Investigation and Analysis of Librarian Diversity in Collaborative Research: A Case Study of Nanjing Agricultural University Library (Postprint)

Authors: Tang Huiyan, Chen Rongrong, Zheng Xinyan, Zhang Qian, Cai Zizheng, Liu Yuxian

Date: 2023-08-26T00:00:00+00:00

Abstract

[Purpose/Significance] Librarian diversity and teamwork constitute two important indicators in the ClimateQUALTM organizational climate assessment in the United States; however, the relationship between them has not been thoroughly investigated. This study employs nearly ten years of scientific research collaboration data from Nanjing Agricultural University Library to conduct an in-depth analysis of librarian diversity manifested in collaborative activities, thereby providing a foundation for further understanding the relevant characteristics of library organizational climate.

[Method/Process] By utilizing librarian attribute information—including age, professional title, educational background, disciplinary background, and position—librarians are categorized into distinct groups. The Simpson index (simpson) and entropy value (entropy) are calculated to measure librarian diversity in scientific research collaboration, and collaboration structure diagrams are constructed to visualize the status of research collaboration.

[Results/Conclusion] When individual librarians are treated as categories in diversity calculations, the Simpson index is 0.994 and the entropy value is 0.6. When categories are defined based on librarian attributes such as age, professional title, educational background, disciplinary background, and position, the Simpson index ranges from 0.82-0.997, and the entropy value ranges from 0.59-0.89. These high Simpson indices and high entropy values indicate that librarians at various levels have participated in scientific collaboration, though such collaboration is not highly concentrated. Analysis of network structures formed at certain collaboration intensities further reveals that collaboration remains relatively loose, with no stable collaboration teams having been established. Librarians with advanced educational backgrounds and senior professional titles

represent a significant force in library research output. Although librarians possess diverse disciplinary backgrounds, a background in library and information science remains a crucial factor in the research output of Nanjing Agricultural University Library.

Full Text

Preamble

Volume 62, Issue 11, June 2018

Investigation and Analysis of Librarian Diversity in Collaborative Research: A Case Study of Nanjing Agricultural University Library

Tang Huiyan[1], Chen Rongrong[1], Zheng Xinyan[1], Zhang Qian[1], Cai Zizheng[1], Liu Yuxian[2],[3]

[1]Nanjing Agricultural University Library, Nanjing 210095
[2]Tongji University Library, Shanghai 200092
[3]School of History and Archives, Yunnan University, Kunming 650091

Abstract

[**Purpose/Significance**] Librarian diversity and teamwork are two important scales in the ClimateQUAL™ organizational climate assessment, yet the relationship between them has not been thoroughly investigated. This paper analyzes librarian diversity reflected in collaboration using nearly 10 years of research cooperation data from Nanjing Agricultural University Library, providing evidence for understanding relevant characteristics of library organizational climate. [**Method/Process**] Librarians were categorized based on attributes including age, professional title, education level, disciplinary background, and position. Simpson's index and entropy values were calculated to measure diversity in research collaboration, and collaboration structure diagrams were drawn to display the state of research cooperation. [**Result/Conclusion**] When individual librarians were used as categories in diversity calculations, Simpson's index was 0.994 and entropy was 0.6. When categories were defined by librarian attributes (age, title, education, disciplinary background, position), Simpson's index ranged between 0.82–0.997 and entropy between 0.59–0.89. High Simpson's index and high entropy values indicate that librarians at different levels all participated in scientific collaboration, but the collaboration was not highly concentrated. Analysis of network structures formed by certain collaboration intensities also revealed that cooperation remains relatively loose, without the formation of fixed collaborative teams. Highly educated and senior-title librarians constitute an important force in library research output. Although librarians have broad disciplinary backgrounds, a background in library and information science remains a significant factor in research output at Nanjing Agricultural University Library.

Classification Number: G251.6
Keywords: library; librarian research collaboration; diversity; Simpson's index; entropy
DOI: 10.13266/j.issn.0252-3116.2018.11.007

1. Research Background

Influenced by organizational climate surveys and assessments in American libraries [1-2], domestic research on library organizational climate represented by ClimateQUAL™ began in 2010. Scholars such as Zheng Dejun, Bao Ping, and Tang Huiyan from Nanjing Agricultural University conducted in-depth explorations [3-6]. Based on localized adaptation of the ClimateQUAL™ indicator system, Tang Huiyan et al. initiated three rounds of theoretical and practical exploration at Nanjing Agricultural University Library in 2011, 2014, and 2017. By adjusting management measures according to quantitative indicator data and its changes, they played a positive role in improving library organizational climate and enhancing innovation capacity [7].

