



ISNI Overview

The Management of Scholarly Identity

Baltimore, April 4th 2012

Beat Barblan

Director, Identifier Services, Bowker

Identification of Public Identities

The identification of the identities used publicly by parties involved throughout the media content industries in the creation, production, management, and content distribution chains across multiple fields of creative activity.

A Bridge Identifier

The ISNI is not intended to provide direct access to comprehensive information about a public identity but can provide links to other systems where such information is held.

ISNI Overview



- Incorporated in the UK in December 2010
- Published as an ISO Standard ISO 27729 on 3/15/12!
- OCLC (Leiden) appointed as the ISNI-IA Assignment Agency
 - Created the initial ISNI-IA database (2011)
 - First million ISNIs assigned and being diffused to data contributors
 - Ongoing assignment operations in test through Q1 and part of Q2 2012



The ISNI-IA Founding Members



- **CENL (Centre of European National Libraries)**
 - 48 European national libraries, represented by BL and BnF
- **CISAC (International Society of Authors and Composers)**
 - 225 societies from 118 countries (87% musical composers)
- **IFRRO (International Federation of Reproduction Rights Organisations)**
 - 135 organisations from 74 countries
- **IPDA (International Performers Database Association)**
 - 37 societies from 28 countries
- **OCLC**
- **Proquest/Bowker (BIP, COS, Dissertations, PAD)**

Sample use cases for ISNI



- **Researchers need identifiers**
 - Grant applications
 - Profile management
- **Rights management**
- **European Arrow Project**
 - ISNI with ISTC for registration of digitization rights information
- **Supply Chain**
 - NISO I2 committee recommended the ISNI Assignment system for all institutions in the Digital Supply Chain
- **Music Industry**
 - Need for a unique global level identifier shared by record labels and distributors

- **International**
- **Cross domain**
 - Creators and other contributors in all disciplines (authors, editors, translators, illustrators, composers, actors, performers, artists, researchers)
 - Organizations that are part of the supply chain of created works (publishers, aggregators, retailers, database vendors, libraries,...)
- **Centralized registration**
 - Global network of registration agencies

WORK
(intellectual creation)

