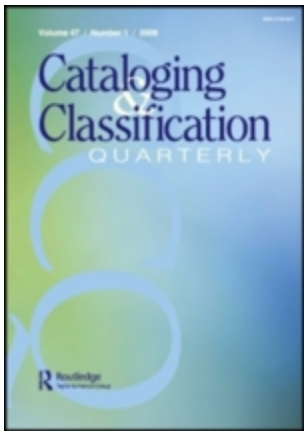

Authority Control Today: Principles, Practices, and Trends

Journal:	*Cataloging & Classification Quarterly*
Manuscript ID	WCCQ-2020-1273
Manuscript Type:	Special Issue
Keywords:	Authority Control, Catalog maintenance < Library catalogs / OPACs, Controlled vocabularies, Cooperative cataloging, Information retrieval, Library catalogs

SCHOLARONE™
Manuscripts

Authority Control Today: Principles, Practices, and Trends

Authority control enhances the accessibility of library resources, improving users' ability to efficiently find the works most relevant to their information search. While authority control and the technologies that support its implementation continue to evolve, the underlying principles and purposes remain the same. Written primarily for a new generation of technical services librarians, this paper illuminates the importance of authority control in cataloging and library database management, discusses its history, describes current practices, and introduces readers to trends and issues in the field, including future applications beyond the library catalog.

Introduction

Authority control is a key aspect of the cataloger's work. Through the practice of authority control, an index of authorized terms is formed in the catalog, allowing for creators and subjects to be named and recalled by consistent terminology. Doing the work to identify preferred terms for new names and subjects to add to the catalog in such a way that similar terms are differentiated and variant forms are linked together ensures the library's patrons will receive more accurate and complete results in their information search. The maintenance of this controlled vocabulary is important, as language and nomenclature are living things.

In the days of the card catalog, authority maintenance was a very time-consuming and manual process, with each change requiring the original term to be crossed out and replaced by the updated term handwritten on every affected card from the catalog or the typing or printing of entirely new cards to incorporate the changed terms into the card catalog. This process is now considerably more efficient and continues to improve as technology develops and as cooperative practices evolve. A basic understanding of the historical and current practice of authority control,

the importance of such work, and an introduction to current trends and the potential evolution of authority control in the future will serve the catalog librarian in the practice of this important work and will allow for optimal utilization of the database management tools of today's library catalog.

History of authority control

The history of authority control is intertwined with the history of library cataloging, including the development of theory, principles, and practices for creating, maintaining, and sharing a library catalog. The library catalog is the r er of all bibliographic items (i.e. information objects) available in a library. This register records various kinds of information about each information object in the library's collection and has taken on varying forms and styles over the years as cataloging principles and practices were discovered and developed.

In the mid-19th century, Anthony Panizzi, a key figure in the development of the library catalog, established a 91-rule plan to "systematize the operation of cataloging."[1] These rules created a prototype of modern day catalogs with entry headings, descriptions, references, and notes recorded for each item in the catalog. Panizzi's 91 rules demonstrated the value of creating uniform headings for recording information about library materials in the library catalog. This creates bibliographic entries with consistent access points used to find and collocate library materials.

Charles A. Cutter's "objects and means" for the catalog[2] codified the catalog's purpose of connecting patrons to library materials and defined the kind of information required to achieve this end. The Statement of Principles adopted by the International Conference on Cataloguing Principles (ICCP) organized in Paris, France (also known as the "Paris Principles")[3] and Seymour Lubetzky[4] showed the value of simplifying and standardizing cataloging rules to create

For Peer Review Only

a universal standard allowing interoperability between library catalogs. Authority control as practiced presently is rooted in these and other key figures and developments that pursued efficiency through consistency in the construction and use of the library catalog.

Up until the late 20th century library catalogs were analog in form. These forms ranged from handwritten tablets, indexes and manuscripts to published books and files of index cards. Each form of the catalog listed library materials alphabetically by the author's name, title, subject, or call number. These same analog forms for recording library materials stored the approved terms for names, titles, and subjects that were used as headings in the creation of catalog entries. For example, each card in a card catalog contained at the top the primary way the item could be found (i.e. the main entry heading) followed by other descriptive detail and physical location details to help patrons find the item they needed (see Figure 1). In these formats, items in the library catalog authored by the same person would be collocated together in the list or index card file because the form of the person's name was consistently recorded and filed. Entries (either in a list or on a card) also contained references to other related headings and items, helping guide the patron to their desired item.

The development of computer technology enabled printing the analog catalog, making the production and maintenance of the catalog more efficient. These technological advances also led to the development of MARC[5] as an encoding standard for both bibliographic and authority metadata[6]. Developed by Henriette D. Avram in the 1960s while working for the Library of Congress, MARC "attempted to both convert and manipulate the data stored on a catalog card."[7] In 1974 the Library of Congress began issuing lists of newly created and updated name headings for use in other libraries. The first MARC-based authority records were created by the Library of Congress in 1977, and the first edition of *Authorities: A MARC Format* was published a few

years later in 1981, making available a national standard for recording, storing, and sharing electronic authority records. In a relatively short amount of time, MARC developed into an international standard that increased sharing and interoperability between libraries and their catalogs.[8]

As computer technology improved and computer networking technology developed, the library catalog was converted to an electronic environment, removing the need to publish an analog catalog list or card file. Files of authorized headings were also converted into this new digital environment. In a computer environment that involves various networks, creating, updating, and sharing catalog metadata became much more consistent and efficient. Libraries adopted integrated library systems developed by 3rd-party software companies that connected the library's many functions to the metadata in the catalog (e.g. acquisitions, circulation, and discovery). These systems provide patrons a digital (or online) interface for searching and browsing the catalog. Authority metadata is utilized in these search interfaces to guide patrons to the preferred form of terms and headings that are used in catalog records so that patrons know how and what to search to meet their information needs. While enabling access to vast quantities of information, the creation and growth of the WWW and online search engines are recent developments in the history of the library catalog, posing unique challenges and opportunities that continue to shape authority control today.

Authority control defined

Authority control is the process of organizing information by using a singular form or spelling of a name or topic. This process must be understood within the context of the library catalog and its essential functions. The library and its staff are in the business of acquiring, processing, maintaining, and circulating physical and digital information resources (i.e. monographs, e-

books, journals, reference materials, scholarly articles, archival collections, audiovisual items, databases, maps, etc.). In order to make these resources accessible, catalog librarians create, update, and maintain metadata through careful description and structured information display to assist users in discovering those library resources that will best serve their needs. Cataloging involves description of the information resource, subject analysis for content access and classification to determine the resource's location, what the item is about, and bring together other related resources. Metadata for an information resource is organized in a bibliographic record. Bibliographic records are collected together into a database called the library catalog. The bibliographic record contains the metadata to describe, differentiate, relate, and locate an information resource. In this way, the bibliographic record acts as a surrogate for the information resource, allowing the library patron to learn about a particular item and decide whether it will meet their need without having to examine each potential resource on the shelf.

