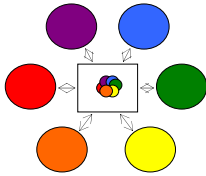# VIAF:
Virtual International Authority File

by Dr. Barbara B. Tillett
Chief, Cataloging Policy & Support Office
Library of Congress
for Tennessee Library Association
April 5, 2006

Agenda

- Conceptual models
  - FRBR, FRAD, FRSAR
- Authority control
  - VIAF
- Cataloging principles
  - IME ICC Statement
- New cataloging codes
  - RDA

Now onward to authority control and the VIAF project…

Objectives – VIAF

- Facilitate sharing
- Reduce cataloguing costs
- Simplify creation and maintenance of authority records internationally
- Enable users to access information in the language, script, form they prefer

The concept of a virtual international authority file has been discussed since the 1970's within IFLA – first as part of an idea for a single shared file and more recently in terms of linked national or regional authority files. The virtual international authority file within IFLA is intended to meet these objectives:

-to facilitate the sharing of the workload to reduce cataloguing costs. Our community has expanded, especially in Europe these days, where libraries are viewed with archives, museums, and rights management agencies as "memory institutions." We hope authority files could be freely shared among all communities. Shared authority information has the added benefit of reducing the global costs of doing authority work while enabling controlled access and better precision of searching.

Other objectives for authority control are

- to simplify the creation and maintenance of authority records internationally and

- to enable users to access information in the language, scripts, and form they prefer or that their local library provides for them...

Authority control benefits

- "Precision" in searching
- Syndetic structure of references to help navigate (the variant forms of name/title/subject/etc.)
- Displays to collocate works
- Links to forms used in particular resources
- Bring library catalogues into the mix of tools available on the Web

The benefits or virtues of authority control have been debated and restated for decades. When we apply authority control, we are reminded how it brings precision to searches, how the syndetic structure of references enables navigation and provides explanations for variations and inconsistencies, how the controlled forms of names and titles and subjects help collocate works in displays, how we can actually link to the authorized forms of names, titles, and subject that are used in various tools, like directories, biographies, abstracting and indexing services, and so on…We can use the linking capability to include library catalogues in the mix of various tools that are available on the Web.

Controlling forms used for access and displays provides consistency for users.

Virtual International Authority File (VIAF)
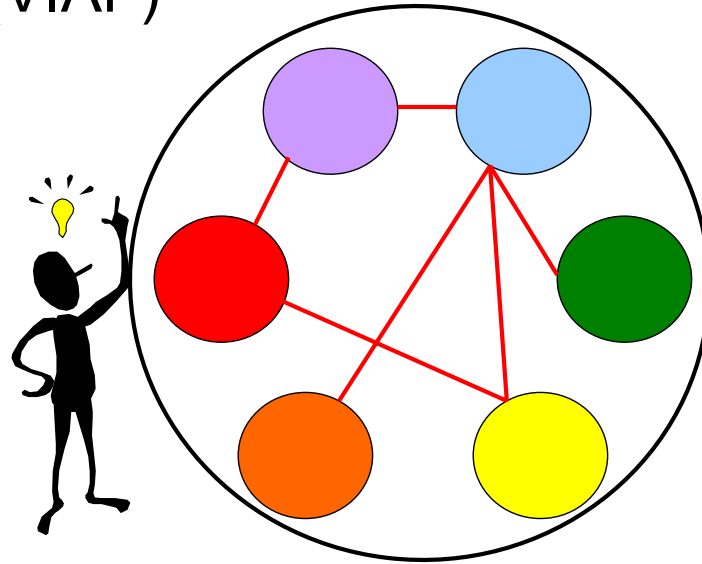
There are many technological capabilities that are coming together now and we are really at the brink of making a virtual international authority file a reality...

This is virtual because it is not really a file itself, but a linked system that connects existing Authority Files.

IFLA UBC authority principles

- Each country responsible for authority headings for its own personal and corporate authors
- National authority records available for everyone to use

- Same form and structure used worldwide

We're also making an historic change to how we view Universal Bibliographic Control (UBC). The IFLA UBC principles for authority control are parallel to those for bibliographic control, namely that

- each country is responsible for the authorized headings for its own personal and corporate authors and

- the authority records created by each national bibliographic agency would be available to all other countries needing authority records for those same authors. Even more, that the same headings would be used worldwide.

In the 1960's and 1970's when this was really catching on, technology had not yet advanced to make such sharing practical on an international level. Plus the lack of funding for an international center to manage such a program prevented that visionary concept from becoming reality. As for the same form being acceptable worldwide, the IFLA developers at that time were primarily from North America and Europe and apparently did not acknowledge the necessity for multiple scripts.

New view of UBC

- **User perspective**
  - ☐ Display script/language(s) of one's own country
- **National bibliographic agency responsibility**
  - ☐ Bibliographic and authority control in own country; available to everyone
- **International record sharing**
  - ☐ Link forms established in "national" or "regional" authority files to create a virtual international authority file

For the past few years a new view of Universal Bibliographic Control is emerging from several working groups within IFLA. This new perspective reinforces the importance of authority control, yet puts the user first…It's a practical approach that recognizes a user in China may not want to see the heading for Confucius in a Latinized form, but in their own script. Similarly users in Russia or the Ukraine would want to see the heading in your own script and language.

