The Problem of Anonymity in Archives: A Literature Review

Arşivlerde Anonimlik Sorunu: Bir Literatür Değerlendirmesi

Emily ROSS*

Abstract

Archivists processing documents rely on factors such as authorship and provenance to contextualize their materials and render them searchable. But in my past experience as an archives user, I repeatedly came across instances of anonymity: letters and diaries by unnamed authors or to unknown recipients, photographs of unknown subjects. In some cases this anonymity is a loss of information that was once there, but in other case it enabled the material to come into existence in the first place: such as in the case of satirical political poetry, for which a writer might face legal censure. In this literature review, the issue of anonymity in the archives is explored, both in a pragmatic sense (recommended strategies for managing it), and a philosophical sense (according anonymous documents the same status as documents with known authors).

Keywords: *Anonymity, Anonymous, Author, Material of unknown origin, Pseudonym*

Öz

Arşivcilikte belgelerin işlenmesi, materyallerin kavramsallaştırılması ve aranabilir hale gelmesinde önemli olan yazarlık ve kaynak (menşe) gibi faktörlere dayanmaktadır. Ancak bir arşiv kullanıcısı olarak deneyimlerimizde defalarca isimsiz yazarlar tarafından bilinmeyen kişilere atfedilmiş mektuplar ve günlükler, bilinmeyen konuların fotoğrafları gibi anonim eserlerle karşılaşmaktayız. Bir zamanlar bu anonimlik bazen bir bilgi kaybı olarak görülürken bazen de belgenin yazarının yasal kınamaya maruz kalabildiği hicivli siyasi bir şiir örneğinde olduğu gibi eserin ilk etapta ortaya çıkmasını sağlamaktaydı. Bu literatür taramasında arşivlerde anonimlik sorunu hem pragmatik anlamda (konunun yönetimi için önerilen stratejiler) hem de felsefi anlamda (anonim belgelere uygun olarak bilinen yazarlı belgelerle aynı statüde) ele alınmaktadır.

Anahtar sözcükler: *Anonimlik, Yazarlar, Aslı bilinmeyen materyaller, Takma ad*

*   Master of Library Studies student (Archives Concentration); Simmons College, USA. (rosem710@gmail.com)

Introduction

Provenance is a foundational archival tenet, prescribing that records should be "maintained according to their origin and not 'intermingled' with those of another provenance" (Roe, 2010, p. 15). Establishing a originator-document relationship can be productive for both entities: knowledge about the originator can help explain the document, while the document can in turn cast new light on the originator (Hurley, 1995, p. 236).

While provenance makes perfect sense as a basis for the physical storage of materials and itemization of them within finding aids, when it comes to cataloging a collection difficulties can arise because provenance does not easily or reliably map onto a single metadata field (title, creator, and so forth). The key piece of provenance data may fall into different fields depending on the particular collection; as records may have originated from a single individual or organization (creator), or have been assembled but not created by an individual or organization (collector), or owned by an individual or organization who neither collected or created them (donor), or other such iterations. When the provenance of material is clearly apparent, such idiosyncrasies can be compensated for by explanatory notes and careful phrasing, such as "Isadora Duncan Collection" (collection originates from Isadora Duncan) versus "Collection about Isadora Duncan" (collection relates to Isadora Duncan but does not originate from her) (RDA Toolkit, 2010). When enough information about the origins of archival materials is available to be fitted into the slots of an electronic catalog often more tailored towards managing books, such glossing over the mismatch between provenance and metadata can be made to work. However, a downside of the centrality of the principle of provenance is that is in less ideal scenarios when aspects of an item or collection's provenance are unknown, cataloguing material in a standardized and useable way is potentially problematic.

As Griffin observed, anonymity "puts in question ... almost all of our ingrained assumptions and procedures for dealing with texts" (Griffin, 1999, p. 466). When cataloging a print text, the absence of a known creator is a mild inconvenience: the creator slot is usually left blank and instead the main entry for a record is its title.[1] However, the recommended practice in archival cataloging is to derive the title of a collection from the name of its originator,[2] meaning that the originator is effectively responsible for both 'author' and 'title' access points, and thus that works with an anonymous creator are left with neither, which makes cataloging them in a standardized way significantly challenging. It should be noted that while electronic cataloging highlights this difficulty, it also alleviates it by creating the possibility for researchers to retrieve anonymous materials from searches on a particular subject, or chronological period, or via a full text keyword search.