When catalogers perform authority control (also referred to as authority work) they establish, through verification and validation, controlled headings or terms for various entity types (i.e. people, places, corporate bodies, families, series, works, expressions, subjects, and genres) used in information resource description. Authority control creates a database of consistent, unique terms that includes variant terms, related terms, other associated attributes, and sources of information. These various elements are recorded together in an authority record. A database of authority records can also be referred to as an authority file. The established or authorized form chosen in an authority record is the form used within bibliographic records in the description and subject analysis of an information resource acquired by a library. Using these controlled terms creates predictable and consistent metadata that can be used to differentiate between similar entities and collocate related resources.

Authority records consist of four major components: the authorized form of the entity, variant forms (i.e. synonyms), attributes and information about the entity, and source information to support the choices made in establishing the entity. Authority records can be formatted using various metadata encoding standards. Libraries primarily encode authority records using the MARC 21 format for authority data[10]. Figure 2 illustrates a MARC authority record for a person entity. Each piece of information in the MARC record is housed in a numeric tag that is machine-readable (see Figure 5). The form of the personal name established in the 100 tag is the authorized form for this person, meaning that anytime an information resource by or about this person is added to the library catalog the metadata description for the item will use this exact form of their name. The 4XX tags contain variant forms of the name that provide "see from" references to guide patrons and staff to the established form in the 1XX tag. 5XX tags generate "see also" references that represent a relationship between the entity in the authority record and another established entity such as a recognized pseudonym under which the author also writes.

Within the authority record, X tags contain attributes associated with the person including related locations (e.g. birth and death place), areas of professional activity or expertise, occupation, associated organizations, gender, language, and fuller form of the name. The 6XX tags provide additional notes that guide catalogers in using the data in this record. In particular, the 670 tag identifies sources that show evidence for the decisions and metadata included in the record. Most importantly, these sources provide evidence for the form of the name chosen in the 1XX field. Tags 000, 001, 005, and 008 represent fixed field data to identify information about the record itself, including the type of record, when the record was created, and how the record can be used. Tags 010 and 035 are control numbers that uniquely identify the authority record in library systems and databases and facilitate record overlay when updating authority records. Tag

For Peer Review Only

040 identifies the language of description and the content standard for the metadata. It also identifies the institutions that contributed to the creation and maintenance of the record. Dates related to the entity, including birth and death dates, are coded in the 046 tag. Any 9XX tags store local information specific to the source database for the record. In addition to personal names, authority records are created for place names, corporate bodies, families, series, works, expressions, topical and geographic subjects, and genres[11]. The MARC format provides a sophisticated encoding standard for recording, maintaining, and sharing authority metadata.

Importance of authority control

The importance of authority control lies in its ability to support users' information retrieval needs through the authorized terms in the authority file acting as a controlled vocabulary with consistent, reliable, and unique terms. This brings precision to searches and collocates related materials in results lists. The structure of authority records with cross references and hierarchically related terms collocates works on the same topic and improves navigation between related concepts. It also allows for linking between library resources and other tools, especially online.[12] End users benefit from the predictability of consistent naming and more precise results.

Library staff also benefit from the consistent application of authority control practices within the catalog. Whenever an item needs to be added to the catalog that has the same author as another work already cataloged, the time spent describing the new item by the cataloger is decreased if the name has already been established in the library's authority file. Maintaining a current copy of authority records in the local file is important. Outsourcing the portion of authority control work that can be automated can improve catalogers' time and resource management,[13] allowing catalogers to spend their limited time on the portion of authority work that requires their especial expertise. Establishing entities in a prescribed and structured way

For Peer Review Only

ensures that both catalogers and patrons know how to differentiate between similar entities when cataloging or searching the library catalog for information resources.

Barbara Tillett explained the essential nature of authority control within libraries over three decades ago, stating "In order to accomplish the finding and gathering functions, the catalog must have authority control. Authority control is inherent to a catalog and without it, a file cannot be considered a catalog."[14] Technical services librarians today must be just as passionate about this essential function of their jobs when communicating with administrators about authority control to ensure continued support for this work.

How authority control is used in libraries

In the process of cataloging an information resource the cataloger chooses main entry headings to provide a predictable heading for access, to uniquely identify the resource (e.g. by choosing the item's author, series title, subject, additional contributors, etc.) and to collocate related resources. The terms chosen for these headings are selected from an authority database. If the chosen heading to best represent the information resource exists in the library's local authority database, the cataloger may simply re-use that existing heading. If the author entity does not yet exist in the local authority database, the cataloger will add a new authority record to the file, whether by creating a new record or downloading an existing authority record from an external source. If the library participates in a cooperative cataloging program, they may also establish the named entity in the cooperative authority database.

Libraries maintain a local authority file using a variety of means including batch loading records, partnering with library vendors, participating in cooperative cataloging programs, and manually creating or editing authority records one-by-one. A library's local authority database is typically maintained as part of their integrated library system (ILS) alongside other library

For Peer Review Only

metadata including bibliographic records. To disambiguate seemingly similar but different resources and collocate related resources, the ILS forms relationships between authority records and bibliographic records by matching on the authorized form of a heading from the authority database. When a term used in a bibliographic record matches the authorized form from an authority record, a link is created and indexed in the system that also includes links to and from related variants and references defined in the authority record.

As an example, consider an authority record representing Steve Jobs, the founder of the corporate body Apple, Inc (see Fig). Bibliographic records for information resources attributed to Steve Jobs or about Steve Jobs will include the form of his name found in the 1XX field of the authority record as a heading in the bibliographic record (see Fig). The authority record for Apple, Inc. includes a link to Steve Jobs as the founder of the corporate body (see Figure 3). The ILS does the work to index and collocate these related records for later search and retrieval. The library catalog or discovery system that a library uses to help patrons search, browse, and discover information resources can utilize the information recorded in the authority record to collocate materials with bibliographic metadata that shares the same authorized headings, direct patrons from variant forms and terms to the authorized form, and disambiguate similar headings in search results (see Figures 6-7).

A library's local authority database stands in contrast to cooperative (or aggregate) authority databases maintained on a regional, national, or international scale. Libraries can participate in cooperative cataloging programs allowing them access to a pool of shared authority records as well as the ability to contribute new or updated records to the shared database. The Library of Congress's Program for Cooperative Cataloging (PCC) maintains various cooperative cataloging initiatives including the Name Authority Cooperative Program (NACO) and the

For Peer Review Only

Subject Authority Cooperative Program (SACO). These two programs create and maintain a national database of authority records that can be leveraged by libraries in the creation and maintenance of their own local catalogs. Libraries participating in these cooperative programs can automate bringing records from the shared cooperative databases into their local authority database.

Learning how to perform authority work is more often accomplished on the job than in library school, especially to the level of competence that is required to effectively maintain good authority control.[15] This can be achieved through a combination of in-house training, attendance at workshops or online trainings, staying up to date on library literature related to authority control, and regular, practical application through mentored authority work during cataloging activities.[16]

Automated authority control

Many libraries use vendors to obtain authority records and to purchase authority control services as means of automating the clean up of headings in their bibliographic database and the maintenance of their local authority file. Due to the labor-intensive nature of authority control, outsourcing some of this work can increase efficiency and control costs.[17] The number of options for outsourcing authority control services has decreased in recent years to a handful of vendors,[18] the majority of which provide similar services. When a library is shopping for an authority control vendor, a formal request for proposal (RFP) may be issued, allowing for vendors to respond to the library's listed requirements and formally bid their services for comparison. However, a simple price quote request may be sufficient for the library's needs, as long as there is thorough communication between the library and each potential vendor about the specific services that will be included and documentation of these options.