Yet to still get the benefits of shared authority work and creation of bibliographic records that can be re-used worldwide, we can link authorized forms of names, titles, and even subjects through the authority files of national bibliographic agencies and other regional agencies to create a virtual international authority file. These are several models for how this might work and we need to do more pilot projects of prototypes of thee models to test which would be best to pursue.
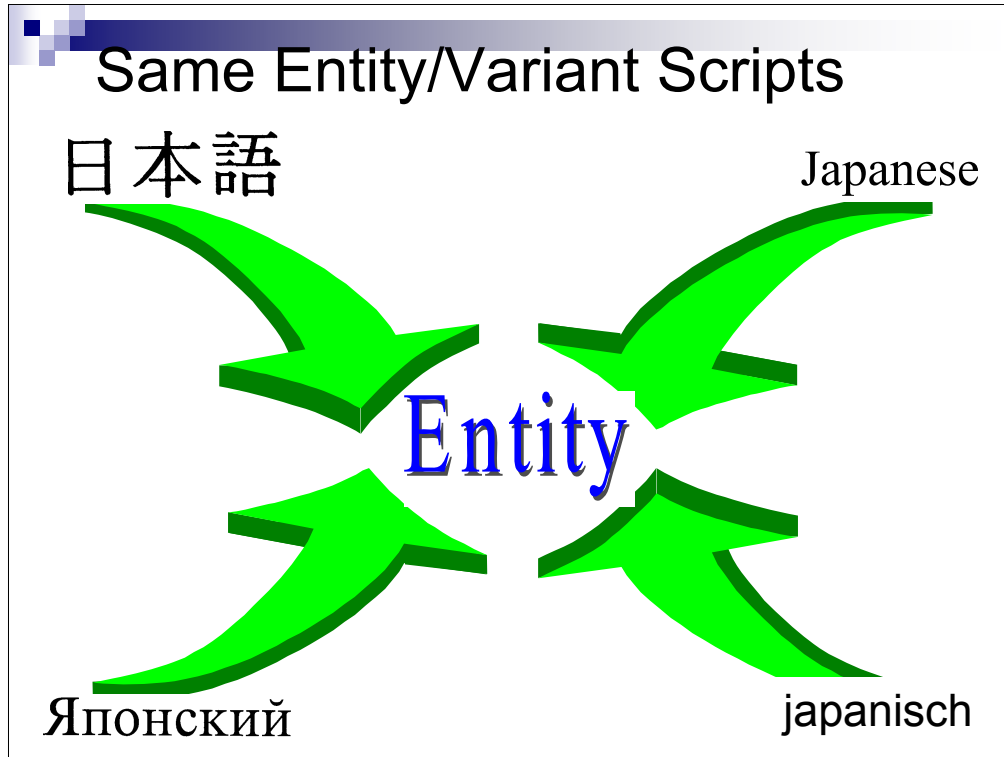
In order to be of most use to the library users in each country, the scripts should be the scripts they can read!

This slide shows that the names we give to an entity can be expressed in many languages and in many scripts. For example, we could write it in English or German with a roman script, in Russian in Cyrillic scripts, or in Japanese (in any of three scripts!) and in many other languages and scripts.

Transliteration may serve as a way for some users to be able to decipher records, but much better is the accuracy of using original scripts.

We should now provide at least cross references for variant forms of headings in variant scripts when that is appropriate. We should eventually be able to display the script and form of a heading that the user expects and wants.

I believe that many catalogers within IFLA realize the value of preserving parallel authority records for the same entity. This allows us to reflect the national and cultural needs of our individual users, and at the same time to allow us to set up the syndetic structure of cross references and authorized forms of headings to be used in our catalogues intended for a specific audience following specific rules. It also allows us to include variants in alternate scripts, at least as cross references for now.

In working internationally we quickly see the challenges to sharing authority information.

Entities

- Different entities – same name

Winston Churchill

Ideally a particular entity will have the same established form in all authority files. But we know this isn't always true. Different entities may be assigned the same established form of name.

Winston Churchill is the name of the famous English prime minister, an English nobleman from the 1600's, a grand master engraver, even the name of an orchid and a fuschia flower – among many others of the same name.

Entities

- Different forms of names for same entity

    Barbara Ann Barnett

    Babs Barnett

    Barbara Ann Barnett Tillett

    Babs

    Dr. Tillett

    Barbara B. Tillett

    Tillett, Barbara B.

Also different forms of the name may be established for the same entity.

Bibliographic identities

- 李滉
  - □ Yi Hwang (McCune-Resichauer)
  - □ I Hwang (Revised romanization)
- 퇴계; 退溪
  - □ T'oegye (Pen name)
- 경호; 景浩
  - □ Gyeongho (Courtesy name)

- 이황 Hangul
- 李滉 Hanja

There is also the concept of bibliographic identities that is intended to differentiate between the different persona that are used by a person or group of people in different circumstances. This could be a pseudonym versus a real name or an official name or a courtesy name, etc., such as we see here for Yi Hwang who wrote under the name T'oegye but also had the courtesy name of Gyeongho. The name can be expressed in several scripts, including several transliterated forms depending on the standard used for romanization.

If all forms are possible…

- Need to explain variants to end-user to avoid confusion (explain pseudonyms, abbreviations, etc.)

Lewis Carroll

Dodgson, Charles Lutwidge, 1832-1898

When we control all the possible forms variations for the names of an entity and associate then with the bibliographic records for the bibliographic resources that they have some role in creating, producing, or owning, we need to explain that to the user. For example, if the user chooses one form, and we then launch a search, as the Getty Museum and Institute does, that includes all the variant forms and retrieves them all, the user sees records with the variety of names and may find this puzzling.

I searched in the Getty system under Lewis Carroll, the author of "Alice's Adventures in Wonderland." Why am I getting back information about this mathematician, Charles Lutwidge Dodgson, 1832-1898?

You need to tell the user of variant names used for the same person or corporate body or work.

Alternatively it might be better to recognize each "persona" and establish a name for each one with its variant forms and link them with see also references.

Entities

- **Crosswalks and Mappings - not always 1-1**
  - **Differences among cataloguing rules**
    - **Ships**
    - **Events**
    - **Meetings of corporate bodies**
    - **Undifferentiated names**
  - **Differences among communities**
    - **Libraries, archives, museums, etc.**

As we look at linking we must recognize that different cataloguing rules have differences in what they consider entities - choices are not universal, for example, German rules (*Regeln für die alphabetische Katalogisierung*- RAK) do not recognize that the ships logs can be under an entry for the name of the ship as they can be in AACR2, so the Germans would not have an authority record for ships' names. Similarly for events. For meetings of corporate bodies, the German rules would not create a heading for the entity that AACR2 creates in as a hierarchically subordinate heading for a meeting under the name of the corporate body.