---

1   RAD 1.1B3b, in Section 4, Table 3: 'Content Standards with Provisions for Anonymity.'
2   ISAD(G) 3.1.2 (International Council on Archives, 2000).

In this paper, I will discuss issues and scenarios specific to managing a work with an anonymous creator. Depending on the material in question, archivists may catalog at different levels of granularity: a single entry may be the surrogate for vast collection of material or for a single item. Different types of anonymity can exist at any level of granularity, but with vast collections plural anonymous creators may well be involved which adds a level of complexity beyond the scope of this essay. Conversely, anonymous items within a larger collection do not present as much difficulty because they can be included within that collection's broader provenance, so I will not discuss those either. For the purposes of this essay, I will be discussing anonymity at a level of granularity between these two: discrete anonymous items of sufficient size to require their own catalog entry. In the next section (2. What Anonymity is Not) I will discuss some of the assumptions around anonymity, then in section 3 I will itemize types of anonymity, then in section 4 I will summarize relevant content standard instructions for dealing with anonymity, before concluding with my judgements on the matter.

What Anonymity is Not

Given that anonymity is customarily defined in the negative—the lack of an author—it seems appropriate to approach assumptions about anonymity in the negative case.

Anonymity is not just a historical phenomenon

It is erroneous to assume that anonymity is only a problem with historical documents. While various scholars have proposed that anonymous creatorship has petered out since the invention of the printing press or of copyright law (Griffin, 2007, p. 466), this is not actually the case. Both the printing press and copyright law were around in 1830 and yet 80% of novels published that year were done so anonymously (Griffin, 2007, p. 466). While anonymous novels may no longer be in vogue, anonymity is still the norm rather than the exception for many genres of archival materials, such as ephemera (posters, greeting cards), photographs, and memorabilia. As is discussed below, Digital Anonymity is ubiquitous on the internet, and adds another level of challenge to managing and cataloging virtual resources.

Anonymity is not context-free, and names are not context-guaranteed

One of the cited values of provenance is that knowing the creator or collector of work provides insights to the context of that work, which can then aid with interpretation. Certainly having a creator whose background is known provides a verifiable shortcut to the probable dating and culture of origin of materials, but the accuracy of such assumptions is not guaranteed: material may date from only a portion of a creator's life, may have been collected while travelling overseas, etc. Actual dates and contexts of specific pieces can only be ascertained by looking within the collection, and similar searches through content for context can also be used for anonymous materials.

Anonymity is not an indicator of unreliability

Anonymity has a bad reputation. Anonymity has been used as a mask for criminal behavior (such as graffiti) or antisocial behavior (such as cyber bullying), and such associations may be behind the sense of unease people seem to have about anonymous texts:

> Rogers (2002): "today authorship and authority have become inextricably linked … anonymity and pseudonymity have become suspect behavior" (p. 233)

> Robson (2008): "The absence of the name can too easily be taken for an absence of value … it is as if there were something suspicious about the anonymous work. Just as a cheque will bounce if one forgets to sign or date it, so a text is seen to be invalid unless it carries a name" (pp. 353, 362).

In the sciences, blind and double blind studies are believed to *improve* the reliability of information, yet in other fields it is thought that anonymity "diminishes if not thoroughly destroys" the value and credibility of information (Etter-Lewis, 1996, p. 117). While the exact details of the context of a text may not be as *verifiable* as if the creator is known, that doesn't make the text itself any more *unreliable*. Another way of looking at this is to re-examine the assumption that a text is reliable and credible, merely because the author is known (Etter-Lewis, 1996, p.118).