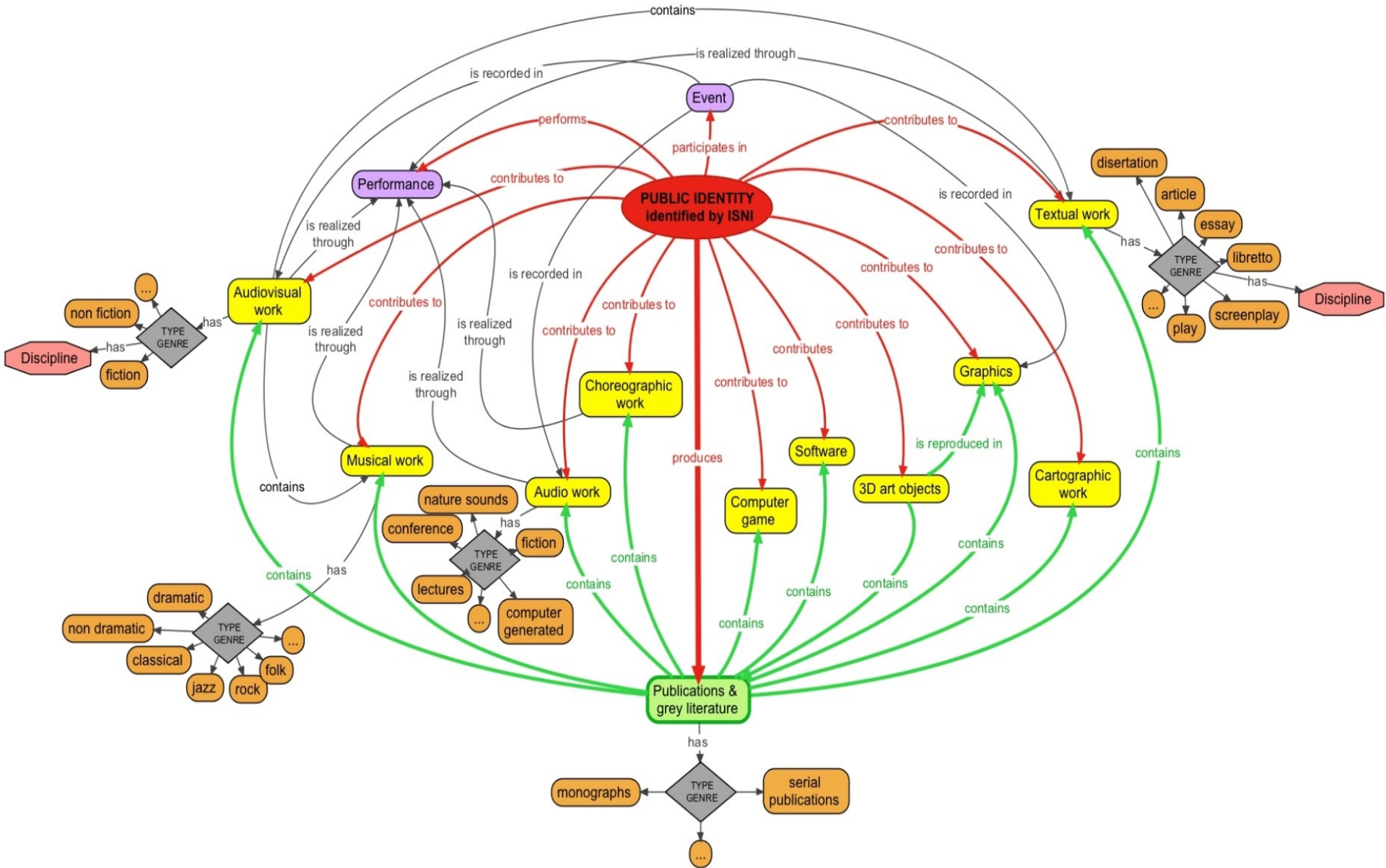
PUBLICATIONS & GREY LITERATURE
(implies manufacturing and had craft)

PERFORMANCE / EVENT

TYPE GENRE
(distribution node)

the possibility to assign any specific GENRE or TYPE if necessary

DISCIPLINE
(the possibility to assign a class from a KOS such as DEWEY)



Search

Select Field: Select Index: Search Terms:

VIAF: The Virtual International Authority File

VIAF, implemented and hosted by [OCLC](#), is joint project of several national libraries plus selected regional and trans-national library agency to lower the cost and increase the utility of library authority files by matching and linking widely-used authority files and making that information available on the Web.

Participating institutions are:

- VIAF (Virtual International Authority File)
 - 12 million+ authority records
 - from 26 national and major research libraries
 - Harvested and managed by OCLC Research; matching algorithms refined over 5-6 years

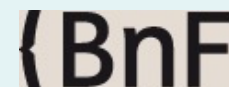
Base cross domain file of the ISNI database, building on work already done

-  [National Library of Australia](#)
-  [Library and Archives Canada](#)
-  [National Library of the Czech Republic](#)
-  [Bibliotheca Alexandrina \(Egypt\)](#)
-  [Bibliothèque nationale de France](#)
-  [SUDOC \(France\)](#)
-  [Deutsche Nationalbibliothek](#)
-  [Getty Research Institute](#)
-  [National Széchényi Library \(Hungary\)](#)
-  [National Library of Israel](#)
-  [Istituto Centrale per il Catalogo Unico \(Italy\)](#)
-  [NUKAT Center \(Poland\)](#)
-  [Biblioteca Nacional de Portugal](#)
-  [Biblioteca Nacional de España](#)
-  [National Library of Sweden](#)
-  [Swiss National Library](#)
-  [RERO \(Switzerland\)](#)
-  [Library of Congress/NACO](#)
 - [Biblioteca Nacional de México](#)
 - [British Library](#)
 - [National Agricultural Library \(U.S.\)](#)
 - [National Library of Medicine \(U.S.\)](#)
 - [National Library of New Zealand](#)
 - [National Library of Scotland](#)
 - [National Library of South Africa](#)
 - [National Library of Wales](#)
-  [Vatican Library](#)