There are also different practices for undifferentiated names - the Germans recently changed their rules to differentiate more names - they more commonly used undifferentiated forms for personal names using just initials for forenames. They still do not require as complete a name or a qualified name to distinguish as the Anglo-American Cataloguing Rules call for. However, even under the same cataloguing rules, say AACR2, when we get more information to differentiate a person, we can make a new authority record to differentiate that person from others groups together under an undifferentiated form of name. This also means that the record for the undifferentiated name can reflect different associated entities over time. And different communities recognize different entities…

Projects to facilitate authority control on a global scale

- EU: AUTHOR Project, LEAF, <indecs>, INTERPARTY
- HKCAN
- IFLA: MLAR, FRANAR, FRSAR
- Dublin Core "Agents"
- DELOS/NSF Working Group "Actors/Roles"
- EAC (Encoded Archival Context)
- Unicode/Multiple Scripts
- NACO/SACO for AACR2 and LSCH Users

Over the past few years there have been several projects that help us get closer to providing authority control on a global scale: In my paper for the LC Bicentennial Conference in November 2000, I described the projects you see on this slide:

*There are several that were sponsored by the European Union, such as the European Project AUTHOR (started in 1995) that converted a sampling of authority records from the 5 participating countries (Belgium, Spain, Portugal, UK, France) to the same communication format, UNIMARC. The LEAF project (now coordinated at the Staatsbibliothek zu Berlin) looked at linking authority files for archival purposes using Z39.50 protocols and OAI protocols. The <indecs> and INTERPARTY projects were looking for cooperative work among libraries, museums, archives, and rights management communities in sharing authority information.

*Within the International Federation of Library Associations and Institutions, the IFLA MLAR (Minimal Level Authority Records) Working Group identified essential data elements needed in authority records (today we'd call these metadata). This work continues by the IFLA Working Group on FRANAR (Functional Requirements for Authority Numbers and Records and FRSAR for subject authority records.

*Within the digital metadata community, there was a Dublin Core "Agents" working group that explored recommendations for dealing with authority information in the digital environment.

*Another development over the past few years has been the acceptance of Unicode within the Microsoft tools, such as Windows that facilitates more global compatibility with multi-scripting .

*There is also the initiative of expanding the international cooperative cataloging projects, NACO and SACO, to worldwide users of the *Anglo-American Cataloguing Rules* and *Library of Congress Subject Headings* - also promoting authority control on a global scale.

Russian National Authority Control Projects

May 2003 International Conference on Authority Control (Russian State Library, Moscow)

- **Personal Names**
  - ☐ National Library of Russia (Russian authors)
  - ☐ Russian State Library (foreign authors whose works have been translated into Russian)
- **Corporate Body Names**
  - ☐ National Library of Russia(1930-)
- **Geographic Names "RuGeo"**
  - ☐ Russian State Library (Ministry of Culture)

There are several initiatives in Russia to build national authority files. Some identified during the May 2003 International Conference on Authority Control held at the Russian State Library in Moscow, include projects for personal names (the National Library of Russia in St. Petersburg taking the lead for names of Russian authors and the Russian State Library taking the lead for foreign authors whose works have been translated into Russian. For corporate body names, the National Library of Russia will take the lead for names since 1930. For geographic names, the project RuGeo is led by the Russian State Library with funds from the Russian Ministry of Culture.

Through such projects we are able to reduce the costs of authority work while improving the potential for more precise searching capabilities for our users.

Systems to help future authority work

- Automatic check of heading against local authority file
- If not found, automatic check against "virtual" international authority file
  - Display found matches for editing or reference
 - Insert authorized forms into local authority record for future link

The cost of authority work is a concern, and international sharing is just one piece of the solution to help reduce costs. Another is to develop better automated system capabilities to support catalogers doing authority work and to even do some of it automatically. Some local systems already provide us with computer-assisted mechanisms for automatic checking of headings against an existing authority file, and we could see this expanded to start a search against a virtual international authority file, if no match was found locally. . Here's where having an international resource will help global costs.

We can also envision the capability of displaying the found matches from the virtual file for a cataloguer to edit or to merge information, if desired, into the local authority record.

Switching for displays

- Library default
- User-selected preferences
  - Client set-up, "cookies," or future method
    - Language
    - Script
    - Culture (country)/spelling
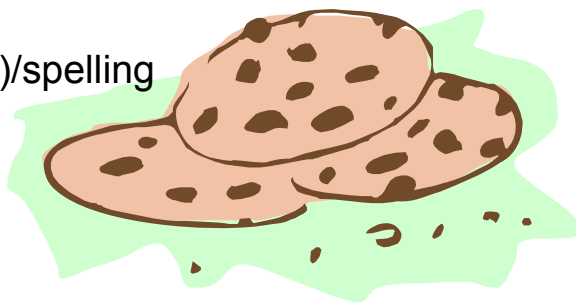
Some systems now provide community specific retrievals to concentrate on the subject needs of a community in selecting resources for online searches, and other systems like "my library" or "my OPAC" even go beyond that to individualize specific retrievals. Those could build in the authority preferences for user preferred scripts and displays for controlled vocabularies.

We want to have the authorized form preferred by a library as the default offered to most users, but we can also envision offering user-selected preferences through client software, or cookies that let the user specify once what their preferred language, script, or cultural preference is - for example for spelling preferences when cultures have variations, like American and British variations of English.

Scenario - Web access

- Original cataloguing with no match found in local authority file
- Launches Web search of virtual international authority file
  - Assumes response times/system reliability will improve in future

There are many ways this could actually be applied, and I've suggested several scenarios in earlier papers. Let's quickly take a look at one scenario.

Let's say the catalogers is doing original cataloging and finds there is no authority record in the local file for a heading he or she needs.

Create bibliographic record

100 1 ‡a Tchaikovsky, Peter I

245   ‡a Piano concerto

Not found in local authority file. Checking VIAF.

A cataloger types in information in the bibliographic record for a heading. The local system checks the local authority file and finds no match, so it tells the cataloguer that the heading was not found and launches a Web search to the virtual international authority file.

RUSMARC-record
Маркер: 00445nx 22001453 450
001:   10326
005: 20001108144619.0
100:   $a20001108arusy0103   ca
200:   1$8rus$7ca$aЧайковский$bП. И.$f1840- 1893$gПетр Ильич$4070
200:   1$8rus$7ba$aChajkovskij$bP. I.$f1840-1893$gPetr Il'ich$4070
300: 0 $aРусский композитор, ученик А.Г. Рубинштейна.
700:   1$310327$8rus$7ba$4070
810:   $aВсемирный биографический энциклопедический словарь. - М., 1998.
801:   0$arus$brnb$c20001108
810:   $aГАК РНБ.