For Peer Review Only

Once a vendor has been selected, the library works with an assigned project manager to develop a project profile which documents the library's selections on a number of processing options related to file transfer, bibliographic record cleanup and enrichment, bibliographic heading and authority file matching, record output, and reporting.[19] Processing samples can usually be requested before or after contract signing to establish expectations for the results of the desired service.

Automated authority control processing of the library's full bibliographic database establishes a master or base file. The library's bibliographic headings are cleaned up and normalized in preparation for matching against the selected authority files, especially Library of Congress Names and Subjects. If a heading matches the 1XX or 4XX from an authority record, the authorized form will be inserted into the bibliographic record and the authority record will be included in the output files. The library receives a cleaned up file of their bibliographic records and all the associated authority records to load into their ILS. Customizable reports are also included, allowing libraries to follow up the automated processing with any needed manual review for headings that could not be confidently matched by the machine. This maximizes the benefits of an automated authority control workflow while allowing library staff to focus their efforts on the authority work that most needs their expert evaluation.

Optional ongoing services can also be contracted for continued automated maintenance of the library's authority file. Bibliographic records that have been added to the library's database since the original base file processing are sent to the vendor to undergo the same process. Vendors may also provide periodic reports and MARC extracts of the authority records which have changed since the library's last processing.

For Peer Review Only

For many libraries, outsourced automation of authority control is a routine aspect of technical services work. With the changing landscape of metadata tools, shrinking cataloging departments, and evolving library collections, some libraries are developing methods for automating authority control in-house, using MarcEdit, SQL queries, and batch processing.[20] For programming-savvy librarians, these methods can be worthwhile to explore. Another potential new trend is the provision of authority control within the ILS itself, as debuted by Ex Libris' Alma.[21] If other library systems begin offering built-in authority control functionality, vended authority control may become a less common workflow.

Beyond the library catalog

Authority control for archives

Libraries often need to consider how archival authority structures complement bibliographic authority records and how these can work together in their discovery system. Many libraries, especially within academic institutions, coexist with archives, which often means that bibliographic records for traditional library materials must also commingle with descriptions of archival materials, whether within the library catalog, through a discovery layer, or in other applications. While collection-level MARC records for archival materials may reside in the library catalog to allow for books and related archival materials to both be found within the same information search, the finding aid remains "the preferred method for describing archival materials."[22][23] Finding aids generally reside outside of the library catalog in another description platform such as an archival management system like ArchivesSpace or in a web-based discovery tool.

Authority work for archival materials must often be undertaken in multiple realms. The collection description in the library catalog will usually conform to traditional library authority

For Peer Review Only

control procedures and utilize authorized terms from library authority thesauri. The corresponding archival description within the library's finding aids database also benefits from authority control. Archival authority records describe persons, families, and corporate entities associated with a body of archival materials and may be created utilizing the content standard in Part II of *Describing Archives: A Content Standard* (DACS)[24], supported through the structural standard Encoded Archival Context-Corporate Bodies, Persons, and Families (EAC-CPF).

The relationship between archival authorities and library authorities has evolved in recent decades, as the concept of "context control" comes to fruition, especially through the archival authority record's inclusion of "administrative histories and biographical sketches of organizations or individuals who create records" and its documentation of "the relationships between records creators and the records themselves."[25] One noteworthy development in the realm of archival authorities is the international cooperative program Social Network and Archival Context (SNAC), which matches archival authority records from various institutions against the Library of Congress Name Authority File and other sources of authorities to merge records for the same entities. This allows for archival authority descriptions to be linked with related archive, library and museum resources, demonstrating the power of coordination between library and archival authority control practices.[26]

Digital collections and institutional repositories

Information retrieval in other areas outside the bibliographic realm also rely just as heavily on the consistent presentation of names and subjects. Digital collection metadata is an example of another data repository for which authority control should be considered important.[27] Description records for digital objects may be included alongside bibliographic catalog records in discovery layer results lists. Names and subjects need to be consistent across platforms in

For Peer Review Only

order to enhance discoverability and meet user expectation, especially to enable bringing together descriptions of physical library materials with a digitized version of the same items. Institutional repositories are another area that may benefit from consideration within the context of a library's authority control workflows and procedures, particularly since the lack of consistency in name forms in such systems "inhibits retrieval of items by a single author."[28]

Issues and trends

Authorized headings as facets

Advances in computer and networking technologies, including the move towards linked data and semantic web, are influencing the evolution of cataloging and authority control. The use of lexical or keyword searching in library catalogs and discovery systems has pushed metadata creators and stewards towards viewing these authorized terms used in bibliographic description as facets[29] that discovery tools can surface to patrons for filtering and improving information retrieval. The controlled vocabulary Faceted Application of Subject Terminology (FAST)[30] developed by OCLC is one example of this trend towards leveraging authorized terms as facets. As a controlled vocabulary for subject analysis derived from the Library of Congress Subject Headings (LCSH),[31] FAST aims to simplify the control, use, and navigation of subject headings. Using FAST headings in bibliographic metadata enables easier indexing and display in discovery systems for use by patrons during the information retrieval process.[32]

Federated Authority Databases

Computing and networking technologies enable a network of authority files and databases that are increasingly interconnected and open. The Virtual International Authority File (VIAF)[33] is an authority aggregator that collects established authorities from various authority databases throughout the world. This service provides a portal for librarians and information scientists to

For Peer Review Only

identify established authorities for personal, corporate, and geographic names as well as works, expressions, and bibliographic title. VIAF also enables reconciliation services to disambiguate between those entities.

Another service receiving increased attention within cataloging and metadata communities is Wikidata from the Wikimedia Foundation.[34] Wikidata is a core service of the WWW and the semantic web[35] that enables establishing and interlinking entities between various authority databases and services to more fully describe and disambiguate people, places, works, subjects, and genres. Similar to VIAF, it provides a portal for bringing together the various authorized forms of an entity to a single location for reference and use.[36] Adding the unique identifier for Wikidata to a library authority file provides access to other authority sources that could help library staff and patrons uniquely identify entities for their work and research.[37] The Library of Congress's Linked Data Service[38] is an example of a national institution providing access to their authority metadata for use by catalogers as well as developers working on library metadata systems and discovery tools.

Other services aid catalogers in uniquely identifying authors and their creations. Open Researcher and Contributor ID (ORCID)[39] is a service that allows scientific and academic authors and contributors to uniquely identify themselves regardless of how their name is referenced in a publication so that their creative works are correctly attributed and collated together. This service acts as a form of authority control to disambiguate authors and their scholarly contributions. Additionally, ISNI[40] is an international standard identifier and service to establish permanent and unique identifiers for the names of creators across multiple domains. This service allows individuals and organizations to establish an ISNI identifier as well as provide a searchable database for identifier lookup. Finally, domain-specific services like

For Peer Review Only

IMDb[41] and MusicBrainz[42] provide persistent, unique identifiers for entities related to specific resource formats like film and music.

