Attachment 1
VIAF Update

Thomas Hickey
Chief Scientist
OCLC Research
Helsinki, 2012

Agreements signed

- Library of Congress
- Deutsche Nationalbibliothek (German National Library)
- Bibliothèque nationale de France (National Library of France)
- Kungliga biblioteket (National Library of Sweden)
- National Library of Australia
- Biblioteca Nacional de España (National Library of Spain)
- Biblioteca Nacional de Portugal (National Library of Portugal)
- Bilioteca del Congreso Nacional de Chile
- Biblioteca Apostolica Vaticana (Vatican Library)
- Library and Archives Canada (Bibliothèque et Archives Canada)
- Getty Research Institute
- NUKAT Center of Warsaw University Library
- RERO (Library Network of Western Switzerland) (RERO)
- Système Universitaire de Documentation (Sudoc/ABES)
- VLACC Brussels network of Flemish public libraries

Waiting for agreements
(former contributors)

- ICCU (Italy)
- National Library of the Czech Republic
- Swiss National Library
- BIBSYS (Norway)
- Országos Széchényi Könyvtár (Hungary)

Files in motion
(new contributors)

- KB (Netherlands)
- DBC (Denmark)
- NDL (Japan)
- Wikipedia
- IZUM (Slovenia)
- РSL (Russia)
- NII (Japan)
### Other possibilities
- National Library of Finland
- National Library Board of Singapore
- National Library of Malaysia
- National Library of Latvia
- Magistral Library and Archive of the Order of Malta (Italy)
- Ionian University (Greece)
- Universidad Central "Marta Abreu" de Las Villas (Cuba)

### Statistics
- Browser interface usage up 75%
- API usage up 400%
- Uniform Titles
- Added BIBSYS, SUDOC, VLACC
- 18 -> 29 million authority records
- 70 -> 112 million bibliographic records
- 14.5 -> 22 million clusters

### Other changes
- Integration with OCLC services started
- Relationship with ISNI
- ODC-By license, public bulk downloads
- Uniform Titles
- Clustering improvements
- xA in place to help control clustering
- Redaction procedures in place
- Better monitoring, limits on simultaneous use

### Coming year
- Integration with ISNI reports
- Better anomaly reports to participants
- Move to new infrastructure
  - Hadoop/Hbase
  - Production-level interface support
- More files
Proposals

• Bring Wikipedia in as full source file
  – Adding VIAF IDs to English Wikipedia
• Link to more outside sites
  – sameas.org
  – freebase.com
• Improve VIAF usability
• Scholars funnel
• Geographics?
• Deities, imaginary characters? ships?
Attachment 2
### Current state

- We do preliminary matching based on dates, titles, ISBNs
- No disambiguation, so we throw out questionable and multiple matches
- Does not participate in clustering
- ~ 250K Wikipedia links in VIAF
  - Project to get those VIAF links in Wikipedia
  - German Wikipedia drives significant traffic to VIAF
- Wikipedia links in VIAF are widely used

### Proposal

- Consider Wikipedia a ‘source’
- Only include Wikipedia names that match a VIAF entity
  - Avoids ‘random’ Wikipedia names
  - More consistent with library authority domain
- Only people for now

### Advantages

- Allows use of VIAF’s disambiguation stage
  - More matches
  - Better matches
  - Better clustering
- Improves one of the more visible aspects of VIAF
- Would allow inclusion of information from Wikipedia in VIAF
Questions

• Should we proceed with personal names?
• Only those already in VIAF?
• Interest in other types of names?
  – Corporates, geographic, books, movies
• Is starting with English Wikipedia reasonable?
  – More VIAF references in German Wikipedia
• License? MARC Codes?

Scholars Funnel in VIAF

Thomas Hickey
Chief Scientist
OCLC Research
Helsinki, 2012

The Idea

• We offer ‘cloud-based’ service to create/maintain authority files
  – Built using the xA platform
• Treat the combination as a ‘scholars’ funnel for VIAF
• Would show up as (a single) source in VIAF

Current state

• Project
  – Lead by Karen Smith-Yoshimura
  – Implemented/supported within Research
• Names
  – Syriac Reference Portal
  – Exploring Fihrist project in the UK
Syriac Reference Portal

- [www.syriac.ua.edu](http://www.syriac.ua.edu)
- Started at U of Alabama
- Lead by David Michelson (now at Vanderbilt)
- We loaded 3K names
  - Mined from WorldCat Identities
- Names are being actively edited

Fihrist --Islamic Manuscripts Catalogue Online

- [www.fihrist.org.uk](http://www.fihrist.org.uk)
- Out of the Oxford & Cambridge Islamic Manuscripts Catalog Online (OCIMCO)
- Encoded using TEI/XML
- A sample now in [viaf.org/hosted/sandbox](http://viaf.org/hosted/sandbox)
 Assertions

• Taps expertise of scholarly communities
  – Closer collaboration between scholars and librarians
  – Highlights importance of unique identifiers
• Scholars Funnel gives these names a permanent home
• Allows cross-project collaboration
  – While still maintaining association of names with projects
• Unconstrained by MARC-8 character set
• Would add valuable information about the names to VIAF

 Questions

• Seen as valuable?
• Interested in seeing information in VIAF?
• What sort of projects would be acceptable in such a funnel?
• License? MARC Codes?
VIAF Council Meeting Minutes
National Library of Finland, Helsinki, Finland
Friday, August 10, 2012

Attachment 3
VIAF - ISNI Interoperation
Discussion Paper

This paper is prepared in collaboration between the OCLC as ISNI Assignment agency, and the Bibliothèque nationale de France (BnF) and the British Library (BL) in charge of the quality control in the ISNI system. It focuses on the interoperation between VIAF and ISNI as strongly interrelated systems. It also outlines the issues, challenges and perspectives that the authority data quality presents in the ever evolving context of the networking environment and the role of the libraries in it.

