541064

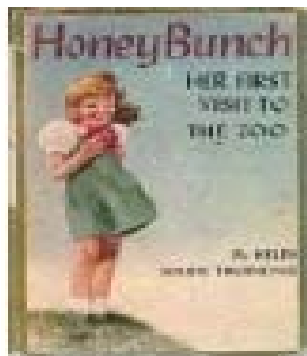
Complicated names



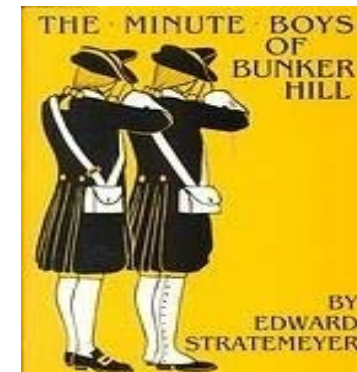
By: Carolyn Keene
(Ed Stratemeyer,
Mildred Wirt, et al)



By: Franklin W. Dixon
(Ed Stratemeyer, et al)

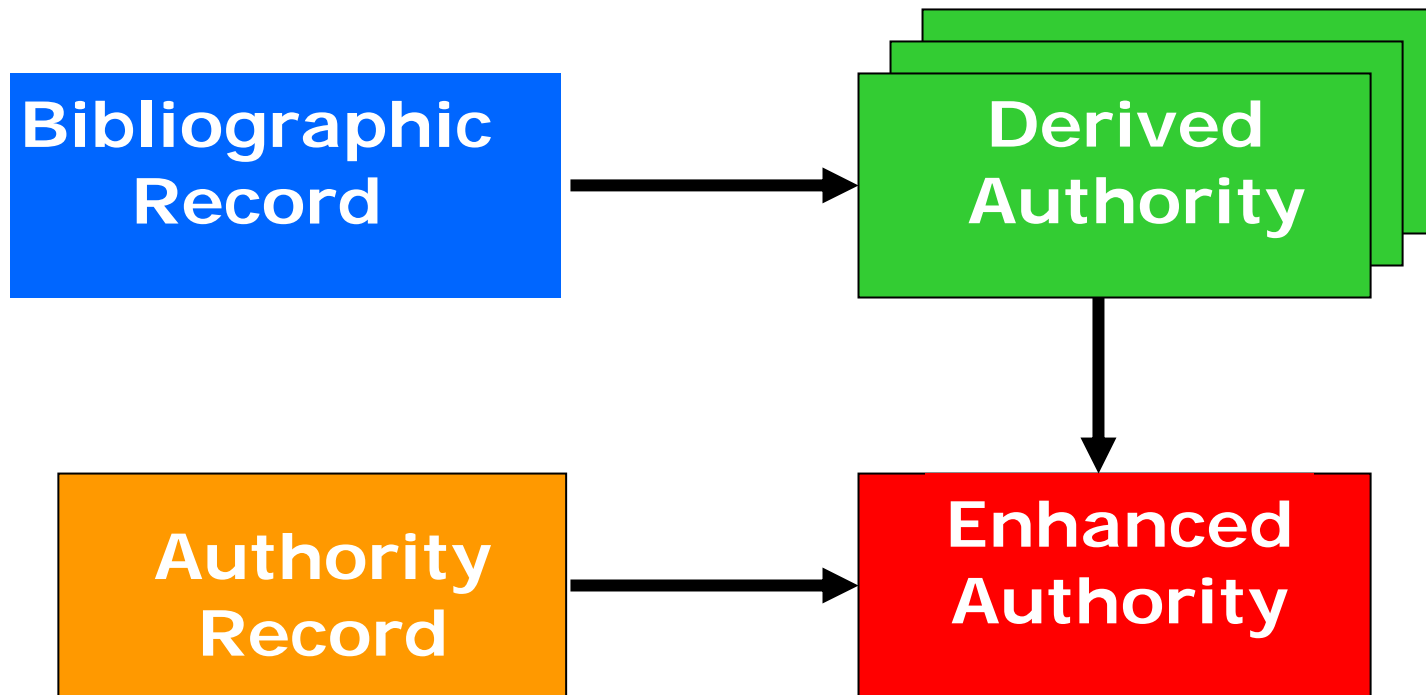


By Helen Thorndike
(aka Mildred Wirt)



By Edward Stratemeyer

Enhancing the authorities



Brief LC authority

010 n 84044261

040 DLC \$c DLC \$d DLC

100 1 Larson, Jack.