To further investigate the impact of various indicators on library organizational climate, this study selected "librarian diversity" and "teamwork" from the ClimateQUAL™ indicator system to examine librarian research collaboration. Research on scientific collaboration in China primarily includes analyses of collaboration in natural sciences, factors influencing research cooperation, and relationships between collaboration and knowledge exchange [8-10], as well as studies on collaboration in library and information science [11]. However, research on diversity in research collaboration remains limited. Harvard University analyzed 2.5 million research papers and concluded that the advantage of research diversity lies in collaboration—papers with diverse authorship and diverse research teams are more attractive [12].

With the globalization of scientific and technological development and increasing complexity of research, collaboration has become a mainstream approach in scientific endeavors. Libraries are institutions that provide information services to readers, and almost all work requires cooperation. Research, which involves challenging the unknown with considerable uncertainty, is also an important way to enhance library service capacity and guide future development. Research collaboration is particularly challenging due to this uncertainty. Most librarians pursue career development through professional titles, and conducting research and publishing articles are prerequisites for promotion. Therefore, librarians have strong motivations to write research papers. During the research and writing process, librarians often collaborate, forming teamwork relationships independent of work relationships. These collaborative relationships exhibit diversity due to differences in librarians' departments, disciplinary backgrounds, titles, education levels, and ages. How does librarian diversity manifest in research collaboration? What characteristics of library organizational climate does this reflect? This study conducts a hierarchical analysis of diversity in librarian and collaborator conditions during research output at Nanjing Agricultural

University Library from 2007–2016 to explore the relationship between librarian diversity and collaborative research.

2. Related Concepts and Research Methods

Diversity is a fundamental attribute and state of existence and development in nature, economy, society, and the human spiritual world. Material forms, geographical environments, biological species, social structures, political systems, industrial divisions, ideological cultures, living habits, behavioral patterns, academic schools, and value systems are all diverse and varied.

Diversity describes characteristics of many systems. Generally, system diversity has three attributes: the number of categories, the balance of quantities across categories, and differences between categories. Over the past 30 years, Nanjing Agricultural University Library has developed rapidly, with its librarian team growing to nearly 80 members, continuously improving educational levels and expanding business scope. The library has nine departments covering three functional areas: campus information construction, modern educational technology, and traditional library services, with significant differences in work content between departments. The evolving nature of work requires continuous upgrading and adjustment of the librarian team. The library is essentially an evolving, complex system where librarians continuously learn, cooperate, and influence each other to complete their mission.

Various methods analyze system diversity. Simpson's index and entropy are two classic diversity indices that involve only the first two attributes (number of categories and their balance). The data in this study makes it difficult to address differences between categories, so we use these two indices to calculate diversity reflected in research collaboration.

2.1 Simpson's Index

Simpson's index is used to calculate species diversity, representing the probability of randomly selecting two different samples from a large community. It measures the number of categories and their balance. The calculation formula is:

$$GM = 1 - \frac{\sum X_{ij}^2}{(\sum X_{ij})^2}$$

where GM is the diversity index and X_{ij} is the i-th collaboration combination. The formula shows that if there is only one collaboration combination, $\sum X_{ij}^2 = (\sum X_{ij})^2$, making the diversity index 0. Therefore, a higher diversity index indicates broader collaboration within a unit, though an excessively high value may also suggest a lack of relatively stable core teams. The distinction between high and excessively high is relative.

2.2 Entropy

The concept of entropy originates from thermodynamics as a measure of system state uncertainty. In information theory, information entropy $S = \sum P_{ij} \ln P_{ij}$ represents system order—the higher the order, the smaller the information entropy and diversity; the higher the disorder, the larger the information entropy and diversity [14]. Zhang Lin et al. studied the relationship between entropy and information distribution dispersion, finding a positive correlation—greater entropy indicates more dispersed information [15]. Similarly, greater entropy indicates greater diversity and more dispersed collaboration.