Up pops the match with a record created at the National Library of Russia in St. Petersburg…. [Aside: The remarkable thing is I did indeed find this record using the Internet and was able to display the roman and Cyrillic characters on my PC!]

Our cataloguer takes a look and perhaps doesn't want all the information but likes a reference or two and wants a link, so

Scenario
Link authorities

100 1 ‡a Tchaikovsky, Peter I

Create and Link?

Yes

No

The local system asks the cataloger if she wants the system to create a basic authority record from the one found and to make a link to it…and we click on "yes".

And the local system would automatically build a local authority record, grabbing the linking information from the virtual authority file - that is the record from St. Petersburg, Russia. The cataloger then adds the MARC field 100, authorized form, according to the locally used cataloging rules, in this case AACR2. And our cataloger can add other fields if needed.

The local system would add the linking 700 field - the MARC format has the 7xx fields in authority records, where we can put the linking authorized form and the record control number and the source information for future linking. This linking of authority files would primarily be among the national or regional authority files of national bibliographic agencies - depending on the model we choose. I'll come back to that in a minute.

Virtual International Authority File (VIAF)

Tchaikovsky
(LC)n79072979

Чайковский
(NLR)10326
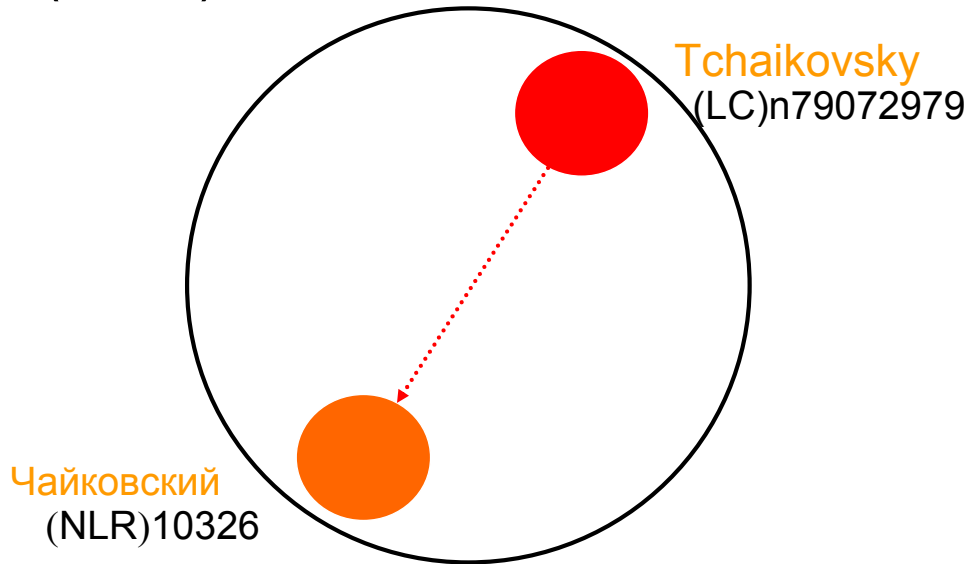
So we've now added another link in the virtual international authority file to the authorized form following AACR2 - note the record control number for the Library of Congress: (LC) n79072979 - and the Russian record for the same entity following the Russian cataloguing rules in Cyrillic script - note the record control number from the National Library of Russia: (NLR)10326.
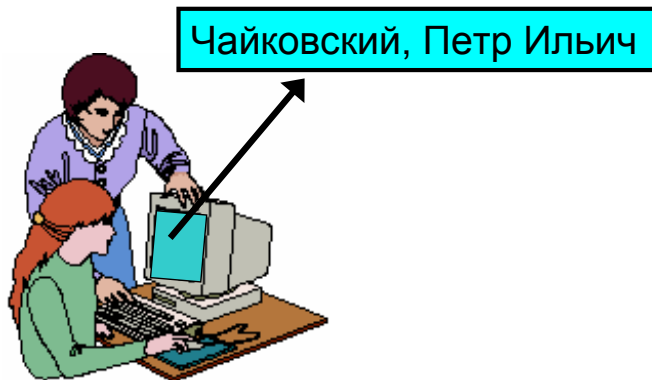
Scenario
Corrects bibliographic record

100 1 ‡a Tchaikovsky, Peter Ilich, ‡ d 1840-1893

245   ‡a Piano concerto

Then our local system updates our local bibliographic record using information from the authority record.

Local system - User view

User's cookie specifies Cyrillic/Russian preferred.

Чайковский, Петр Ильич

When a user comes along, the local system or the "cookies" on the user's system, could specify they want to see the Cyrillic form (click) and (click) we could display it for them…

You can also imagine displaying any script or a Braille keyboard output, or we could provide voice recognition response, built on a user's profile or their "cookie."

Tag	I1	I2	Subfield Data
010			‡a n 80050515
035			‡a (DLC)n 80050515
040			‡a DLC ‡c DLC ‡d DLC ‡d NIC
100	0		‡a Confucius
400	0		‡a Konfuzius
400	0		‡a K'ung Fu-tzu
400	0		‡a Kongzi
400	1		‡a Kong, Qiu
400	0		‡a K'ung-tzu
400	1		‡a K'ung, Ch'iu
400	0		‡a Kʻoshi
400	0		‡a Konfut͡siĭ
400	0		‡a Kongja
400	0		‡a Kung Fu
400	1		‡a K'ung, Fu-tzu
400	0		‡a Confucio
400	0		‡a Конфуций
400	0		‡a 孔夫子
400	0		‡a 孔子
400	0		‡a 孔丘
400	0		‡a こうし
400	0		‡a コウシ
400	0		‡a 공자
670			‡a Jakobs, P. M. Kritik an Lin Piao und Konfuzius, c1983: ‡b t.p. (Konfuzius)
670			‡a Konfut͡siĭ, 1993: ‡b t.p. verso (551-479 B.C.)
670			‡a His Gespräche (Lun yü), 1910: ‡b t.p. (Kungfutse)
670			‡a Web connection ‡u http://www.friesian.com/confuci.htm
700	0		‡a 孔夫子 ‡5 Natl. Lib. of China

This is not a VIAF record, but rather is an example of what a Library of Congress authority record might look like with Unicode capability to include original scripts as cross references in a library's catalog. This just gives you an idea of what it might be like. The Library of Congress will be implementing a Unicode-based version of our integrated library system in about a year from now – we hope.