Leveraging high confidence data from different domains



- **British Library**
 - JISC names (research grant data),: UK theses, (ZETOC (Possible))
- **CISAC**
 - IPI - International Party Identifier (87% musical composers), 2 million records
- **IPDA (International Performers Database Association)**
 - 500,000 performer records
- **IFRRO (International Federation of Reproduction Rights Organisations)**
 - Including access© (Canada), ALCS (UK), CEDAR (Netherlands), CEDRO (Spain), Librius (Belgium), Prolitteris (Switzerland), VGWort (Germany)
- **Proquest/Bowker**
 - BIP (books in print), Theses and Scholar Universe, American professional societies



- **Team at Bibliothèque nationale de France and the British Library**
 - Manual checking of statistical samples
 - Review by data source - input to data policies
 - Establishing the percentage of data errors & characteristics
 - Resolution of queries from RAGs and general public
- **Program analysis**
 - Creation of anomaly checker
 - Calculation of Dewey classification; Creation class

- Database
- Data model and data privacy
- Assignment
- Diffusion
- Enhancements and changes
 - Public interface for comment and URL
 - Contributors may change their own data; integrity checks may cause record to merge or split
- Notification and Diffusion of changes and enhanced metadata

- **Importance of Data Privacy**
- **All data is held with a source code**
- **Source profile determines for each field class**
 - **Whether a field can be displayed or is only for matching**
- **Data may be deleted; except core metadata for assigned ISNIs**

- **Personal names - Primary match fields**
 - Name, name identifiers, dates, titles of resources, title identifiers, co-authors
- **Personal names - Secondary match fields**
 - Institution affiliations, publishers, nationality, gender, partial titles, experimenting with Dewey classification of titles
- **Organisations - Primary match fields**
 - Name, name identifiers, address, organisation type, active dates, associated persons (e.g. Band members)
- **Organisations - Secondary match fields**
 - Titles of resources, affiliated organisations

- **An ISNI is assigned where:**
 - **Metadata from 2 or more independent sources matches with a sufficient level of confidence (match confidence)**
 - **Or metadata has 3 or more VIAF sources**
 - **Metadata is complete and unambiguous**
- **All records have a data confidence level**
 - **Indicating closeness of contact with party behind the identifier**
 - **Rights management societies have highest confidence**

Centralised Registration / Diffused Collection and Management



- **Network of nodes and expertise**
 - RAGs for gathering, completing, assessing the quality of input data , responding to disambiguation responses
 - Reference Databases for input to disambiguation, matching, assignment and correction
 - RAGs / Reference Databases for diffusing ISNIs and promoting usage
 - Assignment Agency stores URLs in database and sends notifications of changes and corrections
 - All ISNI members responsible for Quality; special role for the Quality Team

- **ISNI-IA is advocating one shared scheme**
 - Confusing to have 2+ identifier schemes appearing at the same time for the same identities
 - Dilutes effect of linked data
 - Corrections easier to administer with one scheme
 - Cross domain identification - e.g. writer of scientific articles also book author and song writer
 - ISNI linking identifier in ISO (ISBN, ISTC, ISAN, ISWC ++)
- **Interoperating systems**
 - Method to be negotiated; aim for SYNERGY
 - ORCID's focus is on end user input; ISNI's is on registration

- **ISNI-IA**
 - Not for profit, incorporated in the UK, unprecedented cross domain alliance
 - Funded the creation of the database and assignment system
 - Assignment Agency and RAGs on RAND cost recovery
 - Ongoing costs are modest - no permanent staff, permitting price per ISNI to be as low as possible
- **Emphasis on registration and data quality, building on existing data**
- **Data privacy is respected while core data is open**
 - **diffusion and usage of assigned ISNIs is free and is encouraged**