Linked Data

Developed from the early 2000s, linked data and the semantic web[43] are technologies and best practices for publishing data on the WWW. Central to these technologies and best practices is the use of URIs[44] to uniquely identify an entity rather than relying on the string representation (or label) for a given entity[45]. The use of URIs in authority work within a library setting is an increasing trend to facilitate more implicit linking of entities to other datasets, data repositories, and catalogs that exist online. Some authority control vendors include an URI enrichment option to their services to facilitate this linking work. These links will enable a graph of linked entities that can be related and traversed in ways that reveal new paths of knowledge and understanding that were not present previously. These services and technologies support the move to broaden and expand the role of authority control from being discrete headings used as part of cataloging and bibliographic description towards thinking of cataloging as a process of creating and managing entities and their identity. Following this trend blurs the lines between cataloging and authority work.

Ethical authority control practice

In addition to staying informed about technological changes and the opportunities they present, catalogers must also remain aware of developments within other areas in the field. Catalogers are increasingly aware of the significant power they have in the creation of personal name authority records, as they determine how a creator will be described in the authority record. While differentiation is an important aspect of authority work, careful consideration must be employed in determining which of several possible forms of a name should be used as the authorized form

For Peer Review Only

by which an individual will be known and in selecting the key descriptors to include in authority records. What should be done when an author does not want to be named or to have certain information about themselves revealed? How can historically marginalized individuals be recognized for their contributions to works when little information is known about them? Should persons with multiple intentionally separate identities have those identities combined into a single authority record?

As practitioners grapple with these challenging decisions in their authority work, important conversations are emerging in academic literature, conferences, forums, committees, and working groups. A significant monograph edited by Jane Sandberg, *Ethical Questions in Name Authority Control,* builds on previous ethical discussions in library literature to provide an ethical framework for the library community. Chapters explore such issues as self-determination, privacy, the impacts of colonialism, gender variance, transgender identities, the impacts of emerging technologies, and suggestions for how to solve some of the ethical challenges in this work.[46] To bolster the development of each individual cataloger's judgement, the library community is working toward the provision of better institutional guidance for the ethical application of cataloging principles. A new international committee is developing a code of ethics for catalogers which will address key issues. Another resource produced by this committee is a growing cataloging ethics bibliography, providing catalogers a comprehensive curriculum for developing their awareness of the issues at stake and practical solutions for scenarios they may encounter in their work.[47] Keeping abreast of the evolution of ethical authority control practice through studying the literature, participating in community developments, and applying ethical principles as they are learned will ensure catalogers have the understanding necessary to

For Peer Review Only

ethically provide library users with richly connected resources and useful information about the creators of those works.

Into the future

Various initiatives and pilot projects look to shape the future of authority control and identity management. Directed by the Library of Congress and contracted with Zepheira, BIBFRAME[48] represents the future of bibliographic description that is increasingly plugged into the world of the semantic web and linked data. BIBFRAME is a new encoding standard for bibliographic metadata and description that is aimed at replacing the current MARC 21 format. It promises improved interoperability for cooperative cataloging and sharing as well as connecting to broader non-library WWW communities. The development of BIBFRAME represents a move away from thinking of bibliographic metadata as carefully crafted records toward a focus on each element or attribute in a metadata record as data that can be combined and mixed in varying ways and levels not typical with current bibliographic description. The atomization of bibliographic description requires more reliance on authority control that is designed to provide consistent, predictable library metadata.

The Linked Data for Libraries (LD4L)[49] set of initiatives and grants represent the efforts of major academic institutions along with the Library of Congress and the PCC to develop guidelines for implementing the future of catalog metadata and description with BIBFRAME at its center in a shared, cooperative library community. Practical efforts center around catalog metadata creation, indexing, searching, and display as well as how larger WWW communities like Wikidata connect to these efforts. How authority control will inform and shape this work is an area of future research and exploration.[50]

For Peer Review Only

Recent pilot projects have the potential to shape the future of authority control and authority metadata. The PCC URIs in MARC records pilot[51] initiated by the Library of Congress and the PCC explores applying principles of linked data and the semantic web to more consistent and pervasive identifier creation and identity maintenance. The use of standard identifiers enables connecting traditionally siloed library metadata to the larger WWW and semantic web communities. The PCC Wikidata pilot[52] is exploring the future of identity management by looking at Wikidata both as a source for authority metadata that could enrich existing library metadata and serve as a location for publishing existing authority and bibliographic metadata that would benefit library and non-library communities alike. The shared entity management infrastructure project[53] is a project being developed by OCLC and partners within the library community to create a persistent, shared, and centralized system for managing identifiers and metadata for library information resources. This project aims to create an authoritative database of library metadata to connect library collections to communities and resources on the larger WWW and semantic web. Just as authority control has evolved and adapted with past technology developments these projects and efforts will continue to press forward in improving the creation, maintenance, and discovery of information resources.

Conclusion

As a continually growing and changing database, the library's catalog requires ongoing maintenance, especially through disciplined authority control practice. While the work can be time-consuming and intellectually demanding, the pay-off for the user experience is incalculable. The authority control landscape is an interconnected network of libraries, vendors, databases, and services (see Figure 8). Library patrons of today expect seamless information retrieval and sophisticated navigation. Correct application of authority control best practices assists catalogers

For Peer Review Only

in meeting these needs, while connecting users to the most relevant resources for their information search.

Notes

[1] Seymour Lubetzky, "Panizzi vs the 'Finding Catalog'," *Journal of Cataloging and Classification* 12, no. 3 (1956): 152-156.

[2] Charles A. Cutter, *Rules for a Printed Dictionary Catalog* (Washington: Government Printing Office, 1876), 10.

[3] Pino Buizza, "Bibliographic Control and Authority Control from Paris Principles to the Present," *Cataloging & Classification Quarterly* 38, no. 3-4 (2004): 118.

[4] Seymour Lubetzky, *Principles of Cataloging. Final Report. Phase I: Descriptive Cataloging* (Los Angeles: Institute of Library Research, University of California, 1969), https://eric.ed.gov/?id=ED031273.

[5] **Ma**chine **R**eadable **C**ataloging record.

[6] Library of Congress, "MARC Standards," October 2, 2020, http://www.loc.gov/marc/.

[7] Library of Congress, *The Card Catalog* (San Francisco: Chronicle Books, 2017), 152.

[8] Larry Auld, "Authority Control: An Eighty-Year Review," *Library Resources & Technical Services* 26, no. 4 (1982): 323-25.

[9] Commonly referred to as "data about data," metadata is the sum total of what one can say about a given information object at any level of aggregation recorded in a structured form; see also ABC CLIO, "Metadata," *Online Dictionary for Library and Information Science*, October 2, 2020, https://products.abc-clio.com/ODLIS/odlis_m.aspx#metadata.