There is no particular order to the arrangement of the references, except to place the non-roman scripts following the roman scripts. This model shows English, German, Italian, Chinese, Japanese, Korean, Russian, and transliterations (including Wade-Giles and pinyin for the Chinese, since the Library of Congress just switched to use pinyin).

Distributed - independent AF's with Z39.50 access

SEARCH
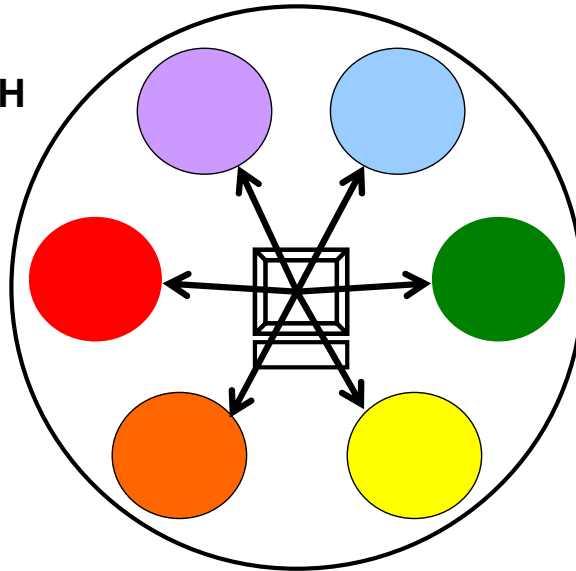
There are several models for a virtual international authority file.

For a distributed model, a searcher would use a standard protocol like the next generation of Z39.50, or SRU/SRW, to search the independent authority files of participating National Bibliographic Agencies or regional authorities.

Central Link
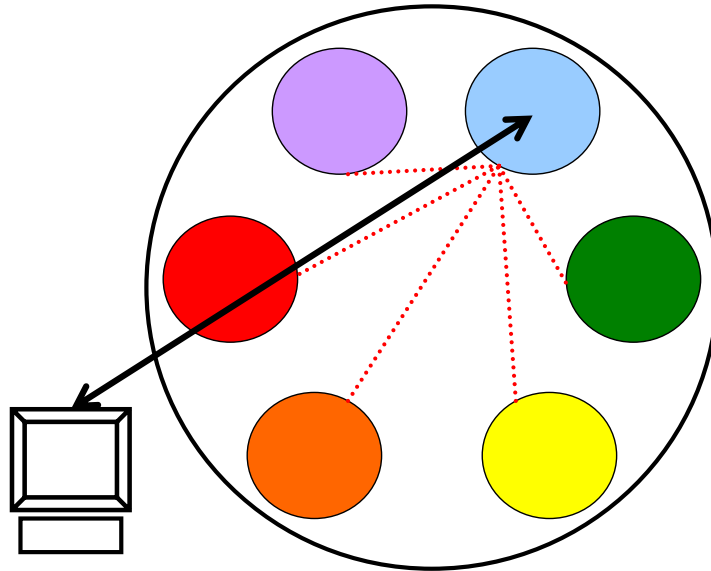
Another model is to have one central authority file and link all others to that, so that work would not need to be done by each national bibliographic agency with all other participants in this international universe. A cataloger would then get access to all the authority records for that entity worldwide by a single search of the central file. If there was not match in that central file, a search could then be made with Z39.50 to the other files.
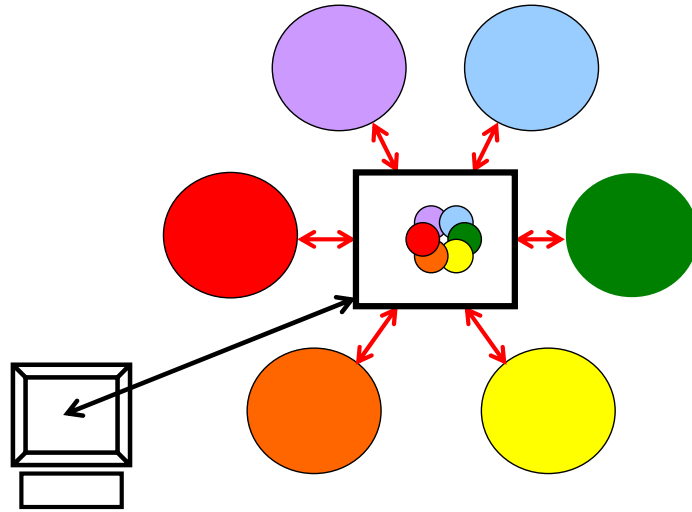
Centralized - Union Authority File (OAI Model)
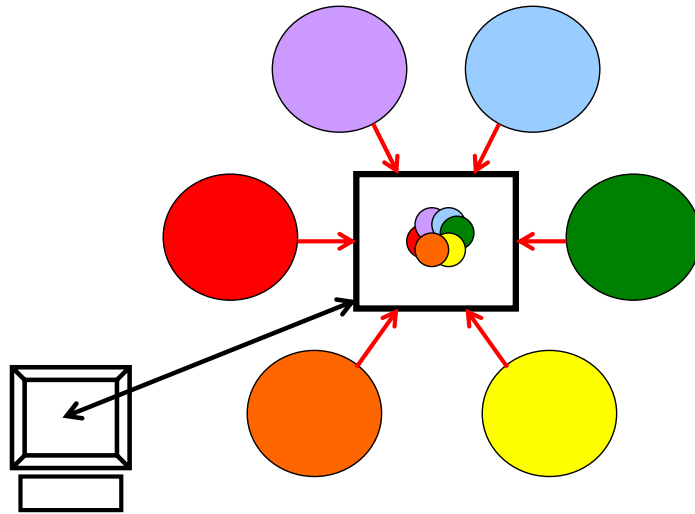
This shows a centralized model: We may find that this model is the best approach in terms of record maintenance - The Open Archives Initiative (OAI) protocol model uses a server with harvested metadata from the national authority files.

That information is refreshed in the server whenever there are changes in the national files. This means the day to day record maintenance activities continue to be managed as they are now by the National Bibliographic Agency (or regional authority). Unless we also build in the linking, we possibly will lose a level of precision in the searching in this model; but there are ways to include the links for entities in this model, too.