Assuming we need to calculate the information entropy of librarian collaboration indicators, the initial data matrix includes m samples (collaboration combinations). The calculation formula is:

$$S = -k \sum P_{ij} \ln P_{ij}$$

where $k = \frac{1}{\ln m}$ and $P_{ij} = \frac{X_{ij}}{\sum X_{ij}}$ is the sample value. The k value standardizes the entropy between 0 and 1.

3. Data Collection and Cleaning

3.1 Data Collection

We collected and organized basic data on all research output by librarians at Nanjing Agricultural University Library from 2007–2016, including formally published papers, monographs, patents, and conference papers. Data primarily came from annual internal statistics, supplemented and verified through CNKI and VIP databases. Internal statistics provided the most comprehensive source. We identified 175 research outputs involving library librarians during 2007–2016, including 171 papers, 2 monographs, and 2 conference papers.

For these 175 records, we collected paper/monograph titles, first author names, publication year, whether it was a core journal, first author information, and collaborator details including age, title, education, position, disciplinary background, and relationship to the first author. Table 1 details the collection objects and methods. Information on titles, age, positions, and education levels may have changed during the 10-year period. For incomplete original information, we obtained accurate data through multiple means: consulting authors, database searches, web searches, and calculations.

3.2 Data Cleaning

Some collected data required standardization. For example, age at publication needed uniform calculation methods, department names varied across periods, and disciplinary background names were diverse. During data cleaning, age was calculated as actual age. The library currently has nine departments, but we

also included two historically significant departments: the Information Research Institute and Campus Network Center, due to past reorganizations. Since librarians graduated from different institutions at different times with varying major names, disciplinary backgrounds were standardized according to the 110 first-level disciplines in the "Catalogue of Disciplines for Degree Conferral and Talent Development (2011)" issued by the Academic Degrees Committee of the State Council and Ministry of Education.

After cleaning, we established target matrices based on author names, disciplinary backgrounds, positions, education levels, ages, and titles, then calculated Simpson's index and entropy values. Some librarians are master's or doctoral supervisors who collaborate with graduate students; these students were excluded from statistics.

4. Data and Calculation Results

From 2007–2016, Nanjing Agricultural University Library published an average of 17.5 articles annually. Among 175 research outputs, 155 had library librarians as first authors (88.6%). Details are shown in Table 2 and Figure 1 [Figure 1: see original paper]. Table 2 shows 75 outputs in the first five years and 100 in the latter five years, indicating an upward trend. In terms of quality, there were 63 core journal papers: 27 in the first five years and 36 in the latter five years, showing significant growth. Since 2012, most outputs had library librarians as first authors, including 1 SCI paper, 1 SSCI paper, and 2 EI papers, indicating strengthened research leadership and internationalization in recent years.

Table 2 and Figure 1 also show overall growth in research collaboration. The library's collaborative leadership has strengthened in recent years, evidenced by rising average number of authors per paper and increasing numbers of collaborative papers. Both total authors and library authors show growth trends, with internal collaboration growing more noticeably than external collaboration, which fluctuated. Particularly in the last five years, all outputs had library librarians as first authors, and external collaboration also increased significantly. The library has continuously strengthened inter-departmental cooperation through programs like Librarian University, Librarian Lectures, and academic salons, while also hiring academic mentors to consciously improve librarians' professional quality, enhance organizational climate, and promote more internal and external collaboration.

The 175 papers included 141 total authors: 66 library librarians (46%) and 75 external personnel (53%). There were 44 single-author papers and 131 multi-author papers (75%). Total collaboration instances reached 416, with most papers having 1–3 collaborators and an average of 2.4 collaborators per paper.

For diversity index calculations, data must be divided into categories. First, using individual librarians as categories, we established a collaboration matrix and calculated Simpson's index as 0.994 and entropy as 0.6. Annual matrices

were also created, showing slowly rising Simpson's index and fluctuating entropy (Figure 2 [Figure 2: see original paper]).

We also calculated diversity indices using different librarian attributes (title, age, education, disciplinary background, department) as categories. Table 3 shows the results: Simpson's index ranged 0.82–0.997, standardized entropy 0.59–0.89, and unstandardized entropy 1.82–5.41. Higher Simpson's index and entropy values for age, disciplinary background, and department categories compared to title and education categories indicate more diverse collaboration across these dimensions. While unstandardized entropy and Simpson's index rankings were consistent, standardized entropy rankings were reversed. Further analysis revealed that different attribute quantities produce different $\ln m$ values, causing standardized entropy rankings to change. However, standardized entropy provides a baseline concept for diversity, offering a measurement standard for attribute-specific diversity.