Contents

1 Introduction ........................................................................................................................................ 2
  1.1 What is ISNI? ................................................................................................................................. 2
  1.2 ISNI and VIAF ..................................................................................................................................... 2
  1.3 ISNI Scope and ISNI Assignment ..................................................................................................... 3
2 How ISNI contributes to enhance VIAF quality .................................................................................. 4
  2.1 Notification of ISNI Assignment ....................................................................................................... 5
  2.2 Notification of ISNI Replacement ..................................................................................................... 5
  2.3 Notification of Duplicates .................................................................................................................. 6
  2.4 Notification of Data Error ................................................................................................................. 10
  2.5 Notification of a Record to Split ....................................................................................................... 13
  2.6 Special treatment for Pseudonyms+ ................................................................................................ 16
3 How VIAF contributes to ISNI ............................................................................................................... 17
  3.1 VIAF additions .................................................................................................................................. 17
  3.2 VIAF Cluster Modifications ............................................................................................................. 18
  3.3 Merge in VIAF .................................................................................................................................. 18
  3.4 Split in VIAF ..................................................................................................................................... 18
  3.5 VIAF members may notify errors in the ISNI database .................................................................. 19
4 Improving the matching in the ISNI database ...................................................................................... 19
  4.1 Detecting anomalies .......................................................................................................................... 20
    4.1.1 Date anomalies .......................................................................................................................... 20
    4.1.2 Matching short forms of names within a record ........................................................................ 20
5 Appendix 1 ISNI Sources .................................................................................................................... 21
6 Appendix 2 ISNI Pseudonyms and Anomaly Checks ........................................................................ 21
1 Introduction

1.1 What is ISNI?

The International Standard Name Identifier (ISO 27729) is a new ISO standard capable of identifying the millions of contributors to creative works, including writers, artists, creators, performers, researchers, producers, publishers and more. The standard “specifies the International Standard Name identifier (ISNI) for the identification of public identities of parties; that is, the identities used publicly by parties involved throughout the media content industries in the creation, production, management, and content distribution chains.” The ambitious goal of this new standard is building directly on the functional legacy of library authority files as represented in the contributing databases that comprise the Virtual International Authority File hosted by OCLC.

ISNI was designed as a bridge between existing proprietary right holder identification systems, and resource discovery tools, such as VIAF, to identify and disambiguate names across the contributors’ information systems. ISNI metadata will link the public identity to all of its other manifestations in different systems, making it possible for industry partners to exchange party information without disclosing confidential information. It also provides a single identifying code for a party involved in multiple creative genres (music, cinema, visual arts, literature, etc.).

The ISNI International Agency, which is responsible for ISNI’s administration and governance, was officially incorporated by its six founding members – International Confederation of Societies of Authors and Composers (CISAC), The International Federation of Reproduction Rights Organisations (IFRRO), International Performers Database Association (IPDA), ProQuest, OCLC and the Conference of European National Librarians (CENL - represented by Bibliothèque nationale de France and the British Library) – as a London-based not-for-profit organisation on December 22nd, 2010. As a precursor to establishing a system of registration agencies for the new standard the ISNI-IA partners set themselves the goal of readying the ISNI for industry adoption by first creating the initial ISNI database. This database was launched in November 2011. ISNI identifiers are being allocated by processing data from the databases of the consortium’s founding members and affiliate organisations. The base cross domain file for ISNI is VIAF.

1.2 ISNI and VIAF

The ISNI database includes a copy of VIAF as base data. ISNI and VIAF are complementary databases and systems, both with a common objective of uniquely identifying identities and their names and in sharing data to ensure widespread diffusion of accurate data to data sources on the web at large and to interlink them. While VIAF was an unprecedented achievement in federating the authority files of the libraries and making them visible to the whole world, ISNI is bringing an international ISO identifier to the VIAF database that will be then shared with the VIAF contributors. Beyond sharing identifiers and
metadata, both VIAF and ISNI are sharing community links, in particular with the Databases of Rights Management Societies and Researchers. The need for libraries and Rights Management Societies to work together in name control has become more pressing, particularly for the complete management of joint digitization programs. The ground stone for interoperation was laid by sharing a copy of the VIAF data to be used as base data for the ISNI database and the assignment of ISO identifiers that are subsequently shared with VIAF. ISNI is doing significant quality work and will also provide notifications of system and data improvements to VIAF as they come to light during ISNI operation.

### 1.3 ISNI Scope and ISNI Assignment

ISNI includes personal and corporate name authority records. It does not include meeting names, geographic names and subjects. It also does not include name/title authority records as these were creating duplicates in the ISNI database. Sparse records with only a name are dropped on loading. ISNI is currently loading updates monthly from VIAF. A new VIAF file is available normally on a monthly basis. There is a file comparison program that is used to determine adds, updates and deletes.

ISNI includes data from 20 non VIAF sources (see appendix).

Records on the ISNI database can have one of 3 statuses; assigned, provisional or suspect. Assigned status is given where there are 2 independent sources on a record or where there are 3 or more VIAF sources. There are currently 1.4 million assigned ISNIs (9% of the VIAF records on ISNI have assigned status).

When creating the ISNI database, the VIAF clusters were loaded as individual records into the ISNI database (based on CBS). (As the clusters were loaded as single records, they were not re-evaluated by the ISNI/CBS matching and merging process.) Records from fourteen ISNI contributors were then loaded to the database and ISNIs were assigned as per the rules above. When VIAF updates are processed, a match is made of the VIAF cluster ID and all fields with a VIAF source are replaced. Some records may have been corrected in ISNI and contain a field that will prevent future updates from VIAF. During subsequent VIAF updates, all new records are assessed against the whole database. This may result in duplicate VIAF records being flagged.

The VIAF matching was emulated in ISNI/CBS. Following initial test, the VIAF matching was adjusted and tightened with an emphasis on avoiding false matches. The team in Leiden also worked on multiple programmatic ways to detect records suspected to include multiple identities. The identified cases of suspect records were further divided into those that need manual checking and those that can be fixed by program.

Over the past year in which OCLC Leiden has been developing the ISNI database and system, the ISNI Quality Team at the BnF and BL (representing CENL) have been examining statistical samples to assess the quality of the matching and of the data itself. The team concentrated on false matches rather than on matching failures, as once ISNIs are diffused, merging is less serious than splitting. They found a 5% error rate in VIAF records with assigned ISNI status (i.e. an estimated 50,000); some of the records
represent matching errors but the majority is undifferentiated name authorities. This error rate does not take into account the VIAF contributors where a high percentage of undifferentiated authorities was detected. Data from these VIAF participants were ignored for the purposes of ISNI matching unless they were the only contributors on the record. It has since been noted that quantity of records that are undifferentiated are encoded as such and thus the blanket rule for the contributor that has provided them can be lifted. (The data are also masked from public display if ignored in matching.) The error rate also does not reflect other problems, notably VIAF clusters that are duplicates and clusters with inconsistent treatment of pseudonyms.