670 Thomson, V. The cat, c1982: \$b t.p. (Jack Larson)

Mining the bibliographic record

LDR 00826ccm 2200289 a 4500
1 ocm10025532
5 20031229650847.0
8 840627s1982 nyuuua n eng
10 \$a 84758340
40 \$a DLC \$c DLC
19 \$a 17706440
20 \$c \$2.95
28 22 \$a 48418 \$b G. Schirmer
45 2 \$b d198006 \$b d198007
48 \$b va01 \$b ve01 \$a ka01
50 00 \$a M1529.3 \$b .T
100 1 \$a Thomson, Virgil, \$d 1896-
245 14 \$a The cat : \$b duet for soprano and baritone / \$c
Virgil Thomson [words by Jack Larson].
260 \$a New York : \$b G. Schirmer, \$c c1982.
300 \$a 1 score (11 p.) ; \$c 31 cm.
500 \$a For soprano, baritone, and piano.
650 0 \$a Vocal duets with piano.
600 10 \$a Larson, Jack \$x Musical settings.
700 1 \$a Larson, Jack.

Language
LC Control Number

LC Classification
Usage
Title
Publisher
Place of Publication

Date of Publication
Material Type
Authenticity

Information in bibliographic records

From the bibliographic records we gain significant additional information about Jack Larson:

- He is a lyricist
- His primary subject area is music
- He was published in the 80s and 90s by G. Schirmer and Belwin Mills in New York
- Worked with Virgil Thomson and Gerhard Samuel
- Jack Larson is the only name he has used on his publications
- Etc.

Enhanced authority record

```
00824nz    2200301n    4500
 0   1 oca01144962
 1   5 19840809154202.7
 2   8 840702n| acannaab|          |n aaa |||
 3  10      $a n    84044261
 4  40      $a DLC $c DLC $d DLC
 5 100 1    $a Larson, Jack.
 6 670      $a Thomson, V. The cat, c1982: $b t.p. (Jack Larson)
 7 903      $a 84758340 $9 1
 8 903      $a 93710923 $9 1
 9 910 11   $a the cat $b duet for soprano and baritone $9 1
10 910 11   $a sun like $b on a poem by jack larsen $9 1
11 921      $a g schirmer $9 1
12 921      $a belwin mills publ corp $9 2
13 922      $a nyu $9 2
14 930      $a jack larsen $9 1
15 940      $a eng $9 2
16 942      $a 234 $9 2
17 943      $a 198x $9 1
18 943      $a 197x $9 1
19 944      $a cm $9 2
20 950 11   $a thomson, virgil $d 1896 $9 1
21 950 11   $a samuel, gerhard $9 1
```

Structure of place names

- First Level Names (Country or State)
 - Natural features
 - Countries, states, and provinces, etc.
 - Extraterrestrial bodies
 - Geographic regions
- Second Level Names
 - Cities and other places below the first level
- Third level Names
 - City sections
 - Neighborhoods
 - Highway interchanges
- Qualifiers