5. Detailed Analysis of Collaboration Patterns

Diversity index analysis shows that librarians of different titles, ages, education levels, disciplinary backgrounds, and departments all participated in research collaboration, but collaboration was not concentrated. What specific patterns exist, and what policy incentives can libraries implement to improve collaboration effectiveness? We examined the collaboration matrices in detail.

5.1 Professional Title

Among 44 single-author papers, 29 were by mid-level librarians and 7 by junior-level librarians (81.8% total). This suggests that junior and mid-level librarians have strong motivations to publish but weaker team awareness or haven't formed effective teams.

Among 131 collaborative papers, authors included 19 junior, 38 mid-level, 33 associate senior, 17 senior librarians, and 31 students (including external collaborators). Library titles have four levels: junior (assistant librarian), mid-level (librarian), associate senior (associate research librarian), and senior (research librarian). Table 4 shows the title-based collaboration matrix. The top three collaboration types are associate senior–associate senior, associate senior–mid-level, and associate senior–senior, indicating that associate senior and mid-level librarians are the main research forces. Although few senior librarians exist and they face no publication pressure, they actively participated in 32 collaborations, helping cultivate research capacity.

The Simpson's index (0.856) and entropy (0.89) for title-based collaboration are relatively high, indicating imbalance across different title levels. Indeed, junior–junior and senior–senior collaborations are relatively rare, while cross-level collaborations are more frequent.

5.2 Age

Nanjing Agricultural University Library staff show an elliptical age distribution, with most librarians aged 31–40, followed by 41–50, and fewer aged over 51 or under 30. We divided ages into 8 five-year brackets from 21–60. Table 5 shows author counts by age bracket.

Table 5 shows that librarians aged 46–50 are most productive, followed by 41–45 and 31–35. Notably, ages 36–40 form a trough. The likely reason: before 35, librarians strive for associate senior titles, publishing more frequently. After 35, some relax after successful promotion or abandon efforts toward senior titles. Ages 41–50 see a new peak as librarians pursue senior titles, which have higher publication requirements and more collaboration. Ages 36–40 represent a prime productive period, making it a key target for research policy incentives.

Table 6 shows the age-based collaboration matrix. The strongest collaboration is between ages 46–50 and 31–35, followed by 41–45 with 36–40, and 46–50 with 41–45. These collaborations likely have internal drivers, as these age brackets represent periods for pursuing associate senior and senior titles.

The Simpson's index for age-based collaboration is 0.953, with standardized entropy 0.65. Interpreting entropy as disorder suggests some orderliness in cross-age collaboration.

5.3 Education Level

Among participating librarians, 3 had associate degrees, 43 bachelor's, 18 master's, and 7 doctoral degrees (some counted multiple times due to degree upgrades). External collaborators included 6 bachelor's, 6 master's, and 20 doctoral degree holders. Table 7 shows the education-level collaboration matrix. Master's and doctoral degree holders are the most important collaborators, far exceeding other combinations, followed by doctoral–bachelor collaborations. Notably, although only 7 librarians held doctoral degrees, they participated in 64 of 131 collaborative articles, demonstrating their strong role in driving research collaboration and output. Master's–bachelor collaborations rank third.

The education-level Simpson's index is 0.82, with standardized entropy 0.79. These similar values suggest that optimizing research capacity requires improving staff educational structure and collaborating more with highly educated external personnel.

5.4 Disciplinary Background

Unlike academic departments, librarians have relatively diverse disciplinary backgrounds, with some holding double or triple majors. Over 10 years, 66 librarians participated in research output across 27 first-level disciplines: 15 in library and information science, 16 in agriculture and related fields (crop science, plant protection, agricultural economics, horticulture), and 35 in other disciplines (education, Chinese language, arts). Table 8 shows the top 5 disciplinary

collaborations. Library and information science collaborations dominate (74 instances), followed by library science–agriculture collaborations, demonstrating that library science background is crucial for research output.