Virtual International Authority File (VIAF)

If we take that centralized union authority file model and instead of having all of the records stored centrally, we instead create a system with a server that knows about the existing authority files in each national bibliographic agency, we could in effect have a virtual international authority. The file itself exists through the linking and interconnections through the networked system.

I am sure you can imagine other variations of these models. And we need to try them out to see which will be best for us in today's Internet environment.

VIAF Proof of Concept

Project underway

- Die Deutsche Bibliothek (DDB)
- The Library of Congress (LC)
- OCLC Online Computer Library Center (OCLC)

A project is now underway to test a variation of a centralized virtual authority file model. In August 2003 the Deutsche Bibliothek, the Library of Congress, and OCLC signed a memorandum of understanding to jointly engage in a research project to test this VIAF concept – a proof of concept project. It is hoped if this proves successful, it can be the basis for a true Virtual International Authority File that links the world's national and regional authority files as a freely available shared resource.

VIAF Proof of Concept DDB/LC/OCLC

OCLC matching algorithms

LCNAF

DDB/PND

Stage 1 Retrospective Files Link
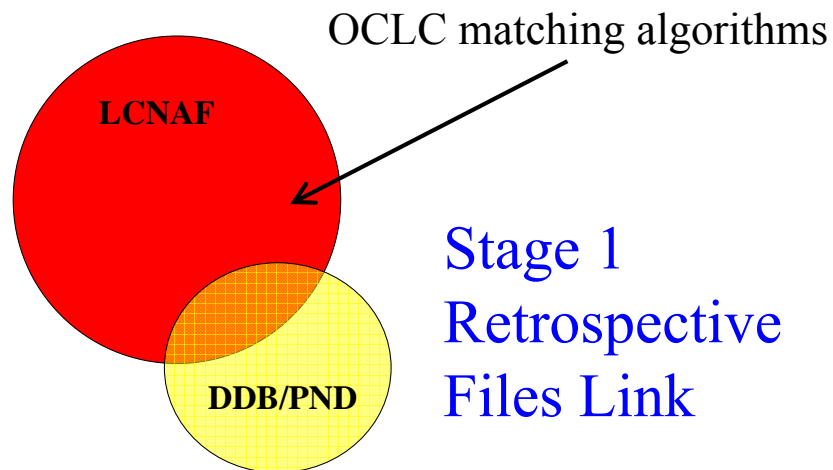
The first stage of this project linked our existing authority records for personal names: LCNAF and the DDB's Personal Name Authority File (PND).

It involved testing OCLC's matching algorithms to see how much a computer can do for us and how much will require human intervention for matching and checking. OCLC's Research Office used information in the bibliographic records (such as dates of publication, subjects, languages, countries of publication, etc.) and information in authority records (authorized and variant forms of names) to do this matching in order to increase the matching rate and to reduce false matches.

Matching and linking

- LC Personal Name Authority File
  - 4 million total authority records

- DDB Personal Name Authority File
 - 2.5 million total authority records

This first stage is nearly completed and the results are being compiled for a report by the end of this year. OCLC received the authority records and the bibliographic records from both the Library of Congress and the Deutsche Bibliothek and has done a first pass to compare and match the records.

This slide shows the total number of authority records that were in the respective files at the end of 2003 when the set of records to be matched was made available. Nearly 4 million records at LC and nearly 2½ million records at DDB.

There was about a 7.8% match of unique entity records, and 64,000 more name pairs were "rejected" because they matched multiple records (such as for undifferentiated names in one file matching two or more records in the alternate file. The error rate in the matching is about 1%.

We still need to determine whether all this matching is worth the cost and overhead in having both the authority and bibliographic records.

VIAF – LEAF Model (Test)

LCNAF

5.8%
Simple match
(LEAF model);
~10% false match

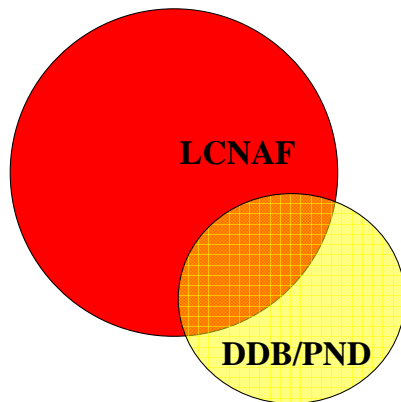Stage 1
Retrospective
Files Link

DDB/PND
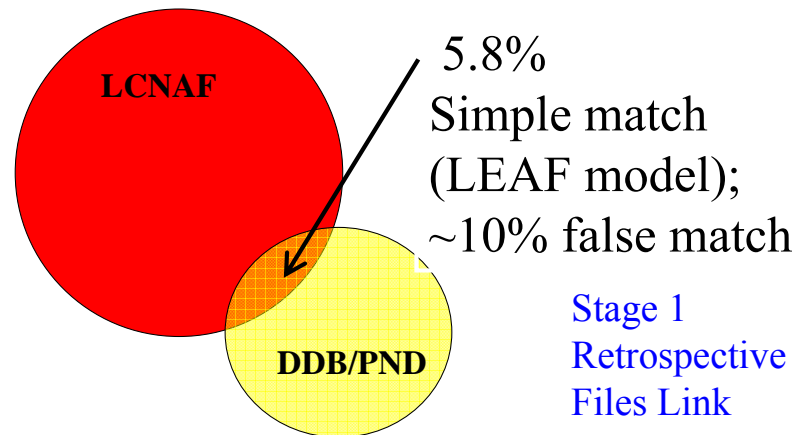
To compare the VIAF Proof of concept model with other models that have been proposed, the matching algorithms and techniques were compared on the same set of personal name authority records from the Library of Congress and the Deutsche Bibliothek.

One such project is Project LEAF – Linking and Enhancing Authority Files – an EU-funded project that is now somewhat operational at the Staatsbibliothek zu Berlin in Berlin, Germany. It began to test linking authority files among archives and libraries, using a very simple matching of the authorized headings.

As you can see on this slide, this simple matching only matched about 5.8% of the records in the two personal name authority files but also led to about a 10% false matching of names for different entities.

