Authority Control for Metadata in Institutional Repositories

Northeast Institutional Repository Day 2021

December 2, 2021

James H Cho

- Assistant Professor
- Catalog and Metadata Strategies Librarian
- Adelphi University Libraries
- 1 South Ave. Garden City, NY
- jcho@adelphi.edu
- Pronouns: He, Him, His

What is an Institutional Repository?

"An institutional repository showcases scholarly and research output to the wider community, and significantly helps in institutional advancement and outreach. It has the benefit of sharing and marketing with others. It also provides an opportunity to raise the profile and brand awareness of an institution, faculty and staff to the global community." (Nagra, 2012)

Welcome to Adelphi University's Scholarly Works

Institutional repositories bring together all of a university's research under one umbrella. Adelphi University's Scholarly Works showcases our campus scholarship and holds works by faculty, students, and staff. You can search for works created by faculty and staff and thesis and dissertations by students.

Open Access | Celebration of Scholarly Research & Creative Works

Works | Academic Units

Adelphi Scholarly Works – Our Institutional Repository

An instance of Esploro from Ex Libris, adopted in June 2020

What are the User Needs for Patrons Searching in an Institutional Repository?

- To see the scholarly output of academic institutions, units, and individual faculty.

- A student who wishes to see if a particular professor is conducts research in an area of interest.

- To see the impact the impact of research production of an university and its departments and individual faculty.

- An academic who is looking for a job and wishes to know what a particular department produces.

- A faculty member who wishes to see what their colleagues are producing.

- To know what a cross section of faculty is producing in a subject area.

- **THEREFORE, THE USER WILL MOST LIKELY BE SEARCHING FOR AN INDIVIDUAL, CERTAIN ACADEMIC UNITS AND DEPARTMENTS AND FOR SUBJECT AND RESEARCH INTERESTS.**

Cataloging with the User in Mind

- Catalogers make resources discoverable and accessible.

- "The convenience [of the user] means that all efforts should be made to keep all data comprehensible and suitable for the users." (International Federation of Library Associations and Institutions, 2016)

- "The essential and defining objective of a system for organizing information, then, is to bring essentially like information together and to differentiate what is not exactly alike." (Svenonius, 2000)

What is Authority Control?

- "The process called authority control was so named because it was thought necessary to determine an authorized form for every entity known by variant forms." (Taylor and Joudrey, 2009)

- "Authority control is needed for collocation – for bringing together everything related to a person, family, corporate body, place or work regardless of what name has been used for one of those entities." (Taylor and Joudrey, 2009)

Why is Authority Control important to Institutional Repositories?

- Because of these user needs (noted above), name and subject authority control is vital to for the accessibility of resources and for the satisfaction of the end users.

- A successful search by the patron is determined by the ability to find research assets by creator name, their academic and unit affiliations, and by subject headings in the search bar.

- Authority control also enables access points/hyperlinks within individual bibliographic records to link to other records and resources with similar metadata.

- **AUTHORITY CONTROL IS THEREFORE VITAL TO THE SEARCH NEEDS OF PATRONS WHEN ENGAGING WITH INSTITUTIONAL REPOSITORIES.**

Name Authority Control

- Name authority control is necessary to identify and group together all of the name variants that a person uses to refer to the same individual.

- Throughout the career of an author, he, she or they may publish under different names. The author may choose to use a full name, a nickname, a married name, different initials, or no initials at all.

- Authority control will allow the patron of institutional repositories to find all of the research assets that the same researchers have published under.

- Also important is to identify the academic department that a faculty or staff belongs to in order to search by academic unit, to further distinguish the researcher, and to group the researcher with other authors within the same unit.

- NAME AUTHORITY CONTROL IS VITAL TO USERS OF INSTITUTIONAL REPOSITORIES SINCE ONE GOAL IS TO ACCESS ALL THE RESEARCH AND SCHOLARSHIP PRODUCED BY THE SAME SCHOLAR.

Controlled Vocabularies

- A controlled vocabulary allows a patron to identify all resources with the same intellectual content.

- Using a controlled vocabulary has the benefit of linking together the records of assets with the same subjects.

- It assists in searching by subject in the public interface and by finding subject headings within individual bibliographic records that links to other records with the same intellectual content.