Due to complexity, we didn't provide a full matrix but calculated Simpson's index as 0.997 and standardized entropy as 0.69.

5.5 Departmental Collaboration

We tracked internal authors by department and external collaborators by institution. The library collaborated with 10 external institutions, primarily other university libraries and enterprises, plus 12 campus departments including administrative offices (Research Institute, Personnel Office, Academic Affairs Office, Pukou Campus Library) and related colleges (Information Science, Public Relations, Economics and Management, Humanities and Social Development, Plant Protection).

Table 9 shows the internal departmental collaboration matrix. Most collaboration occurs within the library. The top three intra-departmental collaborations are: Reader Services Department, Reference Department, and Office. Reader Services accounts for nearly half of staff, making its top ranking reasonable. Notably, User Services and Resource Construction had no internal collaboration over 10 years. User Services primarily maintains campus network operations, collaborating mainly with the Information Applications Department. The top three inter-departmental collaborations are: Reference–Office, Development Research–Office, and Reference–Development Research. The Office (including library leadership) collaborates with all departments. Resource Construction and Reference each collaborate with 6 departments, reflecting their global operational significance.

Overall, internal collaboration has increased while external collaboration growth remains limited. While collaborative climate has improved, libraries should emphasize external collaboration, particularly with highly educated and senior-title personnel, to increase research output.

Department-based Simpson's index is 0.961, with standardized entropy 0.59—the largest difference between the two indices among all attributes, warranting further investigation.

We also collected data on collaborator relationships: (1) collaboration with colleagues within/outside the department or other libraries; (2) collaboration with academic advisors or students while pursuing degrees; (3) collaboration with students or faculty as graduate supervisors; (4) collaboration through family relationships (spouses, siblings).

5.6 Collaboration Network Visualization

Using CiteSpace, we generated a collaboration network visualization (Figure 3 [Figure 3: see original paper]). Node size represents author frequency (141

total nodes). The network includes 126 connections with density 0.0128. One large connected network contains 98 nodes (69.5% of total), with maximum collaboration strength 8 and average 1.5. Thirteen nodes collaborated 3+ times, while 81 collaborated 1–2 times. Some small collaboration groups and isolated nodes (6) represent non-collaborating authors.

The five most productive nodes are labeled A–E. Nodes A (28 publications, centrality 0.15) and B (25 publications, centrality 0.48) both hold senior titles and doctoral degrees, demonstrating their research leadership. Node A has two core collaborators (C and E) but limited overall collaboration and has left the library. Node B has high centrality but no core collaborators, with scattered partnerships. Network density, centrality, and collaboration intensity indicators show that research collaboration remains relatively loose, without stable core teams. Further study of this visualization aims to promote healthy collaboration development.

6. Conclusions and Implications

Through in-depth analysis of Nanjing Agricultural University Library's research collaboration from 2007–2016, we find that librarians have diverse disciplinary backgrounds, multi-generational age structures, and various education levels across positions. Diversity analysis shows that collaboration has become a trend for improving research capacity, with multidisciplinary librarians contributing across the board. Associate senior librarians shoulder heavy responsibilities, while doctoral and senior-title librarians lead professional directions.

However, challenges remain: although collaboration is extensive, it lacks focus and hasn't formed core academic teams. To enhance research capacity, libraries must optimize staff structure, increase the proportion of highly educated librarians, and strengthen collaboration with external highly educated and senior-title personnel.

This study demonstrates that diversity analysis methods—including Simpson's index and entropy—help understand and grasp research collaboration status, identify problems, formulate policies, incentivize collaboration, and promote library development.

References

[1] Kyrillidou M, Baughman S. ClimateQUAL: Organizational climate and diversity assessment[J]. College & Research Libraries News, 2009, 70(3): 154-157.
[2] Li X, Bryan L. On becoming an employer of choice: Gauging employee satisfaction through ClimateQUAL?[J]. Library Trends, 2010, 59(1/2): 256-268.
[3] Zhu Suoling, Bao Ping. The development of library organizational climate assessment in the United States and its implications[J]. Library and Information Service, 2016, 60(6): 104-110.
[4] Bao Ping, Zhou Li. ClimateQUAL™ library service quality evaluation sys-