Anonymity is not a stable referent

While anonymity is most typically understood as the absence of a known author, it is not necessarily the surrogate for a name: the creating entity concealed by its presence may be singular, multiple or corporate. Similarly, anonymity is a relevant concept not just for authorial metadata, but for all types of contributors: collectors, photographers, publishers, and so forth. Rather than being a single "known practice with a known and familiar meaning" (Griffin, 1999) there are multiple different types of anonymity, as will be categorized below in Section 3.

Types of Anonymity

While the literature described different anonymities, I could find no taxonomy of types and so have created my own. My intent is to create useful labels for possible causes of/ motivations for anonymity rather than imply that these are fixed categories which do not overlap in individual instances.

Corporate

Corporate anonymity is a bit of a paradox, in that it is the presence of the corporation's name as creator which renders the actual creator(s) anonymous. This is a typical scenario for materials produced in the course of an employee's work for their organization: reports, web pages and so forth. Individual authorial contributions may be concealed "in order to produce the illusion of an omniscient corporate authority" (Buurma, 2007, p. 20). While it is common practice for corporations to be treated as creators for cataloging purposes, it would actually be more accurate to think of them as publishers or copyright holders of works by anonymous creators.

Defensive

A person might choose not to associate their identity with a document if to do so might cause some kind of negative repercussion. Anonymity thus becomes a shield behind which a person in such a scenario is enabled to speak up. Political graffiti, unsigned or pseudonymous letters to the editor or to advice columns, and student evaluation of courses are all forms of defensive anonymity. Anonymity and pseudonyms have also been employed by marginalized group members themselves to enable an "[un]authorized" or "circumscribed" voice to speak, as with women publishing anonymously or under male pseudonyms (Garvey, 2006, p. 159). Through such practices, female authors could avail themselves of privileges and access accorded to their male counterparts that would be denied to them were their gender known. Another example is oral history material. One of the draw cards of oral history as a research strategy is the possibility of getting access to information not otherwise available: that is private, secret, or has been silenced. In such cases, anonymity is a key element in the creation of a "safe space" needed if participants are "to tell the stories that they do not wish to tell openly" without inhibition (Laoire, 2007, p. 385). Gwendolyn Etter-Lewis (1996) interviewed 88 African American women for an oral history project, and the majority would only agree to participate if she guaranteed anonymity (p. 117). It would be ironic if the anonymity that enabled the creation of such types of valuable archival materials was subsequently the cause of them being re-consigned to oblivion.

Digital

As stated above, a common misconception is that anonymity is a problem most likely to be encountered when dealing with historical documents. In fact, anonymity is so ubiquitous amongst digital sources that it is more like the rule than the exception. Users trying to cite websites have long been frustrated by absent metadata and pseudonyms are in widespread usage in Web 2.0. The inability of existing cataloging methods to cope with absent and nonbibliographic metadata was a main motivation behind the development of Dublin Core (Hillmann, 2005). Complicating the issue of digital anonymity is the fact that many multimedia resources are collaborative efforts which may involve designers, programmers, software and so forth which will require some reconceptualization of the notion of 'creator.' As digital materials are comprising

a larger and larger proportion of archival work, more strategies will need to be devised to manage digital anonymity.

Ephemeral

As with digital materials, with ephemera anonymity is the norm rather than the exception. Items such as posters, invitations, greeting cards, tickets, broadsides, postcards and so forth were intended for temporary use and so no effort was made to record the types of metadata useful for cataloging these items in posterity. The nonstandard sizes of ephemeral materials also make them particularly prone to becoming misplaced from their original order which might have provided some context.

Oppressive

While defensive anonymity can empower the voices of creators, anonymity can also be disempowering in scenarios where the identities of individual members of oppressed groups are obscured or suppressed to serve the purposes of their oppressors. When Durba Ghosh attempted to find native women's stories in the archives of colonial India, she found that such women were frequently designated by generic or in some cases derogatory terms, and when their native names were recorded they were often misspelled or mistranscribed (Ghosh, 2004, p. 299). While the language barrier may have contributed to the latter to some degree, Ghosh (2004) argues that the suppression of native women's names was not mere oversight but strategy, and had "material, political, and social consequences" for the women in society (p.300). In the specific instances she studied, the native women were often concubinal partners of white male colonists, and omitting their names from the record denied that relationship any official status and therefore undermined the chances that these women would be able to succeed in claims for financial support for themselves or the children of that union.