For Peer Review Only

[10] Library of Congress, "Understanding MARC Authority Records: Machine-Readable Cataloging," August 25, 2020, https://www.loc.gov/marc/uma; Library of Congress, "MARC 21 Format for Authority Data," August 25, 2020, https://www.loc.gov/marc/authority/.

[11] Library of Congress, "MARC 21 Reference Materials," August 25, 2020, https://www.loc.gov/marc/uma/pt12.html#pt12; Library of Congress, "Appendix B - Full Record Examples," August, 25, 2020, https://www.loc.gov/marc/authority/examples.html

[12] Corey A. Harper and Barbara B. Tillett, "Library of Congress Controlled Vocabularies and Their Application to the Semantic Web," *Cataloging & Classification Quarterly* 43, no. 3-4 (2007): 53, http://dx.doi.org/10.1300/J104v43n03_03.

[13] Susan L. Tsui and Carole F. Hinders, "Cost-Effectiveness and Benefits of Outsourcing Authority Control," *Cataloging & Classification Quarterly* 26, no. 4 (1999): 44, http://dx.doi.org/10.1300/J104v26n04_04.

[14] Barbara B. Tillett, "Considerations for Authority Control in the Online Environment," *Cataloging & Classification Quarterly* 9, no. 3 (1989): 2, http://dx.doi.org/10.1300/J104v09n03_01.

[15] Rebecca L. Mugridge and Kevin A. Furniss, "Education for Authority Control: Whose Responsibility is it?" *Cataloging & Classification Quarterly* 34, no. 1-2 (2002): 235, https://dx.doi.org/10.1300/j104v34n01_14.

[16] A good starting place for those seeking training in authority work is the PCC NACO training site (https://www.loc.gov/aba/pcc/naco/training/). Textbooks like *Maxwell's Guide to Authority Work* (Chicago: American Library Association, 2002) can also provide a basis for the principles and practices that must be learned. Familiarizing oneself with *Descriptive Cataloging Manual* instructions for name and series authority records (DCM Z1) is also recommended (Library of

For Peer Review Only

Congress, "Descriptive Cataloging Manual, Z1," October 5, 2020, https://www.loc.gov/catdir/cpso/dcmz1.pdf). Many workshops and classes are available for catalogers seeking to learn, especially through library associations and consortia and online courses such as those offered through Library Juice Academy.

[17] Sherry L. Vellucci, "Commercial Services for Providing Authority Control: Outsourcing the Process," *Cataloging & Classification Quarterly* 39, no. 1-2 (2004): 445, https://doi.org/10.1300/j104v39n01_12.

[18] Rebecca L. Mugridge, Rebecca Nous, Nancy Poehlmann and Wendy West, "Benchmarking Vended Authority Control Practices in ARL Libraries," *Technical Services Quarterly* 35, no. 4 (2018): 333, https://doi.org/10.1080/07317131.2018.1509432.

[19] Backstage Library Works, "Profile Guide," October 1, 2020, https://ac.bslw.com/community/wiki/index.php/Profile_Guide.

[20] Mike Monaco, "Methods for In-Sourcing Authority Control with MarcEdit, SQL, and Regular Expressions," *Journal of Library Metadata* 20, no. 1 (2020): 1-27, https://doi.org/10.1080/19386389.2019.1703497.

[21] Rebecca L. Mugridge, Rebecca Nous, Nancy Poehlmann and Wendy West, "Benchmarking," 333.

[22] Michelle Sweetser and Alexandra A. A. Orchard, "Are We Coming Together? The Archival Descriptive Landscape and the Roles of Archivist and Cataloger," *The American Archivist* 82, no. 2 (2019): 336, https://doi.org/10.17723/aarc-82-02-18.

[23] Finding aids contain a hierarchical description of the collection, often divided into related records within series and sub-series, and sometimes described down to the item level, allowing

For Peer Review Only

for relationships between materials within the collection to be preserved and contextual information to be demonstrated.

[24] *Describing Archives: A Content Standard,* 2nd ed. (Chicago: Society of American Archivists, 2013).

[25] Jean Dryden, "From Authority Control to Context Control," *Journal of Archival Organization* 5, no. 1/2 (2007): 4, https://doi.org/10.1300/j201v05n01_01.

[26] Aranzazu Lafuente Urién, "Archival authority control: an introduction to Encoded Archival Context for Corporate Bodies, Persons and Families (EAC-CPF)," September 20, 2020, http://www.apex-project.eu/index.php/en/articles/184-archival-authority-control-an-introduction-to-encoded-archival-context-corporate-bodies-persons-and-families

[27] Jeremy Myntti and Nate Cothran, "Authority Control in a Digital Repository: Preparing for Linked Data," *Journal of Library Metadata* 13 (2013): 99, https://doi.org/10.1080/19386389.2013.826061.

[28] Dorothea Salo, "Name Authority Control in Institutional Repositories," *Cataloging & Classification Quarterly* 47, no. 3-4 (2009): 249-261, https://doi.org/10.1080/01639370902737232.

[29] Distinct information elements that are subclasses to more general classes of organized information enabling nonlinear ways of narrowing down search results; see also ABC CLIO, "Facet," *Online Dictionary for Library and Information Science*, October 3, 2020, https://products.abc-clio.com/ODLIS/odlis_f.aspx#facet.

[30] OCLC, "OCLC FAST", August 25, 2020, http://fast.oclc.org.

[31] Library of Congress, "Library of Congress Subject Headings," August 25, 2020, https://id.loc.gov/authorities/subjects.html.

For Peer Review Only

[32] Rebecca J. Dean, "FAST: Development of Simplified Headings for Metadata," *Cataloging & Classification Quarterly* 39, no. 1-2 (2004): 331-52, https://doi.org/10.1300/j104v39n01_03.

[33] OCLC, "VIAF: The Virtual International Authority File," August 25, 2020, https://viaf.org.

[34] Wikimedia Foundation, "Wikidata," August 25, 2020, https://www.wikidata.org.

[35] The next development of the WWW to structure data within web pages so computers can utilize the data in systematic ways; see also W3C, "Building the Web of Data," October 3, 2020, https://www.w3.org/2013/data/.

[36] For an example of this see the entry for William Shakespeare in Wikidata, https://www.wikidata.org/wiki/Q692.

[37] Theo van Veen, "Wikidata," *Information Technology and Libraries* 38, no. 2 (2019): 72-81, https://doi.org/10.6017/ital.v38i2.10886; Nancy Cooey, "Leveraging Wikidata to Enhance Authority Records in the Ehri Portal," *Journal of Library Metadata* 19, no. 1-2 (2019): 83-98, https://doi.org/10.1080/19386389.2019.1589700.

[38] Library of Congress, "Linked Data Service," August 25, 2020, https://id.loc.gov.

[39] ORCID, "ORCID," August 25, 2020, https://orcid.org.

[40] ISNI International Agency, "About ISNI," August 25, 2020, https://isni.org.

[41] IMDb.com, Inc., "IMDb," October 2, 2020, https://www.imdb.com.

[42] MetaBrainz Foundation, "MusicBrainz," October 2, 2020, https://musicbrainz.org.

[43] W3C, "Semantic Web," October 2, 2020, https://www.w3.org/standards/semanticweb/.