This paper attempts to address:

- Notifications of assignment
- Re-diffusion of ISNIs to VIAF participants
- Joint manual processing of suspect error reports
- Work flow for recording errors
- Work flow for transmitting errors to VIAF and VIAF participants and receiving corrections in the ISNI database
- Differences between ISNI and VIAF especially in relation to matching

2 How ISNI contributes to enhance VIAF quality

ISNI is due to send two types of notifications to VIAF – ISNI assignment and error notifications. These comprise:

- Notification of ISNI assignment
- Notification of ISNI replacement
- Notification of duplicates in VIAF
- Notification of data error in VIAF
- Notification of a record to split in VIAF

All these types of notifications are explained in detail below.

The ISNI system already includes an ISNI XML notification schema for notifying ISNI assignment and records needing data correction, duplicates and records with multiple identities, needing splitting.

The formats of these reports with examples are detailed in the document ISNI Data Contributors Reports and Notifications Guidelines.doc., and are available upon request. This document accompanies the XML schema document ISNI notification schema.xsd.

Where an error in ISNI is a result of VIAF contributor data, then the correct procedure is to fix the record at source, upload to VIAF, which will then upload to ISNI.

Where an error in ISNI is a result of VIAF clustering, then the correct procedure is to fix the cluster in VIAF, assign a new identifier and then upload to ISNI.
2.1 Notification of ISNI Assignment

Snapshot reports can be made available with separate reports for ISNI assignment, possible matches and no matches.

Update reports can be provided at regular intervals for new assignments as they are made and for error reports as they are created.

The assignment notifications indicate local identifier (VIAF cluster id), a direct link into the ISNI database, matching data and non-VIAF source (if applicable).

VIAF will store ISNI in the same way that it stores Wikipedia links (as an external link, as opposed to being a VIAF source).

It is expected that VIAF will re-distribute the ISNIs to the contributing VIAF sources and to make the ISNIs available in VIAF outputs, including linked data. Distribution could be via the existing monthly Files but use of the ISNI XML notification Schema is also a possibility.

VIAF Contributors should plan to receive ISNIs into their systems.

*Notification of ISNI assignment will be made from June 2012.*

2.2 Notification of ISNI Replacement

Where two assigned ISNIs have been merged, a notification will be produced indicating deprecated and new ISNI and the VIAF cluster identifier. To date most merges have been made manually by the ISNI Quality Team and most of them involve resolving provisional matches. When a record with a provisional match link is merged with another record, either an assignment notification or a duplicate notification will be made rather than a merge notification. A task to evaluate records with assigned status and eliminate duplicates is planned for Q3 2012.

It is expected that VIAF will replace the ISNIs and notify the contributing VIAF sources. The replacements will be reflected in VIAF outputs, including linked data.

The ISNI XML notification schema will be used as the basis for conveying notifications to change data in the databases of VIAF contributors. VIAF contributors should plan to ingest this format.
Example:

```
<notification>
  <notificationId>116982837_28-11-2011_18:19:15</notificationId>
  <dateTimeOfNotification>28-11-2011 18:19:15</dateTimeOfNotification>
  <localIdentifier>120485229</localIdentifier>
  <identifier>
    <ISNI status="assigned">ISNI 0000 0001 1698 2836</ISNI>
  </identifier>
  <notificationType>
    <error>
      <merge>ISNI 0000 0001 1698 2836</merge>
      <deprecatedISNI>ISNI 0000 0001 1701 2129</deprecatedISNI>
    </error>
  </notificationType>
</notification>
```

VIAF will replace the ISNI, if found, else write the ISNI.

### 2.3 Notification of Duplicates

Where 2 VIAF cluster records have been merged into one ISNI record, a notification will be produced indicating both VIAF identifiers and the associated ISNI. Most times there will be an explanatory note. The ISNI record will retain both VIAF identifiers.

It is expected that VIAF will either add the ISNI to each VIAF cluster or merge the 2 records. The ISNI record has a field that prevents VIAF update. No update from VIAF will occur without prior manual removal of this field, however VIAF will still receive notifications concerning further activity on the record.

Example:

```
<notification>
  <notificationId>120217015_28-11-2011_18:19:15</notificationId>
  <dateTimeOfNotification>28-11-2011 18:19:15</dateTimeOfNotification>
  <localIdentifier>17312093</localIdentifier>
  <identifier>
    <ISNI status="assigned">ISNI 0000 0001 2021 7010</ISNI>
  </identifier>
  <notificationType>
    <error>
      <duplicate>
        <localIDnumber1>166745520</localIDnumber1>
        <localIDnumber2>17312093</localIDnumber2>
      </duplicate>
      <explanation>assigned ISNI = 000000120217010</explanation>
    </error>
  </notificationType>
</notification>
```
The ISNI database record from which the notification was generated is below. The 977 field has been added during the merge process to generate the notification to VIAF. The 670 field has been added manually. The record contains two 035 fields for VIAF with the two identifiers.

001 120217015
003 $0000000120217010$12011-10-10 17:04:41$6assigned$c60 (ISNI)
009 $2VIAF$02012-01-30 14:00:02$cm21
035 ##$5VIAF$017312093
035 ##$5ALCS$01000491
035 ##$5VIAF$0166745520
090 ##$bIPI Number$00188622438$2ALCS (Other identifier or identity)
097 $00000000123851025$aprovisional (Deprecated)
097 $00000000114012174$aprovisional (Deprecated)
370 ##$agb$2VIAF$2BNF
377 ##$aeng$2VIAF$2BNF
400 1#$aAldersley-Williams, Hugh$2VIAF$2VLACC
400 1#$aAldersley-Williams, Hugh$2VIAF$2LC
400 1#$aWilliams, Hugh Aldersey-$2VIAF$2BNF
400 1#$aWilliams, Hugh Aldersey-$2VIAF$2NL$A
400 1#$aWilliams, Hugh Aldersey-$2VIAF$2LC
400 1#$aWilliams, Hugh Aldersey-$2VIAF$2NUKAT
670 ##$bAuthor's web page$uhttp://www.hughalderseywilliams.com/$2ISNI
700 1#$aAldersley-Williams, Hugh$3public$2VIAF$2NAL$A$0000035350658
700 1#$aAldersley-Williams, Hugh$3public$2VIAF$2LCS$0n 87888018
700 1#$aAldersley-Williams, Hugh$3public$2VIAF$2BNF$011240263
700 1#$aAldersley-Williams, Hugh$3public$2VIAF$2VLACC$0000022538
700 1#$aAldersley-Williams, Hugh$2ALCS
700 1#$aAldersley-Williams, Hugh$3public$2VIAF$2NUKAT$0n 00067216
901 ##$a9788874395392$2VIAF$2NLA

