A Study of the Intellectual Structure of Community Archives

Ana Roeschley[1] and Jeonghyun Kim

[1] College of Information, University of North Texas, Denton, TX 76203, USA
AnaRoeshcley@unt.edu

Abstract. Community archives have gained renewed attention as an emerging archival movement. The goal of this study is to map the intellectual structure of community-centric archival research during the 2000-2017 period. To identify and visualize the relationships among topics within the subject areas, we analyzed the co-occurrence index and network structures of terms derived from titles, abstracts, and author-provided keywords in peer-reviewed journal articles and conference proceedings in the field of community archives.

Keywords: Community Archives, Non-institutional Archives, Intellectual Structure, Research Trend, Co-word Analysis, Social Network Analysis

1 Introduction

With rising interest in social history during the second half of the 20th century, the problematic exclusion of underrepresented groups and individuals from the archival record (Caswell, 2014) has been brought to attention. In recent years, scholarship has emerged examining community-led ways to document groups that were underrepresented within or excluded from mainstream archival bodies, such as government, academic, and other institutional archives. Since the emergence of these community archives[1], scholars, archivists, and members of community archives have worked to document the movement. Although existing studies have explored the various aspects of such archives, few have categorized and evaluated these areas of research. This gap has led us to inquire about the nature of the research being produced in the field of community archives. The goal of this study is to answer this research question by mapping the intellectual structure of community archives and exploring directions toward future research.

2 Methodology

This study used co-word analysis and social network analysis techniques to analyze literature on community archives, and to show their intellectual structure and evolution

[1] The scope of what might be defined as a community archive is broad, covering a wide range of different activities and interpretations.

over time. Using co-word analysis, a method that takes into account the frequencies with which the terms have been mentioned (He, 1999), we have attempted to identify the key topics. The topics in the study are derived from words in the titles, abstracts of articles, and author-provided keywords. Social network analysis, the mapping and measuring of relationships among components in a system, was used to structure the terms' network of research on community archives, in which the nodes are the terms while the links represent the co-occurrence of these terms.

2.1. Data Collection

We searched both general and Library and Information Science-related databases, including Library and Information Science Abstracts, Library & Information Science Sources, Academic Search Complete, JSTOR, and Google Scholar. We included conference proceedings papers, peer-reviewed articles, editorials, and introductions to special issues. Book reviews and books were excluded from our data. To aid us in this search, we used the following search terms when searching the databases: community archive(s), non-institutional archive(s), participatory archive(s), grassroots archive(s), activist archive(s), autonomous archive(s), ethnic archive(s), oral history archive(s), and local history project(s). We then imported all articles that fit our criteria into Refworks. This resulted in a total of 366 papers published from 2000 until 2017.

2.2. Data Analysis

Once our data was collected, we exported the publication records in a tab-delimited format and combined texts from titles, keywords, and abstracts. Then the processed text file was imported into QDA Miner where we could determine frequently occurring terms to create a dictionary. Words with similar meanings or those that stemmed from the same cognate were manipulated as synonyms of a single-word entry and abbreviations/acronyms were fixed to their formal version. We then calculated high frequency terms and created a co-occurrence matrix. The visualization map and its network characters were obtained by analyzing the co-occurrence matrix using Gephi. The network structure was analyzed with a community detection algorithm. Additionally, multiple network properties, including network density and centrality, were calculated for the network of words. We performed our analysis over the entire period and also as divided into three periods: 2000-2005, 2006-2011, and 2012-2017.

3 Results

The network was generated based on a term co-occurrence matrix to reflect the relationships among topics in the field of community archives. Figure 1 shows three distinct communities extracted in this study. The relative size of the nodes is proportional to the occurrence frequencies, and the relative size of lines represent the correlation degree among terms.

The largest community, "community archive projects," consists of 251 nodes (38.32%) and 3,784 edges. Here we see terms that illustrate much of the practical work of community archives, including the broader terms "community" and "community archives" in conjunction with terms like "project," "collection," "case study," "experience," and "collaboration." The second-largest community, which can be called "community archives emerging as a distinct area of archival practice," consists of 241 nodes (38.02%) and 2,497 edges, contains the terms "archive," "history," "practice," "archival practice," "activist," "activism," "social movement," "grassroots," and "oral history." This community represents works about the non-institutional archive movement, exploring its roots, emergence, and developments as a professional archival practice. The last community called "the purpose of archives," consists of 105 nodes (16.06%) and 679 edges. The terms "memory," "identity," "process," "record," "power," "creation," "narrative," "cultural memory," "identity," and "participatory archive" are all featured prominently. This community is more about the abstract meaning of community and non-institutional archives for the communities they serve.