Variation in place names

Cologne (Germany) *LC*

Cologne (Allemagne) *BnF*

Köln *SWD*

Tokyo (Japan) *LC*

Tokyo (Japon) *BnF*

Tokio *SWD*

Kilimanjaro, Mount (Tanzania) *LC*

Kilimandjaro (Tanzanie ; volcan) *BnF*

Kilimandscharo *SWD*

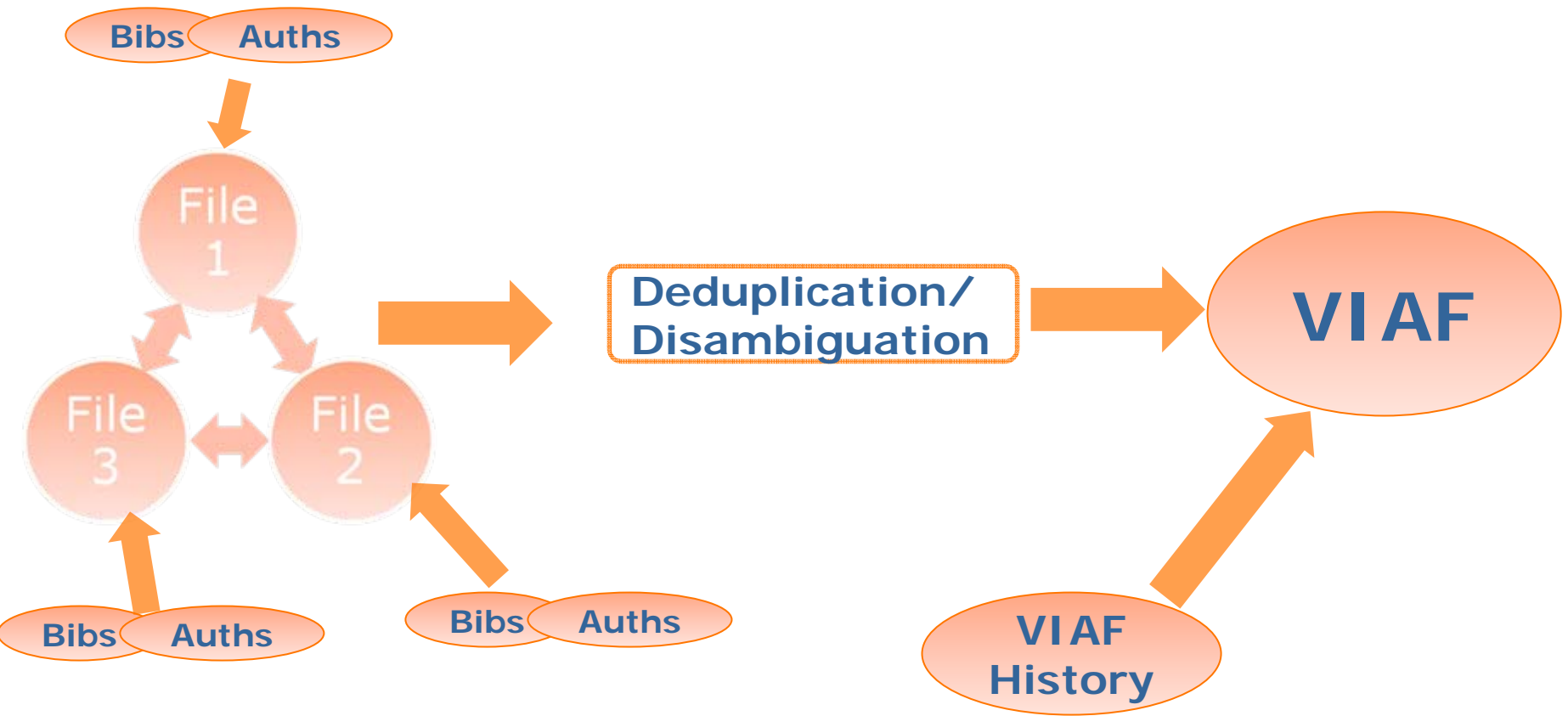
LC record for *Orense, Spain*

- 1 001 n 81110615
- 2 003 DLC
- 3 005 20060331052150.0
- 4 008 811109n| acannaabn |a ana
- 5 010 n 81110615
- 6 035 (OCoLC)oca00655657
- 7 040 DLC \$b eng \$c DLC \$d NIC \$d DLC \$d OCoLC
- 8 043 e-sp---
- 9 151 Orense (Spain : Province)
- 10 451 Ourense (Spain : Province)
- 11 451 nnaa \$a Orense, Spain (Province)
- 12 667 Old catalog heading: Orense, Spain (Province)
- 13 670 Mapa provincial, Ourense, 1:200.000 ... c1997
- 14 670 Columbia gaz. \$b (Orense, province, NW Spain, in Galicia; [capital] Orense)
- 15 675 GONet Aug. 21, 2000
- 16 781 0 Spain\$z Orense (Province)

FAST authority

000 cz n
001 fst01340107
003 OCoLC
005 20090417144539.0
008 060620nn anznnbabn || ana d
034 \$d W1341434 \$e W1341434 \$f N0564002 \$g N0564002 \$2 gnis
040 OCoLC \$b eng \$c OCoLC \$f fast
043 p
151 Pacific Ocean \$z Rowan Bay
550 Bays
670 GNIS, Feb. 10, 2004 \$b (Rowan Bay; bay; 7 mi. N of
Tebenkof Bay, on W coast of Kuiu I., Alex. Arch.;
Wrangell-Petersburg Census Area, Alaska; 56°40'02"N,
134°14'34"W; another Rowan Bay, pop. Place in
Wrangell-Petersburg Census Area)
670 GNIS \$b bay;56°40'02"N 134°14'34"W
688 LC (2008) Subject Usage: 1
688 WC (2008) Subject Usage: 3
751 0 Rowan Bay (Alaska : Bay)\$0 (DLC)sh2004005090