[44] W3C, "URI," August 25, 2020, https://www.w3.org/wiki/URI.

[45] Robert Fox, "From Strings to Things." *Digital Library Perspectives* 32, no. 1 (2016): 2-6, https://doi.org/10.1108/dlp-10-2015-0020.

For Peer Review Only

[46] Jane Sandberg, editor, *Ethical Questions in Name Authority Control* (Sacramento, CA: Library Juice Press, 2019).

[47] Cataloging Steering Committee website, August 30, 2020, https://sites.google.com/view/cataloging-ethics/home?authuser=0.

[48] Library of Congress, "Bibliographic Framework Initiative," August 25, 2020, https://www.loc.gov/bibframe/.

[49] Andrew Woods, "Linked Data for Libraries (LD4L) Gateway," October 2, 2020, https://wiki.lyrasis.org/display/ld4lGW.

[50] Lihong Zhu, "The Future of Authority Control: Issues and Trends in the Linked Data Environment," *Journal of Library Metadata* 19, no. 3-4 (2019): 215-38, https://doi.org/10.1080/19386389.2019.1688368.

[51] Michelle Durocher, "URIs in MARC Pilot," August 25, 2020, https://wiki.lyrasis.org/display/pccidmgt/URIs+in+MARC+Pilot.

[52] Michael Beckett, "Wikidata Pilot," August 25, 2020, https://wiki.lyrasis.org/display/pccidmgt/Wikidata+Pilot.

[53] OCLC, "OCLC Shared Entity Management Infrastructure," August 25, 2020, https://www.oclc.org/en/worldcat/oclc-and-linked-data/shared-entity-management-infrastructure.html.

For Peer Review Only

Figures and captions

Figure 1. A library catalog card for *The rhetoric of Western thought* by James L. Golden. The main entry heading for this book is the author's name in the form “Golden, James L.” The main entry heading was the primary means for indexing and searching in a card catalog.

Figure 2. An example personal name authority record for Steve Jobs highlighting key components of an authority record encoded in MARC 21.

Figure 3. An example corporate body authority record for Apple, Inc. The 500 field (or tag) is a "see also" reference to the personal name record for Steve Jobs.

Figure 4. An example bibliographic record for the book *I, Steve*. The highlighted fields show the authorized headings chosen for use in the record, using the forms established in the library catalog's authority database.

Figure 5. A list of the categories of fields used in a MARC authority record. The highlighted heading field is expanded to show the kinds of headings represented by the final two characters in the 1XX tag. Adapted from Library of Congress, “What is a MARC authority record?,” October 5, 2020, http://www.loc.gov/marc/uma/pt1-7.html#pt4.

Figure 6. An example catalog display for a “see from” reference using the variant headings in an authority record. The catalog uses the variant heading to direct the patron to search with the preferred heading.

Figure 7. An example catalog display for a “see also” reference. After finding resources created by the corporate body “Apple, Inc.,” the catalog utilizes the “see also” references in the authority database to suggest related headings (i.e. Jobs, Steve, 1955-2011; Wozniak, Steve, 1950-; Apple Computer, Inc.) leading the patron to additional resources in the library catalog.

For Peer Review Only

Figure 8. A diagram highlighting key areas of the authority control landscape. The diagram lists example libraries, companies, and services associated with these areas. Each area works together to create and manage authority metadata that facilitates information retrieval for patrons in a library's search and discovery systems.

For Peer Review Only

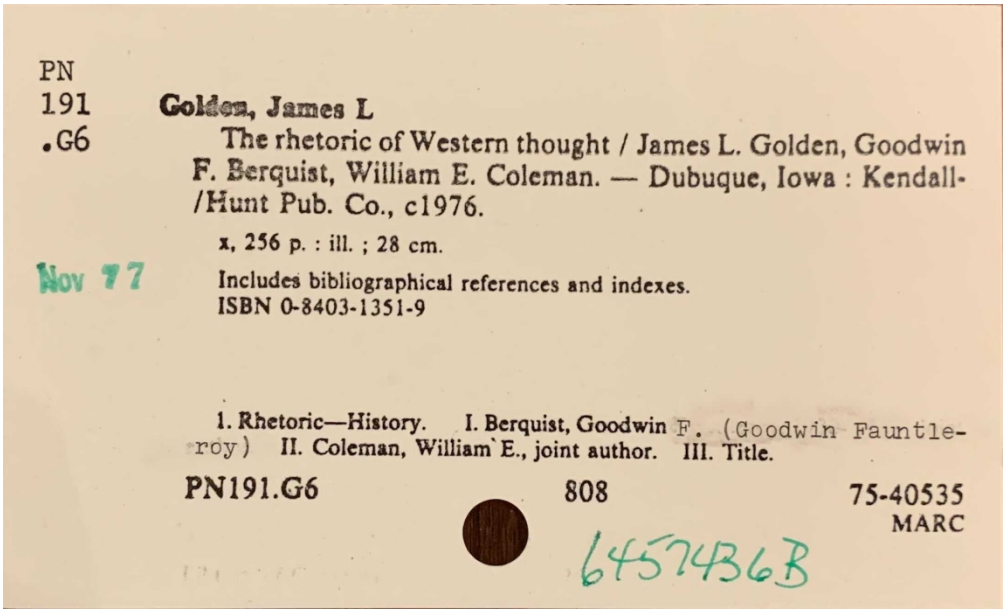

Figure 1. A library catalog card for The rhetoric of Western thought by James L. Golden. The main entry heading for this book is the author's name in the form "Golden, James L." The main entry heading was the primary means for indexing and searching in a card catalog.

177x107mm (300 x 300 DPI)