- CONTROLLED VOCABULARIES ARE IMPORTANT TO USERS OF INSTITUTIONAL REPOSITORIES SINCE ONE GOAL IS TO SEE ALL THE RESEARCH BEING PUBLISHED AT AN INSTITUTION AND ACROSS DEPARTMENTS AND UNITS WITH THE SAME INTELLECTUAL CONTENT.

Authority Work

- "In order to have authority control it is necessary for someone to do authority work, which requires that someone identify all variants for a name or title of a work and make the necessary decisions about which variants represent the same names, which should be the authorized form, and which should be references." (Taylor and Joudrey, 2009)

- Catalogers are accustomed to conducting authority work at the local level.

- In institutional repositories, authority work must be done at the local level because of the lack of capabilities in IR platforms.

Institutional Repository Platforms

- Common Institutional Repository platforms include Digital Commons, Dspace, Fedora, and Esploro.

- Different platforms have different capabilities, but they all have similar problems regarding name and subject control.

- **Esploro** is the Institutional Repository platform used at Adelphi University. It will serve as an illustration on how to make local solutions to the problems of name and subject authority control.

Name Authority Control in Esploro

- As a first step, human resources records were consulted for preferred names and academic unit affiliations.

- Metadata in Esploro, including names, are character-based, so the decision was to use the full name.

- Jane Elizabeth Doe was used in place of Jane E. Doe. The full name will match with Jane Doe, Jane E. Doe, J. Doe, and J. E. Doe. But if Jane E. Doe is used as the preferred name, it has the potential to be confused with Jane Emily Doe.

Controlled Vocabulary in Esploro #1

- In Esploro, there is a feature to add research topics to the research asset records. This is very limited. There is another field to add keywords when there are no appropriate research topic.

- As we have seen, metadata in Esploro is character based. Any slight change creates separate access points/hyper links.

- For example, "non-profit organizations" is different from "not-for-profit organizations". These two terms will create two different access points/hyper links.

Controlled Vocabulary in Esploro #2

- Library of Congress Subject Headings was implemented as the controlled vocabulary in Esploro so that all works with the same intellectual content can be grouped together.

- For example, some researchers used the terms "elderly", "seniors", or "the aged". The Catalog Librarian identified the LCSH term to encompass all of these term variants.

- There is debate as to who should be creating controlled vocabulary terms. Some argue that researchers know their work most intimately. Others argue that the expertise of catalogers are needed.

- These strategies are not mutually exclusive. It can be thought that researchers are conducting subject analysis in providing keywords and catalogers are translating these terms into a controlled vocabulary.

- At times, LCSH terms are not sufficient in term of granularity, especially in STEM disciplines. This is an instance where we decided to keep the keywords applied by the researchers, since they know their work best. Even if they are not an official controlled vocabulary term.

Limitations of the Controlled Vocabulary in Esploro

- At times, LCSH terms are not sufficient in term of granularity, especially in STEM disciplines. This is an instance where we decided to keep the keywords applied by the researchers, since they know their work best. Even if they are not an official controlled vocabulary term, the researcher applied keyword was used.

- One drawback is that it is impossible, unlike LSCHs in MARC records, to combine all the variants, synonyms, homonyms, narrower and broader terms into one heading. Because of this lack of feature in Esploro, the search experience is somewhat compromised. However, being able to link records and assets with the same intellectual content is the great benefit of using a controlled vocabulary in this way.

Work Flows – Cataloging Oversight

- There are many ways research deposits are batch uploaded into institutional repositories.

- In Esploro, these included Smart Harvesting, Smart Expansion, and migrating data from a legacy system.

- Because of limited time, staff and resources, these ways of populating metadata and research assets frees up librarians to focus on cataloging oversight.

- Authority work should be conducted after the fact, at the level of cataloging oversight. To ensure that both name and subject authority control is implemented and sufficient to user needs.

References

- International federation of library associations and institutions. (2016). *Statement of International Cataloging Principles.* *https://repository.ifla.org/handle/123456789/80*

- Nagra, K.A. (2012). Building Institutional Repositories in the Academic Libraries. *Community & Junior College Libraries*, *18*(3-4), 137–150. https://doi.org/10.1080/02763915.2012.799028

- Svenonius, E. (2000). *The intellectual foundation of information organization.* MIT Press.

- Taylor, A. G. & Joudrey, D. N. (2009). *The organization of information.* Libraries Unlimited.