In some cases oppressive anonymity goes hand-in-hand with defensive anonymity, as when immigrants to a new culture change their names in order to avoid the disadvantages they might be prone to in a prejudicial society. In India, fair-skinned mixed race children of colonists and native women often "moved up the social scale by erasing any trace of an indigenous past and suppressing their maternal lineage" (Ghosh, 2004, p. 303). While these re-namings are 'chosen' (motivated by the negative consequences that might otherwise be applied), re-namings can also occur under compulsion. For instance, Native American Indians who might previously had multiple names for different life stages or in different circumstances were assigned names and surnames by their colonizers as a step towards assimilation (Scott, Tehranian, & Mathias, 2002, pp. 18, 20, 23). The Western tradition of married women taking on their husband's surname is another instance where the suppression of the names of historically disempowered groups has served to perpetuate patrilineal control. Administrative convenience, whether by bureaucrats or catalogers, is insufficient justification for excluding alternative names and other ways must be found to preserve and acknowledge all aspects of an individual's identity.

Unintentional

Unintentional anonymity is the default consequence of the associations between names and documents being lost. Typical examples include: uncaptioned photographs, which the owner didn't label because they did not see the need to record their knowledge; misplaced items which are no longer attached to the document that would have identified them; or miscellaneous items, often ephemeral, whose provenance and significance is unclear. If the donor or a close relative of theirs is still alive, they might be able to shed some light on the materials, but in other cases the archivist will have to appraise the items and see if their intrinsic worth makes them worth keeping.

Protocols for Cataloging Anonymity

In an ideal world, the problem of anonymity could be solved by some literary detective work. If the donor or a close family member is still alive, they may be able to contribute lost metadata. In some rare cases, enough identifying details are present in the manuscript that it is possible through careful research to reveal the identity of the creator.[3] In other cases, more forensic means may be used. "Diplomatics" is the careful study of the characteristics of a document "by reviewing its text (language, vocabulary and structure), the physical nature of the material the record is written on, the penmanship used, and any other physical clue" in order to make an educated guess of the document's culture of origin and date (McGuire, 2005, p. 30). When the creator of a document is in dispute, researchers might also use orthographic analysis to attempt to make a determination of authorship, comparing diction choices, spelling errors and so forth between a work known to be by a particular author and a work that might possibly have been written by them.[4] However, all these methods are labor intensive and archivists are unlikely to have time to employ them, unless the document in question is of sufficient importance to make establishing authorship a high priority.

Thus, in most cases, the archivist will have to find a way to catalog the anonymity rather than alleviate it. However, when I tried to research how to catalog anonymity, I had trouble finding official guidelines and those that exist are surprisingly vague. In contrast, I found a number of processing manuals written by libraries and archives with their own idiosyncratic rules about how anonymity should be managed: the kind of lack of standardization that content standards were supposedly created to prevent. In order to prevent you becoming as confused as I was, I have summarized the instructions I found in a series of tables: 'Content Standards with no provision for anonymity', 'Examples of institutional processing manuals with provisions for anonymity', and 'Content Standards with provisions for anonymity' (Table I, Table II).

---

3   For example, Fisher (2004).
4   For example, Wells et. al. (2006).

Content Standards with no provision for anonymity

AACR2 (RDA Toolkit, 2010)
ISAD(G) (International Council on Archives, 2000)

Table I. Example of Institutional Processing Manuals with Provisions for Anonymity

Institution	Relevant note	Instruction
University of Southern Mississippi (Bibliographic Services, 2011)	#7_Creator	If the author/creator is unknown, enter [Unknown]. When the last name of the author/creator is unknown, enter the author's first name and [last name unknown].

Other strategies can be deduced from the range of types of names for anonymous manuscripts; which might be devised from the collector, the content, [5] the geographic location,[6] the repository,[7] the copyright holder, the donor, or other distinctive element.