910 ##$a@Zoomorphic :$bnew animal architecture$2VIAF$2NLA
910 ##$a@Zoomorphic :$bnew animal architecture$2VIAF$2VLACC
910 ##$a@Zoomorphic :$bnew animal architecture$2VIAF$2BNF
910 ##$a@World design :$bnationalism and globalism in design$2VIAF$2NLA
910 ##$a@World design :$bnationalism and globalism in design$2VIAF$2VLACC
910 ##$a@World design :$bnationalism and globalism in design$2VIAF$2LC
910 ##$a@New American design :$bpproducts and graphics for a post-industrial age$2VIAF$2NLA
910 ###@New American design :bproducts and graphics for a post-industrial age$2VIASF$2VLACC
910 ###@New American design :bproducts and graphics for a post-industrial age$2VIASF$2LC
910 ###@Corporate identity$2VIASF$2NLA
910 ###@Corporate identity$2VIASF$2BNF
910 ###@British design$2VIASF$2NLA
910 ###@British design$2VIASF$2LC
910 ###@British design$2VIASF$2BNF
910 ###@The @most beautiful molecule :bthe discovery of the buckyball$2VIASF$2NLA
910 ###@The @most beautiful molecule :bthe discovery of the buckyball$2VIASF$2LC
910 ###@Periodic tales :ba cultural history of the elements, from arsenic to zinc$2VIASF$2NLA
910 ###@Periodic tales :ba cultural history of the elements, from arsenic to zinc$2VIASF$2LC
910 ###@Panicology$2VIASF$2NLA
910 ###@Panicology$2VIASF$2LC
910 ###@Steelcase Design Partnership (New York, N.Y.)$2VIASF$2NLA
910 ###@Steelcase Design Partnership (New York, N.Y.)$2VIASF$2LC
910 ###@Cranbrook Academy of Art$2VIASF$2NLA
910 ###@Cranbrook Academy of Art$2VIASF$2LC
910 ###@Museum of Modern Art (New York, N.Y.)$2VIASF$2NLA
910 ###@Museum of Modern Art (New York, N.Y.)$2VIASF$2LC
910 ###@Wellcome Trust (London, England)$2VIASF$2NLA
910 ###@Victoria and Albert Museum$2VIASF$2NLA
910 ###@Notenboom, Valerie$2VIASF$2VLACC
910 ###@Hollington, Geoff$2VIASF$2NLA
910 ###@Chartered Society of Designers$2VIASF$2NLA

950 ###@Briscoe, Simon$2VIASF$2NLA
950 ###@Briscoe, Simon$2VIASF$2VLACC
950 ###@Briscoe, Simon$2VIASF$2LC
950 ###@Steelcase Design Partnership (New York, N.Y.)$2VIASF$2NLA
950 ###@Steelcase Design Partnership (New York, N.Y.)$2VIASF$2LC
950 ###@Cranbrook Academy of Art$2VIASF$2NLA
950 ###@Cranbrook Academy of Art$2VIASF$2LC
950 ###@Museum of Modern Art (New York, N.Y.)$2VIASF$2NLA
950 ###@Museum of Modern Art (New York, N.Y.)$2VIASF$2LC
950 ###@Wellcome Trust (London, England)$2VIASF$2NLA
950 ###@Victoria and Albert Museum$2VIASF$2NLA
950 ###@Notenboom, Valerie$2VIASF$2VLACC
950 ###@Hollington, Geoff$2VIASF$2NLA
950 ###@Chartered Society of Designers$2VIASF$2NLA
970 ###@1959-06-17$2ALCS

Notification block 976 – assignment notification; 977 – error notification
976 ###@112-01 15:34:25$68874395392$disbn$e1.000000 $f1.000000
976 ###@112-01 15:34:25$6world design nationalism and globalism in design$dttitle$e0.917800 $f1.000000
976 ###@112-01 15:34:25$6british design$dttitle$e1.000000 $f1.000000
976 ###@112-01 15:34:25$6new american design$dttitle$e1.000000 $f1.000000
976 ###@112-01 15:34:25$6new american design products and graphics for a post-industrial age$dttitle$e0.858513 $f1.000000
976 ###@112-01 15:34:25$6scenes of mild peril$dttitle$e0.021399 $f0.000000
976 ###@112-01 15:44:02$6globalism nationalism and design$dttitle$e0.055232 $f0.000000
976 ###@112-01 15:44:02$6new american design $dttitle$e0.021399 $f0.000000
VIAF will write xA records from these notifications that will force a merge of the records. Before writing the record, the local contributor identifiers from the 2 VIAF cluster IDs would be collected.

Example – Last, Nella

A record in xA has been entered manually with links to the NLA, LC, DNB and SUDOC records to force clustering. Previously only the LC and NLA records were clustering. Note that the xA entry is not in the record distributions.

From the perspective of VIAF data contributors, a merged record in ISNI and then in VIAF may manifest as a change to the VIAF cluster identifier and possible also the ISNI if one ISNI had been received previously.
2.4 Notification of Data Error

Data errors are currently signaled by the ISNI Quality Team or by data contributors. From July 2012 onwards, general users will also be able to add notes and URLs. These will pass to the Quality Team. The Quality Team adds data error notification fields with an explanatory note. At the same time, the team may or may not add an update protection field.

It is expected that VIAF will notify the contributing VIAF source or sources to fix the data at source so that eventually the replacements will be reflected in VIAF outputs, including linked data.