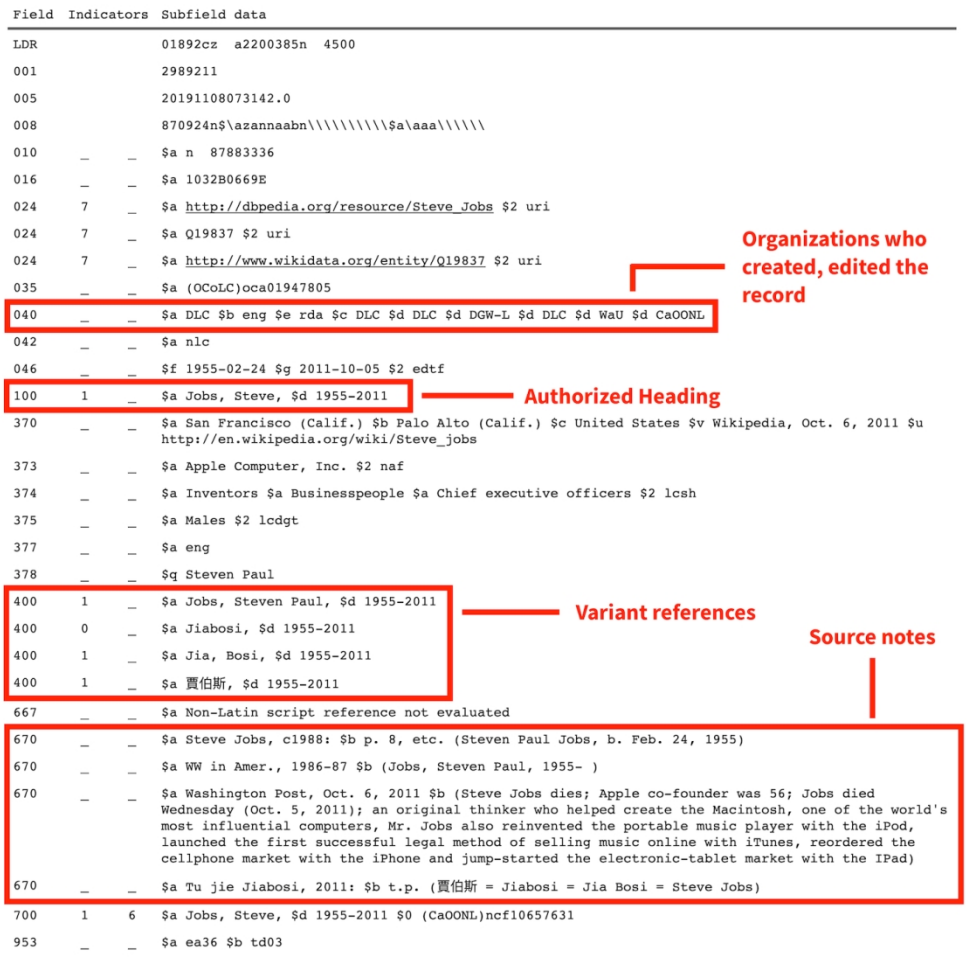

Figure 2. An example personal name authority record for Steve Jobs highlighting key components of an authority record encoded in MARC 21.

127x125mm (300 x 300 DPI)

Field	Indicators	Subfield data
LDR		01845nz a2200241n 4500
001		10979646
005		20190216073347.0
008		190215n$\azannaabn\\\\\\\\\$a\ana\\\\\c
010	_ _	$a no2019024399
035	_ _	$a (OCoLC)oca11821630
040	_ _	$a UPB $b eng $e rda $c UPB
046	_ _	$s 2007-01-09 $2 edtf
110	2 _	$a Apple Inc.
368	_ _	$a Corporations, American $2 lcsh
370	_ _	$e Cupertino (Calif.) $2 naf
371	_ _	$a 1 Apple Park Way $b Cupertino $c CA $e 95014
372	_ _	$a Computer software $a Computers $a Household electronics $a iPhone (Smartphone) $a iPad (Computer) $2 lcsh
377	_ _	$a eng
378	_ _	$q Stephen Richards
410	2 _	$a Apple (Firm)
500	1 _	$w r $i Founder: $a Jobs, Steve, $d 1955-2011
510	2 _	$w r $i Predecessor: $a Apple Computer, Inc.
670	_ _	$a Wikipedia, 13 February 2007 $b (Apple Inc.; American multinational technology company headquartered in Cupertino, California that designs, develops, and sells consumer electronics, computer software, and online services; founded by Steve Jobs, Steve Wozniak, and Ronald Wayne; formerly known as Apple Computer Company (1976-1977) and Apple Computer, Inc. (1977-2007); in January 2007, Jobs renamed the company Apple Inc., reflecting its shifted focus toward consumer electronics, and announced the iPhone, which saw critical acclaim and significant financial success; headquarters: 1 Apple Park Way, Cupertino, California); industry: Computer hardware; Computer software; Consumer electronics; Digital distribution; Semiconductors; Fabless silicon design; Corporate venture capital; products: Macintosh; iPod; iPhone; iPad; Apple Watch; Apple TV; HomePod; macOS; iOS; watchOS; tvOS; iLife; iWork; Final Cut Pro; Logic Pro; GarageBand)
670	_ _	$a Macworld San Francisco 2007 Keynote Address, podcast viewed 14 February 2019 $b (on January 9, 2007 Steve Jobs announced the company name change from Apple Computer, Inc. to Apple Inc.)

Figure 3. An example corporate body authority record for Apple, Inc. The 500 field (or tag) is a "see also" reference to the personal name record for Steve Jobs.

160x127mm (300 x 300 DPI)

Field	Indicators		Subfield data
LDR			01374cam a22003494a 4500
001			16973178
005			20120517151639.0
008			110922s2011 ilu b 000 0deng
010	_	_	$a 2011039328
020	_	_	$a 9781932841664 (pbk.)
020	_	_	$a 1932841660 (pbk.)
020	_	_	$a 9781572846937 (ebook)
020	_	_	$a 1572846933 (ebook)
040	_	_	$a DLC $c DLC $d DLC
042	_	_	$a pcc
043	_	_	$a n-us---
050	0	0	$a HC102.5.J63 $b J63 2011
082	0	0	$a 081 $2 23
100	1	_	$a Jobs, Steve, $d 1955-2011.
245	1	0	$a I, Steve : $b Steve Jobs, in his own words / $c edited by George Beahm.
260	_	_	$a Chicago, Ill. : $b Agate, $c c2011.
300	_	_	$a 160 p. ; $c 18 cm.
504	_	_	$a Includes bibliographical references.
520	_	_	$a "A collection of direct quotes from Steve Jobs on topics related to business, technology, Apple, and life"--Provided by publisher.
600	1	0	$a Jobs, Steve, $d 1955-2011 $v Quotations.
610	2	0	$a Apple Computer, Inc.
650	_	0	$a Businesspeople $z United States $v Quotations.
650	_	0	$a Success in business $v Quotations.
700	1	_	$a Beahm, George W.

Figure 4. An example bibliographic record for the book I, Steve. The highlighted fields show the authorized headings chosen for use in the record, using the forms established in the library catalog's authority database.

154x127mm (300 x 300 DPI)

Field	Definition
0XX	Standard numbers, classification numbers, codes
1XX	Headings (authoritative and reference)
2XX	Complex see references
3XX	Complex see also references
4XX	See from tracings
5XX	See also from tracings
6XX	Notes
7XX	Linking entries
8XX	Alternative graphics
9XX	Reserved for local implementation

X00	Personal names
X10	Corporate names
X11	Meeting names
X30	Uniform titles
X48	Chronological terms
X50	Topical terms
X51	Geographic names
X55	Genre/form terms

Figure 5. A list of the categories of fields used in a MARC authority record. The highlighted heading field is expanded to show the kinds of headings represented by the final two characters in the 1XX tag. Adapted from Library of Congress, "What is a MARC authority record?," October 5, 2020, http://www.loc.gov/marc/uma/pt1-7.html#pt4

170x63mm (300 x 300 DPI)