Table II. Content Standards with Provisions for Anonymity

Standard	Relevant note	Instruction
DACS (Society of American Archivists, 2007)	2.3.6	If the name of the creator, assembler, or collector is not known, or if the repository has assembled the materials, do not record a name. [Instead construct a title according to 2.3.18-22]
	2.3.18-20	[Instructions about using specific terms to designate genre of material]
	2.3.21	Optionally, supply a brief term or phrase that most precisely and concisely characterizes the unit being described. The term or phrase should incorporate the form(s) of material that typifies the unit and reflects the function, activity, transaction, subject, individuals, or organizations that were the basis of its creation or use. e.g. Collection of California vacation albums
	2.3.22	[Use phrasing to] clearly indicate that the subject of the collection is not the collector. e.g. Collection on Isadora Dunanc *Collection is about Isadora Duncan, she is not the collector.* Collection of Robert Browning materials *Collection is materials by Robert Browning, he is not the collector.*
Dublin Core (Metatags Company, Inc., 2013)	DC.Creator	[Creator is one of] 15 optional metadata-elements that can be randomly repeated or left out. i.e. When creator is absent, leave the field blank and use other fields to identify the resource.

---

5   For example, Schmid's 'Archiv': collections of anonymous photographs by subject (Durden, 2003).
6   For example, The Finnsburgh Fragment (McGuire, 2005).
7   For example, British Library MS Poet. 5.

Table II continued

Standard	Relevant note	Instruction
RAD (Bureau of Canadian Archivists, 2008)	1.1B3b	Record the name(s) of the person(s), family (families), or corporate body predominantly responsible for the creation of the fonds or collection as a whole. […] if the name of the collector is not known, or if the repository has assembled the material and the archival unit is distinct, do not record a name element. If the repository has assembled the material, and if the archival unit is generic (i.e., likely to be ambiguous outside of a single repository), record the name of the repository. e.g. University of Saskatchewan Archives photograph collection (*not* Photograph collection)
	1.1B3a	If, as instructed in rule 1.1B3b, no name element has been recorded because the name of the collector is unknown or because the repository has assembled the material, include in the nature element an indication of the theme, content, etc. of the collection. e.g. Canadian philatelic dealer's price lists collection If, as instructed in rule 1.1B3b, no name element has been recorded because the name of the collector is unknown or because the repository has assembled the material, and the subject of the collection is the name of the person, express the title of the collection in a way that clearly indicates that the subject of the collection is not the collector. e.g. Collection about Isadora Duncan (*not* Isadora Duncan collection)
RDA (RDA Toolkit, 2010)	6.27.1.8	If the person, family, or corporate body responsible for the work is unknown […] construct the authorized access point representing the work by using the preferred title for the work [see 6.2.2]
	6.2.2.7	If [previous instructions for naming works] do not apply, choose one of these options as the preferred title (in this order of preference): a) a title that has been assigned to the work subsequent to its creation or compilation e.g. Codex Amiatinus b) the name of the manuscript or manuscript group if the work is identified only by that name e.g. Dead Sea scrolls c) a devised title. Use the authorized access point representing the repository (see 11.13.1) followed by Manuscript. Add the repository's designation for the manuscript or manuscript group. If the manuscript is a single item within a collection, add the foliation, if known. e.g. British Library. Manuscript. Arundel 384

Conclusion

Cataloging archives by provenance is an uneasy fit with electronic cataloging by metadata framework. Archival materials with unknown creators make this problem even worse, and yet the guidance on how to manage anonymity in some content standards is absent, inapplicable to all cases, and/or confusing. DACS, RAD and RDA offer some possibilities; and the Dublin Core model of using the information available offers an alternative approach. But it seems to me that what is not being explicitly stated here, is that in the absence of existing author/title labels, form and content become the predominant cues for characterization. Reconsidering this issue from the point of view of a user looking for materials on a particular topic rather than by a particular author, it seems logical for the 'aboutness' of collections to be a highlighted search dimension on the library catalog – and subject searches are able to retrieve materials by anonymous authors equally as well as they are when authors are known. Library catalogs may be an ill-fitting portal for collections with clear provenance, but their capacity for subject searching diminishes the 'handicap' of anonymity and makes materials with unknown creators equally able to be found and used, whichever of the above guidelines an archivist uses to fill in the boxes in the metadata framework.