Example:

```xml
<notification>
  <notificationId>078412374_22-05-2012_18:19:15</notificationId>
  <dateTimeOfNotification>22-05-2012 18:19:15</dateTimeOfNotification>
  <localIdentifier>85520132</localIdentifier>
  <identifier>
    <ISNI status="assigned">ISNI 0000 0000 7841 2372</ISNI>
  </identifier>
  <notificationType>
    <error>
      <dataError>
        <explanation>Date of birth is 1961 (IPDA and wikipedia)</explanation>
        <comment>
          $bwikipedia $uhttp://en.wikipedia.org/wiki/Violeta_Urmanavi %C4%8Di%C5%ABt%C4%97-Urmana $externalInformationFound
        </comment>
      </dataError>
    </error>
  </notificationType>
</notification>
```

The database record from which the notification was generated is below. The 670 and 997 fields have been added manually.

```
001 078412374
002 $aTpx
003 $000000000078412372$12011-10-10 16:15:25$aassigned$c60
008 $an$b|$c|$da$e|$fz$gn$hn$ia$j|$kb$ln$m|$na$oa$pa$qa$r|$sc
009 $2VIAF$b2012-05-22$cjanifer
00A 99999999:10-03-03
00B 99999999:14-02-12 12:00:02.000
00D 99999999:10-10-11
00U utf8
00X 0
```
035 ##$1VIAF$085520132
035 ##$1IPDA$010364546
090 $00000000131968311$aprovisional
098 $0131968319
370 ##$alt$2VIAF$2DNB
370 ##$alt$2VIAF$2BNF
375 ##$afemale$2VIAF
400 1#$aUrmana, Violetta$2VIAF$2LC
400 1#$aUrmana, Violetta$d1961-$2VIAF$2DNB
400 1#$aUrmana, Violetta$d1961-$2VIAF$2BNF
400 1#$aUrmanavičiute, Violeta$IWirkl. Name$d1961-$2VIAF$2DNB
400 1#$aURMANA$2IPDA
670 ##$bhttp://en.wikipedia.org/wiki/Violeta_Urmanavi%C4%8Di%C5%Abt%C4%97-Urmana$2ISNI
670 ##$bhttp://www.violetaurmana.com/discography/discography.php/en$2ISNI
700 1#$aUrmana, Violetta$3public$2VIAF$2LC$0no2001025854
700 1#$aUrmana, Violetta$d1961-$3public$2VIAF$2NKC$0xx0035413
700 1#$aUrmana, Violetta$d1961-$3public$2VIAF$2DNB$0124660029
700 1#$aUrmana, Violetta$d1961-$3public$2VIAF$2BNF$014234976
700 1#$aURMANAVICIUTE, VIOLETA$2IPDA
910 ##$a@Oberto, Conte di San Bonifacio$2VIAF$2BNF
910 ##$a@Oberto, Conte di San Bonifacio$2VIAF$2DNB
910 ##$aLove duets$2VIAF$2DNB
910 ##$aИсполнители$2VIAF$2BNF
910 ##$aThe @nightingale$2VIAF$2DNB
910 ##$ale @rossignol conte lyrique en 3 actes d'après Andersen =$bThe
nightingale$2VIAF$2DNB
910 ##$aLe @château de Barbe-Bleue$2VIAF$2BNF
910 ##$aLa Gioconda dramma in quattro atti di Tobia Gorrio (Arrigo
Boito)$2VIAF$2DNB
910 ##$aDas @Lied von der Erde Symphonie für eine Alt- und eine Tenorstimme
und grosses Orchester =$bThe song of the earth$2VIAF$2DNB
910 ##$a@symphonien p$2VIAF$2LC
910 ##$ala @gioconda enregistrement sonore$2VIAF$2BNF
910 ##$a@Wiener Opernfest 2005 Festkonzert 50 Jahre Wiedereröffnung Wiener
Staatsoper$2VIAF$2DNB
910 ##$a@Werke, Singst. Orch$2VIAF$2DNB
910 ##$a@Violetta Urmana - Lieder$2VIAF$2DNB
910 ##$a@Sämtliche Orchesterlieder =$bComplete orchestral songs$2VIAF$2DNB
910 ##$a@Symphony no. 9 en D minor, op. 125 "Choral"$2VIAF$2BNF
910 ##$a@Symphony No. 2 "Resurrection"$2VIAF$2BNF
910 ##$a@Solovej$2VIAF$2DNB
910 ##$a@Scenes from Der Ring des Nibelungen$2VIAF$2DNB
910 ##$a@Ritrovato first commercial recording of famous arias and ensemble
from Puccini's major operas in original, revisited, or alternative
versions$2VIAF$2BNF
910 ##$a@Requiem op. 5 grande messe des morts ;$bfor tenor, choir and
orchestra$2VIAF$2DNB
910 ##$a@Requiem für vier Solostimmen, gemischten Chor und
Orchester$2VIAF$2DNB
910 ##$a@Oberto$2VIAF$2BNF
Some of the data errors are signalled by anomaly programs. In the example below, the program name has been included at the beginning of the explanation.

```xml
<notification>
  <notificationId>354604473_22-05-2012_18:19:15</notificationId>
  <dateTimeOfNotification>22-05-2012 18:19:15</dateTimeOfNotification>
  <localIdentifier>50602308</localIdentifier>
  <identifier>
    <ISNI status="assigned">ISNI 0000 0003 5460 4478</ISNI>
  </identifier>
  <notificationType>
    <error>
      <dataError>
        <explanation>isni_csf_date_anomalies::Inconsistent 970 date '1945-07-14' and 943 date '194.'; Inconsistent 970 date '1945' and 943 date '194.'</explanation>
      </dataError>
    </error>
  </notificationType>
</notification>
```

VIAF will create an xA record using the VIAF iD and the notification information and store it in xA for review.

The ISNI XML notification schema will be used as the basis for conveying notifications to change data in the databases of VIAF contributors. VIAF contributors should plan to ingest this format.
2.5 Notification of a Record to Split

Where a record contains more than one identity, a notification will be sent. Notifications are entered manually by the Quality Team and are also generated by anomaly detection programs.

It is expected that VIAF will:

- Fix errors due to clustering or:
  - notify the contributing VIAF source or sources to fix the data at source
- Assign 2 new VIAF cluster IDs and deprecate the old. (The record has been protected from further VIAF batch update). Is there a field for Deprecated VIAF cluster ID?

Example of a notification entered manually:

```xml
<notification>
  <notificationId>081949227_31-12-2011_18:19:15</notificationId>
  <dateTimeOfNotification>31-12-2011 18:19:15</dateTimeOfNotification>
  <localIdentifier>5367603</localIdentifier>
  <identifier>
    <PPN>081949227</PPN>
  </identifier>
  <notificationType>
    <error>
      <split>
        <comment>
          <externalInformationFound>http://thomasarmstrong.com/bio.php</externalInformationFound>
          <externalInformationFound>http://en.wikipedia.org/wiki/Thomas_Armstrong_%28conductor%</externalInformationFound>
        </comment>
      </split>
    </error>
  </notificationType>
</notification>
```

The database record from which the notification was generated is below. The 977 field in this record is used for notification and has been entered manually. It was detected manually when checking records from the Authors’ Guild that failed to match (Authors’ Guild identities would not be expected to have a death date).