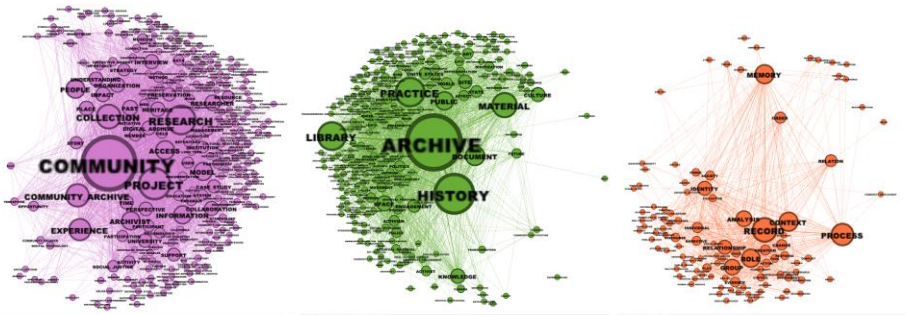

Fig. 1. The three community clusters identified in this study

Figure 2 shows the network structure of the terms represented in each of the three stages: 2000-2005, 2006-2011, and 2012-2017. The network has grown through the periods as the rates of growth for nodes and edges increased between the periods; it is true that the field has rapidly evolved and diversified over time as demonstrated by the increasing number of terms used in publication. It is also evident that the average weighted degree values show a dramatic increase over time (i.e., 23.86 -> 78.82 -> 199.77), which presents stronger and deeper relations among the terms. Modularity and the average clustering co-efficient have decreased over time; this decreased rate of connectedness indicates that newly appeared terms are less cohesive. This also implies that more terms from different groups have been appearing together over time, indicating an increasing trend of interdisciplinary research conducted within the scholarship of those existing independent of mainstream archives.

The terms "archive" and "community" are shown as central terms over the three stages as its degree of centrality and betweenness centrality ranked high. But the use of the term "community archive" greatly increased during the 2006-2011 and 2012-2017 periods (degree of centrality: 10 -> 128 -> 466). During the 2000-2005 period, which includes 19 nodes and 99 edges, the centrality values of "history" ranked high; the term

"history" is a more common term with the use of "historical memory" and "oral history archive," indicating that much of the literature during this period was concerned with the historical importance of records in non-institutional archives. This also implies that much discussion has been floating around oral history collections documenting varied aspects of the local group or community experience during this period.

The 2006-2011 period data, with 148 nodes and 4,068 edges, reveals a growing presence of indigenous perspectives and cultural diversity in the literature as seen in the emergence of terms like "indigenous people" and "queer archive." Further, this period marked the rise of research on non-traditional archival materials as represented by the terms "music archive," "film," "sound archive," and "audiovisual archive." We also observed that the term "community archive" has moved to the core of the network since this period (degree of centrality: 10 -> 128 -> 466).

The 2012-2017 data includes 274 nodes and 19,379 edges. The literature published during this period signifies a growing awareness of the political issues in community and non-institutional archives. This is signaled by terms including "symbolic annihilation" and "social movement." Additionally, the centrality values of the terms that represent more diverse indigenous and ethnic identities, such as "LGBTQ," "black," "woman," "Indian," and "immigrant," were ranked higher than in other periods.

Fig. 2. Network structures for three consecutive periods

4 Conclusion

Due to the nature of community archives as an emerging field of study, we were limited by a small sample size. In identifying three subfields that are not very distinct from each other, we were able to identify community archives as a focused field in its exploratory and theoretical stages. We found that case studies are a common method of research in the scholarship of community archives, but we found little research on assessment or infrastructure of community archives. In addition to this research, there is considerable potential for study on the use and value of community archives.

References

1. Caswell, M. (2014). Toward a survivor-centered approach to human rights archives: Lesson from community-based archives. *Archival Science*, *14*(3/4), 307-322.
2. He, Q. (1999). Knowledge discovery through co-word analysis. *Library Trends*, *48*(1), 133-159.