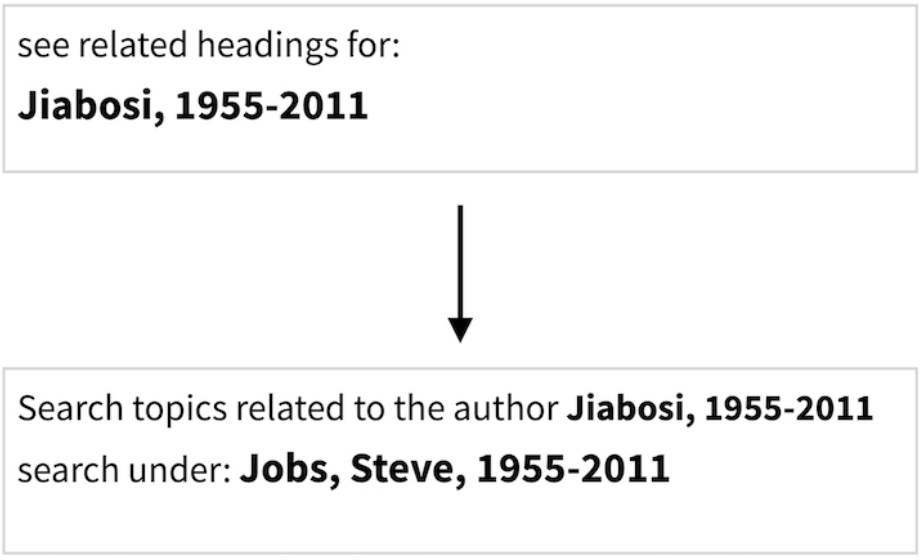

Figure 6. An example catalog display for a "see from" reference using the variant headings in an authority record. The catalog uses the variant heading to direct the patron to search with the preferred heading.

65x42mm (300 x 300 DPI)

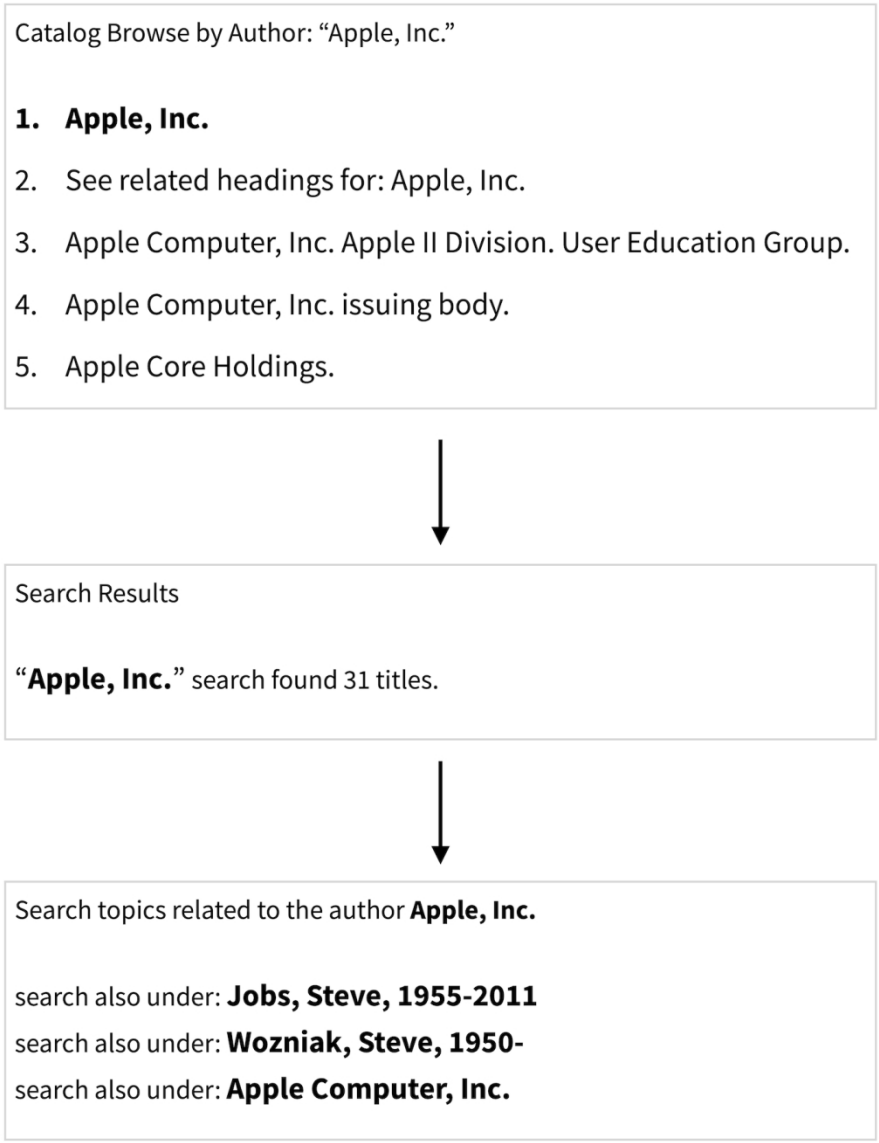

Figure 7. An example catalog display for a "see also" reference. After finding resources created by the corporate body "Apple, Inc.," the catalog utilizes the "see also" references in the authority database to suggest related headings (i.e. Jobs, Steve, 1955-2011; Wozniak, Steve, 1950-; Apple Computer, Inc.) leading the patron to additional resources in the library catalog.

127x162mm (300 x 300 DPI)

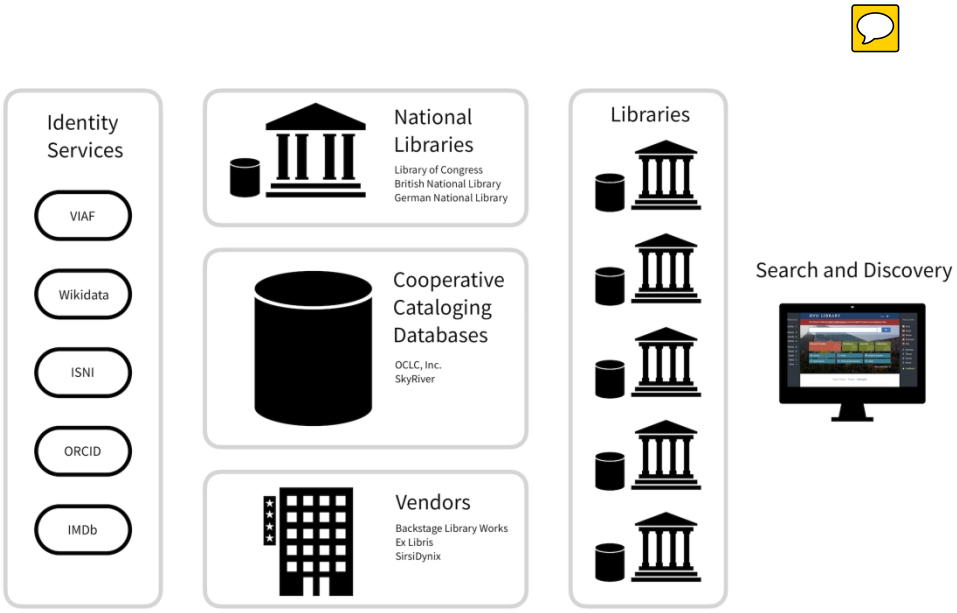

Figure 8. A diagram highlighting key areas of the authority control landscape. The diagram lists example libraries, companies, and services associated with these areas. Each area works together to create and manage authority metadata that facilitates information retrieval for patrons in a library's search and discovery systems.

176x99mm (300 x 300 DPI)