001 081949227
003 $00000000081949226$12012-01-31 17:34:24$asuspect$c39
In the example below, an error has been found in the VIAF clustering. The VIAF cluster can be fixed without requiring referral to a VIAF contributor.

<notification>
<dateTimeOfNotification>31-12-2011 18:19:15</dateTimeOfNotification>
<localIdentifier>95036002</localIdentifier>
</notification>

http://en.wikipedia.org/wiki/Thomas_Armstrong_%28conductor%29


944 $sajm$2VIAF$2DNB
970 $sa$1898$b1994$c1 lived$2VIAF
VIAF will create an xA record using the VIAF iD and the notification information and store it in xA. Subfield c in the ISNI notification field 977 (061C) will record either the VIAF source in error with a mixed identity or a VIAF cluster error where none of the contributing sources contains a mixed identity. E.g. $cNLA (NLA undifferentiated) $cVIAF (VIAF cluster error). VIAF will be able to automatically make two xA records for notifications with $c contents being a VIAF source. Otherwise, the record will be stored for manual review.

The ISNI XML notification schema will be used as the basis for conveying notifications to change data in the databases of VIAF contributors. VIAF contributors should plan to ingest this format.

2.6 Special treatment for Pseudonyms+

The ISNI standard states that ISNIs are assigned to public identities. In order to align library data with data from rights management sources, it is necessary to separate pseudonyms and real names. A suite of programs has been devised to find records and to adjust them for optimal ISNI assignment. Name variants that are detected as pseudonyms or real name forms are moved to related names and where necessary, separate, linked records are made. The documents ISNI Persons, identities and names.doc and ISNI data quality.doc provide more information. There is also more detailed information in Appendix 2 of this document.

Examples:

```xml
<notificationType>
  <error>
    <split>
      <explanation>multiple preferred names?</explanation>
    </split>
  </error>
</notificationType>
```
These records have been marked as suspect in ISNI database. The records that have VIAF and non VIAF sources will be corrected by the ISNI Quality Team. Where these are only VIAF sources, updates from VIAF are not blocked. There will be 10,314 notifications of this type.

A joint project to review these records is to be discussed with the VIAF contributors at the next meeting in Helsinki, August 2012.

```xml
<notificationType>
  <error>
    <split>
      <explanation>related, not variant</explanation>
    </split>
  </error>
</notificationType>
```

Where the records are adjusted for ISNI purposes as above, they are protected from further VIAF update. Notifications are sent for information purposes only. It is not expected that either VIAF or VIAF contributors will act on these notifications because coding as variant rather than related is not an error. It is estimated that there will be 70,000 notifications of this type.

```xml
<notificationType>
  <error>
    <dataError>
      <explanation>cross_reference added</explanation>
    </dataError>
  </error>
</notificationType>
```

As above, these are for information only.

## 3 How VIAF contributes to ISNI

Each month a full file is received from VIAF. ISNI processes the file against the previous month’s file to detect changed and inserted records. The process also detects deleted VIAF clusters (in the old file and not in the new file). VIAF produces a statistical analysis each month indicating:

- Clusters added (to ISNI these are new records) The ISNI and VIAF counts could be compared
- Clusters dropped. The ISNI and VIAF counts could be compared
- Changed clusters (3 different counts – those in both files, those different or missing the the previous file)

### 3.1 VIAF additions
The ISNI load matches on VIAF cluster ID, thus new VIAF records are loaded and matched with the ISNI database.

### 3.2 VIAF Cluster Modifications

These include several cases

- Records changed where there are no new sources
- New sources added to an existing cluster (e.g. in May 2012 = 1,770,307)
- Sources deleted from cluster and no longer on VIAF (e.g. in May 2012 = 9,455)
- Sources deleted from a cluster and moved to another cluster (could be a cluster of one) (e.g. in May 2012 = 184,805)

As already stated, the ISNI load compares VIAF cluster ID and replaces the content entirely. Thus the first 2 cases are correctly handled. For the next 2 cases, the ISNI load will also handle them correctly in most cases, definitely for records where there are only VIAF sources.

VIAF produces a report indicating the deletion and movement of source IDs. This report could be used to identify ISNI records containing non-VIAF contributors that need review in the ISNI database. (The report would exclude records with ISNI notification ID – see 3.4)

### 3.3 Merge in VIAF

- Record will be received as a changed record, including old and merged data
- The record will be matched on VIAF cluster ID within the ISNI database
- The record will include an ISNI notification ID if applicable (i.e. ISNI notification was the origin of the merge). For technical reasons, VIAF cannot report the deprecated VIAF cluster ID.
- ISNI import will replace the VIAF portion of the record. The import will also notice if the merged record is still on file because it will have an “orphan VIAF cluster ID”.
- The record with the “orphan VIAF cluster id” will go through the ISNI delete process. If it is provisional and has no non VIAF sources, it will be deleted completely. Else, the VIAF portion of the record will be deleted in the ISNI database and the record marked for re-matching (see JIRA CBS-3890).

### 3.4 Split in VIAF

- 2 or more records will be received by ISNI.
- If the split was the result of an ISNI notification, all records will have the same ISNI notification ID.
- The “sources deleted from a cluster” reports will also inform the nature of a split
- The records will be matched on VIAF cluster ID within the ISNI database
When ISNI sends a split notification, all non VIAF sources have been stripped from the record and the record is marked suspect and may be protected from update. If a record or records are received with an ISNI notification Id, it will be possible to process the records and replaced and new records and also to remove the suspect status.

### 3.5 VIAF members may notify errors in the ISNI database

VIAF members will notice errors in ISNI that cannot be fixed simply by fixing VIAF. From July 2012, errors can be alerted in the ISNI public interface by clicking the “add comment” button on the left of the full record display.