Other VIAF Models

LCNAF

DDB/PND

* InterParty
* Project AUTHOR

Stage 1
Retrospective
Files Link

Another model that was proposed and described in 2004 was the InterParty Project, which was an EU-funded demonstration project to link authority files among diverse organizations – libraries, archives, museums, rights management organizations, publishers, etc. Unfortunately, that project did not specify any technique or data requirements necessary to support a matching algorithm.

Yet another, earlier project to test sharing authority information was the EU-funded Project AUTHOR, which took samples of authority records from seven national libraries in Europe to convert them to UNIMARC format and make them all searchable at the same time. There was no attempt to match the records for the same entity, other than to display records from the various files that matched the text string keyed in by the user.

# Enhancing the Authorities
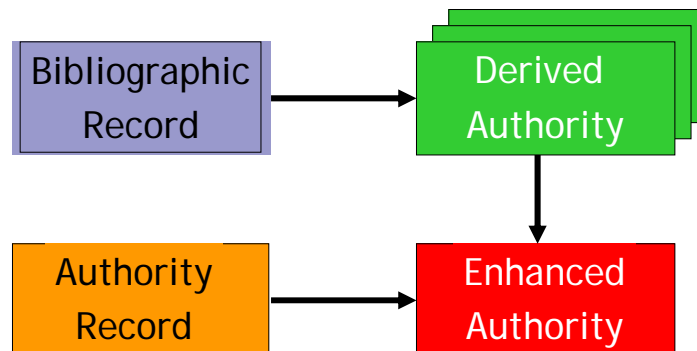
During the matching process, both the authority records and bibliographic records are examined. For every personal name used in either as a main entry or added entry in the bibliographic records, a derived authority record was created. In addition to the name, the derived authority record includes a coded summary of the material published based on information from the associated bibliographic records. A bibliographic record with multiple personal names will generate multiple derived authority records.

All of the derived authority records for a particular person will be clustered with the authority record for the individual.

The contents of the all the derived authority records for the individual are added to form the enhanced authority record. OCLC will use the enhanced authority records as the database for the VIAF.

VIAF Proof of Concept DDB/LC/OCLC

LCNAF

VIAF authority

DDB/PND

Stage 2 Build searchable database
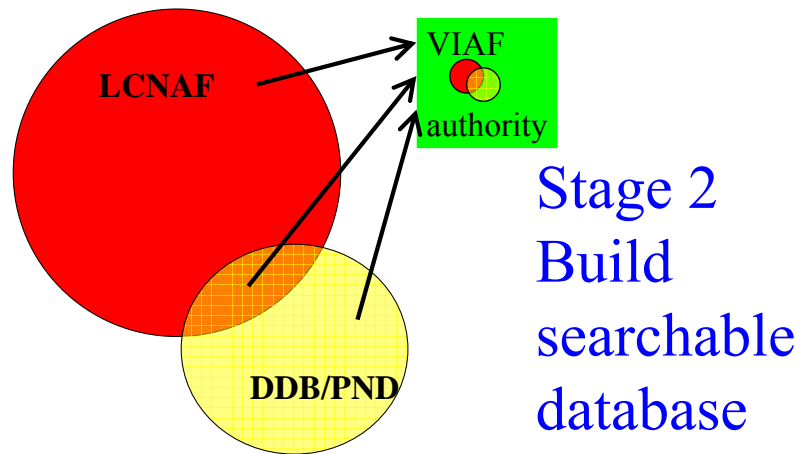
In Stage 2 of the project, we are building this searchable database. OCLC is using the open source software, Site Search to provide access to this database as this part of the test. As we make the links we will be building one or more servers with this "metadata" – one will be housed at OCLC, probably another at the OCLC European office (PICA) and another at the DDB. We were not planning on having one at LC for this project.

VIAF Proof of Concept DDB/LC/OCLC

LCNAF

OAI
Server(s)

DDB/PND
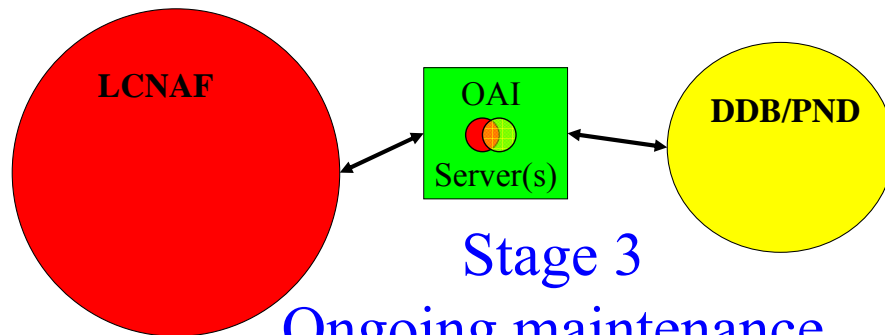
Stage 3
Ongoing maintenance
- Harvest metadata
(OAI protocols)

In Stage 3, we want to test the concept of using the Open Archive Initiative (OAI) protocols to do the ongoing maintenance of updating the information in the server by harvesting metadata for new and updated or deleted information in the home authority files.

VIAF Proof of Concept DDB/LC/OCLC

LCNAF

OAI

Server(s)

DDB/PND

공자

Stage 4
End user interface
- Unicode display switch
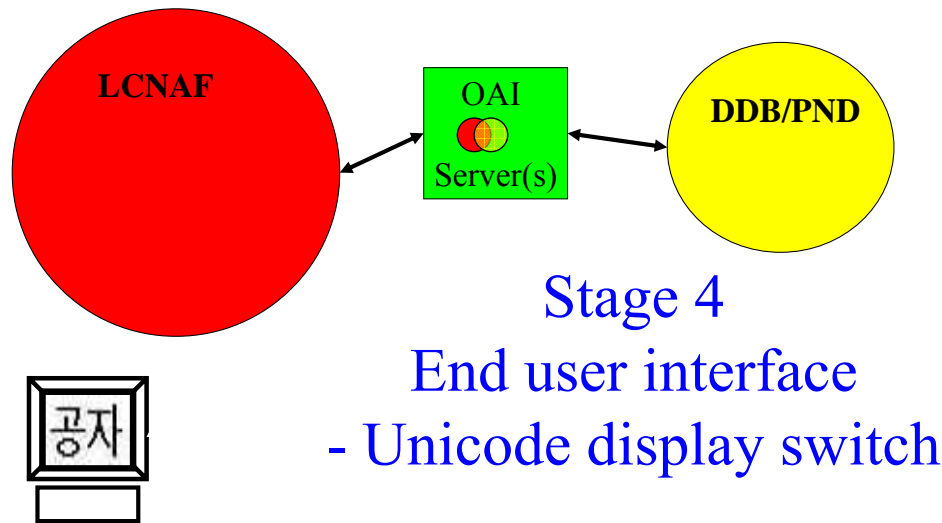
A possible last stage (Stage 4) to test in the proof of concept would be the end user display capabilities to switch the user's preferred form of language and script that would be displayed on his or her machine.