tem[J]. Journal of Academic Libraries, 2010(5): 96-100.
[5] Zheng Dejun, Zhang Zhenghui. Analysis of library organizational climate assessment application: Taking ClimateQUAL™ application as an example[J]. Library and Information Knowledge, 2011, 142(4): 32-39.
[6] Zheng Dejun, Hu Xiaohui. Development of a localized scale for library organizational climate assessment[J]. Library and Information Knowledge, 2013, 154(4): 66-75.
[7] Tang Huiyan, Ni Feng, Chen Rongrong. Exploration of library organizational climate optimization: A case study of Nanjing Agricultural University Library[J]. Library and Information Service, 2016, 60(12): 87-94.
[8] Ai Liangliang. Modern scientific research collaboration from the perspective of Nobel Natural Science Prizes: Cases from 2008–2010[J]. Science and Technology Management Research, 2012, (10): 230-232.
[9] Xu Zhi, Chen Liyu, Wang Sihui. Factors influencing research team collaboration in universities[J]. Science Research Management, 2015, 36(5): 149-161.
[10] Zhao Rongying, Wen Fangfang. Research collaboration and knowledge exchange[J]. Library and Information Service, 2011, 55(20): 6-10, 27.
[11] Wei Ruibin. Research on the status of research collaboration in Chinese library and information science: Based on CSSCI 1998–2004 data[J]. Library and Information Service, 2006, 50(1): 41-43, 48.
[12] Cai Liying. The advantage of research diversity lies in collaboration[J]. World Science, 2014(11): 47-48.
[13] Zhang Junhua. Diversity values and the rise of Zhejiang merchants[J]. Business Economics and Management, 2011(4): 44-49.
[14] Qiu Wanhua. Management Decision-Making and Applied Entropy[M]. Beijing: Mechanical Industry Press, 2002.
[15] Zhang L, Glanzel W, Liang L. Tracing the role of individual journals in a cross-citation network based on different indicators[J]. Scientometrics, 2009, 81(3): 521-838.

Author Contributions

Tang Huiyan: Implemented research and wrote the paper
Chen Rongrong: Calculated Simpson's index and entropy values
Zheng Xinyan: Collected basic statistics, established matrices, and drew collaboration structures
Zhang Qian: Collected basic statistics and established matrices
Cai Zizheng: Collected basic statistics and established matrices
Liu Yuxian: Provided research ideas, verified technical details, analyzed data, revised drafts, addressed reviewer comments, and finalized the manuscript

---

Librarians' Diversity in the Collaboration: A Case Study of Nanjing Agricultural University Library

Tang Huiyan[1], Chen Rongrong[1], Zheng Xinyan[1], Zhang Qian[1], Cai Zizheng[1],

Liu Yuxian[2],[3]

[1]Library of Nanjing Agricultural University, Nanjing 210095
[2]Library of Tongji University, Shanghai 200092
[3]School of History and Archives, Yunnan University, Kunming 650091

Abstract: [Purpose/significance] Librarians' diversity and teamwork are two important scales of ClimateQUAL™. The relation between the two scales has not yet thoroughly investigated. The paper uses the collaboration in the publications of the librarians of Nanjing Agricultural University library during the past 10 years to analyze Librarians' diversity in their collaboration. This analysis will provide evidence for the promotion of organization climate in libraries. [Method/process] The paper catalogs the librarians and their attributions such as age, professional titles, educational background, discipline background, position, etc., and then calculates the Simpson index and Entropy to measure the diversity of librarians. It draws the collaborative network to demonstrate the state of collaboration. [Result/conclusion] Using librarians as catalogues, Simpson index is 0.994, and the entropy is 0.6. Using the attributions of the librarians to catalog, the collaborators' Simpson indices are between 0.82 and 0.997, and entropies are between 0.59 and 0.89. The values of Simpson index are very high, while the entropy is relative low but is still larger than 0.59. The higher diversity may indicate the collaborations have conducted with different catalogues of the librarians but have not yet focused. The paper constructs the collaborative network and analyzes the strong links. It is found that the collaboration is still relatively loose, and the strong team has not been formed. Librarians with high education and high title are the important power to improve performance of scientific collaboration. Although the subject backgrounds of librarians is wide, the library and information science is still the important factor to promote the output of research.

Keywords: library; librarian's collaboration; diversity; Simpson; entropy

Note: Figure translations are in progress. See original paper for figures.