ISNI public interface: [http://isni.oclc.nl](http://isni.oclc.nl)

### 4 Improving the matching in the ISNI database

The document “ISNI evaluation.doc” shows how ISNI matching was configured from VIAF matching documentation. The main differences from VIAF matching are:

- From any one record, only the fullest form of name is used for matching
- The database was mined for common surnames and the scores for matching on initials and dates have been downgraded for common surnames such that other strong matching elements are need for a match
- The tolerance for dates was tightened. Tolerance on both birth and death year is no longer accepted.
- A match on name and year only now needs the support of another matching element (title, publisher, etc.). A match on name and full date of birth is regarded as a match.
- An ISBN match (and other resource identifiers) is given the same weight as a title match
- Personal affiliation matches are given higher matching scores than publishers and institution affiliations but lower than titles and ISBNs
- We invented a new field “contributed to or performed” for serial titles, collected works and performers which is given the same weight as publisher, i.e. lower than title.
- [Because of the large numbers of undifferentiated authorities from the DNB and NLA, we adjusted the matching to ignore their titles and other data unless they were the only VIAF contributors on the record]
- Rob Koopman from the OCLC Leiden ISNI team has classified identities using their titles (and the mined data in WorldCat identities) so classification can used to help match data from disparate sources (theses, journal articles, book titles etc.); giving it a score lower than a title match (this is probably not as relevant to VIAF)
- Hetty from the OCLC Leiden ISNI team is now working on configuring institution matching and affiliations matching, e.g. band members and we will share any additions to the algorithms
4.1 Detecting anomalies

4.1.1 Date anomalies

The data anomaly checker found 7,255 records in error (in December it was 12,800). A sample was tested by the Quality Team and approximately half of them are pointing to errors.

<table>
<thead>
<tr>
<th>VIAF ID</th>
<th>ISNI PPN</th>
<th>Cause</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>163458427</td>
<td>122810279</td>
<td>VIAF cluster error</td>
<td>Newton, Dave</td>
</tr>
<tr>
<td>79061036</td>
<td>140258957</td>
<td>VIAF cluster error</td>
<td>Julian, Claude</td>
</tr>
</tbody>
</table>

4.1.2 Matching short forms of names within a record

All the examples seem to have been resolved.

It appears that VIAF matching has been enhanced.

We want to resolve as many anomalies as possible, before running a program to de-duplicate VIAF within ISNI assigned status (by scanning through the name/title index then sending pairs to ISNI evaluation). Whilst this process will probably be more than 90% accurate, it could cause some good records to merge with records that need splitting.
5 Appendix 1 ISNI Sources

<table>
<thead>
<tr>
<th>Agency</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACCE</td>
<td>Access Copyright, Canada</td>
</tr>
<tr>
<td>AGLD</td>
<td>Authors Guild</td>
</tr>
<tr>
<td>ALCS</td>
<td>Authors’ Licensing and Collecting Society, UK</td>
</tr>
<tr>
<td>AMS</td>
<td>American Musicological Society</td>
</tr>
<tr>
<td>BOEK*</td>
<td>Boekenbank, Belgium</td>
</tr>
<tr>
<td>BOWK*</td>
<td>Books in Print</td>
</tr>
<tr>
<td>BLSA*</td>
<td>British Library Sound Archive</td>
</tr>
<tr>
<td>BRTH</td>
<td>British Library Theses</td>
</tr>
<tr>
<td>CEDA</td>
<td>Centrum Dienstverlening Auteurs- en aanverwante Rechten, Netherlands</td>
</tr>
<tr>
<td>CEDR</td>
<td>Centro Español de Derechos Reprográficos</td>
</tr>
<tr>
<td>ICLA*</td>
<td>Irish Copyright Licensing Agency</td>
</tr>
<tr>
<td>IPDA</td>
<td>International Performers’ Database Association</td>
</tr>
<tr>
<td>CISAC*</td>
<td>International Confederation of Societies of Authors and Composers</td>
</tr>
<tr>
<td>Freebase</td>
<td>Under negotiation</td>
</tr>
<tr>
<td>JNAM</td>
<td>Jisc Names Project, UK</td>
</tr>
<tr>
<td>MLA</td>
<td>Modern Languages Association</td>
</tr>
<tr>
<td>MusicBrainz</td>
<td>Under negotiation</td>
</tr>
<tr>
<td>PROL</td>
<td>Prolitteris, Switzerland</td>
</tr>
<tr>
<td>PROQ</td>
<td>Proquest, Scholar Universe</td>
</tr>
<tr>
<td>RING*</td>
<td>Ringgold</td>
</tr>
<tr>
<td>SCHU</td>
<td>Scholar Universe, Proquest</td>
</tr>
<tr>
<td>VGWO</td>
<td>VG WORT, Germany</td>
</tr>
<tr>
<td>VIAF</td>
<td>Virtual International Authority File</td>
</tr>
</tbody>
</table>

6 Appendix 2 ISNI Pseudonyms and Anomaly Checks

Here is more detail on the pseudonym clean up programs that we are doing.

1. Main name mismatch – records flagged and set to suspect. (Incoming records may update suspect records or match with suspect status records, but they remain suspect and thus no ISNI is assigned.).
   a. VIAF only reported to VIAF for fixing, no further action (10,314)
   b. IPDA has 3071 occurrences (all assigned IPDA will be removed and re-loaded after all pseudonym checks and processing)
c. The rest – 966 occurrences – with ISNI and VIAF sources have been manually checked quickly and those that are wrongly flagged will be reversed. The rest will be fixed manually.

2. Cross checking. Variant names where they do not match the main name in the same record are compared with main and related names of other records. For a candidate pair to be determined, the variant name of record A has to match the main name of record B AND record B must have a related name that matches the main name of record A. The candidate pairs are then assessed by the matching algorithm. Where a good match is found a PPN link is inserted into $9 of the field 400 (variant name) of record A.
   a. These records are sampled for manual review to see if the links made are accurate
   b. After review, if positive, the 400 field is converted to 950 (from variant to related name) and a relationship type “other identity, same person” (new type that could be either pseud or real name)
   c. If there is more than one link, possible match fields are created for all records (77A)
   d. The 950 in record A is copied to 700 in record B, unless a 700 for the same source already exists
   e. The 035 in record A is copied to record B unless a 035 for the same source already exists
   f. The 700 in record A is copied to 950 in record B, unless such a 950 already exists. To this existing or copied 950 in record B, a $9 link is added pointing to record A
   g. If only one records contains a date (970) for the source, then it is copied to the other record.
   h. Records A and B are set to assigned
   i. Notifications fields (977) are added to record A “$z related not variant” and to record B “names and links added” in record B
   j. and also update protection fields (009)

3. Cross checking. Variant names coded as pseudonyms or real names where they do not match the main name in the same record are compared with main and related names of other records. If there is a match of the main name, a candidate match is declared. The candidate pairs are then assessed by the matching algorithm. Where a good match is found a PPN link is inserted into $9 of the field 400 (variant name) of record A. Also the variant name with links will be matched with other variant names in the same record, and where a match is found, the link will be copied. Processing as for 2 above.
Statistics: 3973 VIAF only sources, 446 VIAF and non VIAF