At this time we believe this stage may be omitted from the project, as the technology is not quite yet ready.
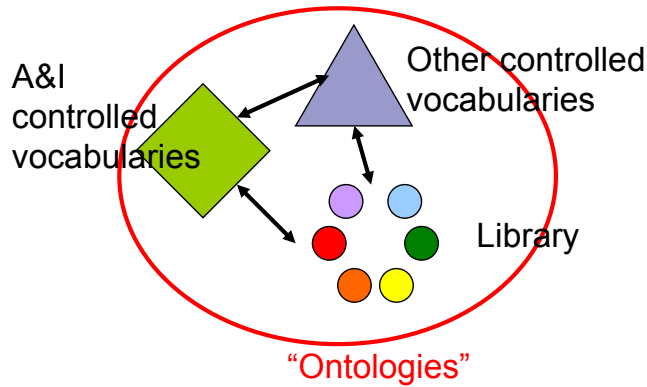
Virtual International Authority File (VIAF)
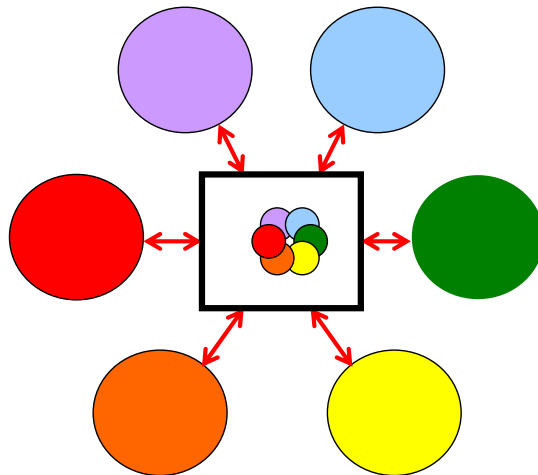
We prefer this model for now, as it seems to hold the most promise for maintenance, but there are concerns about scalability, given the overhead of matching with bibliographic records. However, if it proves successful, we would like to include connections to all the major authority files worldwide, including those from other communities like abstracting and indexing services, archives, museums, publishers, etc. We especially look forward to testing this strategy using non-Latin script authority information. We also want to test it on other types of authority records: such as for Corporate names, Geographic names, and Uniform titles. Subjects might seem appropriate, but we know that abstract concepts often don't translate uniquely from language to language. Therefore, it is unlikely that VIAF model would be extended to include subjects.

We really hope we can preserve local forms of names this way and link different records that use varying cataloguing codes and yet still meet users needs.

Semantic Web Building Blocks

A&I controlled vocabularies — Other controlled vocabularies — Library — "Ontologies"

For the future, we can envision a shared international authority file being an integral part of a future "Semantic Web." You may have heard about this in a Scientific American article a few years ago now by Tim Berners-Lee, founder of the Internet. The idea is to make the Internet more intelligent for machine navigation rather than human navigation of the Web. It involves creating an infrastructure of linked resources and the use of controlled vocabularies, they are calling "ontologies." These ontologies could be used to enable displays in the user's own language and script.

Here's where libraries have an opportunity to contribute to the infrastructure of the future Web - we already have controlled vocabularies in our various authority files. Those would be linked with other controlled vocabularies of abstracting and indexing services, of biographical dictionaries, of telephone directories, and many other reference tools and resources to help users navigate and to improve the precision of searches, so users could find what they're looking for.

Semantic Web Building Blocks

"Ontologies"

End-user
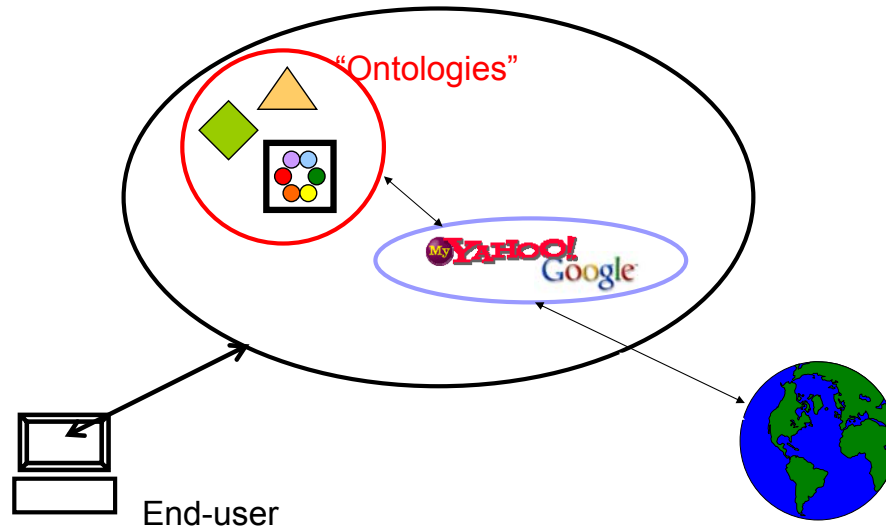
You can see that we would also build in the search engines and future tools that as a collective resource would connect us to the entire digital world.

All of this, of course, would have built-in, appropriate security and privacy assurances and ways to identify and acknowledge resources that we can trust and rely on, and somehow, miraculously, all the copyright issues will be resolved - we are definitely talking future!

But it's great to think about the possibilities and opportunities for testing this out and to think about how we can improve upon our dreams.

The Web has brought us a new way to convey information. The new twist is that our catalog - that is our PC where the online catalog is displayed, is also the device for viewing the actual digital objects and connecting to the entire digital world.

 A freely available Virtual International Authority File could be used by Web systems to improve the precision of users searches and help enable display of the language and script for names and subjects that a user prefers.

Questions?