VIAF data flow



Current state

- Personal name files from 6 libraries (9 files) loaded
- Names are clustered
 - 10 million names
 - 8.5 million clusters
- Identifiers assigned:
 - <http://viaf.org/viaf/77390479>
- Preliminary work done on geographic names
- Unicode throughout
- UNIMARC and MARC-21 supported

What makes a match?

1,705,555 Title
846,722 Double date
123,487 Joint author
71,851 LCCN
24,587 Partial date and partial title
11,010 Partial date and publisher
9,179 Partial title and publisher
6,415 Name as subject
3,168 Standard number

Next steps for VIAF

- More participants (continuing)
- More name types
 - Geographics (started)
 - Corporates
 - Imaginary characters
 - Families
 - Titles
 - ... everything but topicals
- Linked data
- Better searching
- Rights agencies

Discussion

- How would you use VIAF?
- How important is VIAF?
- How important for RLG Partners?
- How important for OCLC Research?
- How important for OCLC?
- How important to be integrated with cataloging?
- How important to be integrated with WorldCat?

Interesting records

Uriel Simon

Leonid Sobinov

Galilah Ron-Feder-Amit

Claudette Colbert

Elizabeth George Speare

VIAF example

400's: Alternate Name Forms (11)

- ‡a Hurban Vajanský, Svetozár 
- ‡a Gurban-Vai`a`nskii, Svetozar, ‡d 1847-1916 
- ‡a Gurban-Vai`a`nskii, Svetozar ‡d 1847-1916 
- ‡a Gurban-Vajanskij, Svetozar ‡d 1847-1916 
- ‡a Hurban Vajanský, Svetozár ‡d 1847-1916 
- ‡a Vai`a`nskii, Svetozar Gurban-, ‡d 1847-1916 
- ‡a Vai`a`nskii, Svetozar Gurban- ‡d 1847-1916 
- ‡a Vajanský, S. H. ‡d 1847-1916 
- ‡a Vajanský, S. H. ‡q (Svetozár Hurban), ‡d 1847-1916 
- ‡a Vajanský, Svetozár Hurban, ‡d 1847-1916 
- ‡a Vajanský, Svetozár Hurban- ‡d 1847-1916 

VIAF

Beta

Virtual International Authority

Search: All fields

Search

47 headings found for *tillett*

	Name	Sample Title
1	Allison, Graham T., 1940-   Allison Graham Tillett 1940-....  Allison, Graham T. 	essence of decision explaining the cuban missile crisis
2	Schneer, Jonathan 1948-  Schneer, Jonathan  	social conflict and the political order in modern britain
3	Tillett, Barbara B. 1946-  Tillett, Barbara B.  	ifla cataloguing principles steps towards an international cataloguing code report from the 1st ifla ...
4	Tillett, Robert B. 	wurtland greenup county kentucky population 1 288
5	Strange, John Stephen  Strange, John Stephen, b. 1896 	reasonable doubt
6	Tillett, Benjamin  Tillett, Ben, 1860-1943 	who was responsible for the war and why
7	Tillett, Michael  	new year

[100 †a Tillett, Barbara B.](#)

[100 †a Tillett, Barbara B. †d 1946-](#)

[100 †a Tillet](#)

MARC-21

Leader	00000nz##a22.....n..4500
001	77390479
005	2009-03-03 17:03:27
008	880923n a nnaa ##### #aaa#### d
400	1 †aTillett, Barbara Ann Barnett †d1946- †2 DNB
400	1 †aTillett, Barbara †d1946- †2 DNB
700	1 †aTillett, Barbara B. †2
700	1 †aTillett, Barbara B. †d1946- †2
700	1 †aTillett, Barbara B. †2
901	†a 3598242751 †93 †2 SELIBR †2 DNB

400's: Alternate

Selected Titles

Countries of Pu

Publication Stat

Selected Publis

MARC-21

UNIMARC

History of VIAF ID:77390479 (3)