4. Variant name matches related name within the same record. OR. Variant name exists with a code of “pseud” or “real name” or language equivalent. This check is done after the cross record checking and resolution, above.
   a. The fields are converted to 950 with a relationship type of “other identity, same person”
   b. A new record (B) is created with data from A – 700 (from 950), 950 from 700, 035 from 035 and 970 from 970. ISNI Status is set to assigned. No notification field is made but an 009 protect field is created.
Statistics: Estimates 70,000
5. A procedure that tests for record incompatibilities will be run. Currently, only a report is produced. Setting the suspect status is a possibility. The incompatibilities are:
   a. Both birth and death dates are only years and neither matches exactly
   b. Identity is active within 10 years of birth date
   c. More than 120 years between date of birth and date of death
   d. Among names in the record shorter forms match but longer ones mismatch (this check in this program may have been superseded by the main name mis-match check)

<table>
<thead>
<tr>
<th>Version</th>
<th>Date</th>
<th>By</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2012-02-01</td>
<td>Janifer Gatenby</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2012-05-23</td>
<td>Janifer Gatenby</td>
<td>Following meeting in Boston Spa 16.04.2012 and further data loading</td>
</tr>
<tr>
<td>2.1</td>
<td>2012-05-25</td>
<td>Janifer Gatenby</td>
<td>Added section on VIAF to ISNI</td>
</tr>
<tr>
<td>2.2</td>
<td>2012-06-15</td>
<td>Janifer Gatenby</td>
<td>Following review with Thom Hickey and Jenny Toves</td>
</tr>
<tr>
<td>3</td>
<td>2012-06-20</td>
<td>Janifer Gatenby</td>
<td>Following review with BnF and BL</td>
</tr>
<tr>
<td>3.1</td>
<td>2012-06-28</td>
<td>Anila Angjeli (BnF)</td>
<td>Following review with BnF and BL</td>
</tr>
<tr>
<td>3.2</td>
<td>2012-07-02</td>
<td>Janifer Gatenby</td>
<td>Minor editorial changes</td>
</tr>
<tr>
<td>3.3</td>
<td>2012-07-16</td>
<td>Anila Angjeli (BnF)</td>
<td>Following review with BnF and BL</td>
</tr>
</tbody>
</table>
Attachment 4
VIAF and ISNI interoperation

VIAFC meeting, Helsinki, 10 August 2012

An international standard: ISO 27729

A network

A central database of identifiers

What is it?

A governance ISNI-IA

...for a global identifier:

**ISNI: 0000 0001 2101 4127**
ISNI uses all VIAF records for:
Personal names and Corporate bodies

ISNI does not use VIAF records for:
Meetings, Geographic Names, Uniform title records

ISNI includes:
More musicians – composers and performers
More academics
(data uploaded from the other ISNI founding members)
Assigned • (Green)

- Assigned
  - 3 or more VIAF sources = c.850,000
  - 1 or 2 VIAF sources + at least one non VIAF source = 340,000
  - At least 2 non VIAF sources

- Possible matches = 245,440
- Provisional = 15 million
- Suspect = 12,000

ISNI system includes
- Confidence levels (data and match)
- Assignment system (online addition)
- Record links to related names
- Public and Private data
- Direct maintenance
- Sampling, reviewing, merging, splitting operations (OCLC Leiden + Quality Team)
- Notification system
Data quality in ISNI and in VIAF

ISNI matches VIAF data with other ISNI contributors’ data

Finding: Good source data to

ISNI brings to VIAF

More in « VIAF-ISNI interoperation: discussion paper »

• ISNI identifiers including confirmed data
• Improved data quality and confidence
  – Anomaly reports – 7,000 date anomalies (>50% represent real errors)
  – Merge, split and data error reports (c. 5,000)
  – Matching improvements
    • Dates, common surnames, longest name form, weightings, new elements
  – Detection non correctly exploited source data
    • parallel main names, name variant conversion, related names conversion, missed data
  – Pseudonyms
    • Feedback, record links (c. 70,000)
• More widely diffused linked data
• Proposal for inter-operation
Example: Data error reported

ISNI: 0000 0000 7841 2372
Identifier: VIAF, BNF, DNB, LC, NKC, IPDA
Name: Urmana, Violeta 1961-
Name variant: Urmana, Violetta
Urmana, Violetta 1961-
Urmanavičiūte, Violeta 1961-

- IPDA passport derived date of birth = 1961

(agree with Wikipedia)

VIAF date = 1959

Example: Record to split

Two identities in the same VIAF record – detected by Authors’ Guild record matching VIAF record with a death date

Record no.: 13380323X
Identifier: AGLD: 7227
Name: Armstrong, Thomas PhD
Title: Power of Neurodiversity: Unleashing the Advantages of Your Differently Wired Brain, The
What is expected from the libraries

- Ingest ISNIs
- Act on specific notifications of error
- Assist in reviewing records flagged for attention (e.g. possible mixed identities due to date anomalies)
- Flag undifferentiated records input to VIAF (or remove them from the supplied data)
- Report back to VIAF when a record containing an ISNI is split into two identities (include this in the routine authority file maintenance)
Thank you!
VIAF Council Meeting Minutes
National Library of Finland, Helsinki, Finland
Friday, August 10, 2012

Attachment 5
Criteria for Joining VIAF

What is VIAF? VIAF is an international service designed to connect the world’s major name authority files based on cooperative contributions and maintenance by VIAF member libraries. See more about VIAF on the OCLC Web site.

What are the requirements for contributing to VIAF? To apply to become a contributor, your organization must have:

- A substantial authority file and a bibliographic file, both of which you maintain and expand regularly
- Authority data that represents languages/scripts needed to complement the existing VIAF contributions. The current members are eager to add authority data for languages beyond English, French, and German, and especially authority data for non-roman languages/scripts.

How do you become a VIAF member?

1. Complete and submit the application at [https://www.oclc.org/forms/viaf.en.html](https://www.oclc.org/forms/viaf.en.html). The form must be signed by an authorized person at your organization to verify your interest. Your application will be reviewed by current VIAF members.
2. Submit a sample file of authority and bibliographic records to viaf@oclc.org.
   The sample file will be processed and your organization will be able to display and examine the results.
   **Note:** There is no charge and no obligation to participate when you apply for VIAF membership.
3. If test processing of your sample files is satisfactory, and your application is accepted by current VIAF members, your organization can sign a contract, entering into agreement with OCLC to:
   - Send your full bibliographic and authority files to OCLC.
   - Send updates to your records to OCLC on a regular cycle.

The contract you receive will detail the rights and responsibilities of VIAF membership.