References

Bibliographic Services. (2011). Best Practices (General Documents). *In The University of Southern Mississippi Libraries.* The University of Southern Mississippi. Retrieved March 19, 2013, from http://www.lib.usm.edu/techserv/metadata/bestpractices/.

Bureau of Canadian Archivists. (2008). *Rules for Archival Description.* The Association of Canadian Archivists. Retrieved March 19, 2013, from http://www.cdncouncilarchives.ca/RAD/RADComplete_July2008.pdf

Buurma, R. S. (2007, Autumn). Anonymity, Corporate Authority, and the Archive: The Production of Authorship in Late-Victorian England. *Victorian Studies,* 15-42.

Durden, M. (2003). Photography, Anonymity, and the Archive: Joachim Schmid. *Parachute,* 109.

Etter-Lewis, G. (1996). Telling from behind her hand: African American women and the process of documenting concealed lives. In R. Josselon (Ed.), *Ethics and Process in the Narrative Study of Lives* (Vol. 4, pp. 114-28). Thousand Oaks: Sage Publications.

Fisher, R. C. (2004). "The Grandmother's Story": Oral Tradition, Family Memory, and a Mysterious Manuscript. *Archivaria,* 57, 107-30.

Garvey, E. G. (2006). Anonymity, Authorship, and Recirculation: A Civil War Episode. *Book History, 9,* 159-78.

Ghosh, D. (2004). Decoding the Nameless: Gender, subjectivity, and historical methodologies in reading the archives of colonial India. In *A new imperial history; culture, identity, and modernity in Britain and the Empire 1660-1840* (pp. 297-316).

Griffin, R. (1999, Autumn). Anonymity and authorship. *New Literary History, 30*(4), 877-95.

Griffin, R. (2007). Working with Anonymity: A Theory of Theory vs. Archive. *Literature Compass, 4*(2), 463-69.

Hillmann, D. (2005, November 7). *Using Dublin Core.* Retrieved March 14, 2013, from http://www.dublincore.org/documents/usageguide/#whatis

Hurley, C. (1995). Problems with provenance. *Archives and Manuscripts, 23*(2), 234-59.

International Council on Archives. (2000). *ISAD(G): General International Standard Archival Description. Committee on Descriptive Standards.* Retrieved March 19, 2013, from http://repositories.lib.utexas.edu/bitstream/handle/2152/14140/ISAD%28G%29.pdf?sequence=2

Laoire, C. N. (2007). To name or not to name: reflections on the use of anonymity in an oral archive of migrant life narratives. *Social & Cultural Geography, 8*(3), 373-90.

McGuire, S. (2005). The Finnsburgh Fragment and provenance: A recordkeeping discussion. *Archifacts,* 27-36.

Metatags Company, Inc. (2013). *Dublin Core Metadata Initiative - DC Metadata Element Set.* Retrieved March 19, 2013, from Metatags.org: http://www.metatags.org/ dublin_core_ metadata_element_set

RDA Toolkit. (2010). RDA Co-publishers. Retrieved March 19, 2013, from http://www.rdatoolkit.org/

Robson, M. (2008,). The Ethics of Anonymity. *The Modern Language Review, 103*(2), 350-63.

Roe, K. D. (2010). *Arranging and Describing Archives and Manuscripts.* Chicago: Society of American Archivists.

Rogers, P. (2002, Spring). Nameless names: Pope, Curll, and the uses of anonymity. *New Literary History, 33*(2), 233-45.

Scott, J. C., Tehranian, J., & Mathias, J. (2002). The Production of Legal Identities Proper to States: The Case of the Permanent Family Surname. *Comparative Studies in Society and History,* 4-44.

Society of American Archivists. (2007). *Describing Archives: A Content Standard.* Retrieved March 19, 2013, from http://files.archivists.org/pubs/DACS-2007.pdf

Wells, S. (2006). *Shakespeare and Co.: Christopher Marlowe, Thomas Dekker, Ben Jonson, Thomas Middleton, John Fletcher, and the Other Players in his Story.* New York: Pantheon Books.