Infographic Representation and Empirical Study of University Library Service Data: Postprint

Authors: Qian Yin, Qu Jianfeng

Date: 2023-08-27T00:00:00+00:00

Abstract

[Purpose/Significance] To integrate and analyze service data generated from the daily operations and services of university libraries, and to visualize it through infographics, thereby enabling users to more clearly and easily comprehend both the service data and its underlying implications.

[Method/Process] After identifying the target audience for the infographics, data is collected from multiple channels, then subjected to cleaning, integration, and analysis to extract the themes to be conveyed, with appropriate visualization techniques employed for infographic creation.

[Results/Conclusion] Practical application demonstrates that effective visual expression facilitates a clearer understanding of the utilization patterns of various library resources, clarifies user demands for different library services, predicts future directions for library operations, and provides an evidential basis for decision-making in library management.

Full Text

Infographic Expression and Empirical Research on University Library Service Data

Qian Yin, Qu Jianfeng Shanghai Jiao Tong University Library, Shanghai 200240

Abstract

[Purpose/Significance] This study integrates and analyzes service data generated from the daily operations and services of university libraries, visualizing it through infographics to enable users to more clearly and easily understand

the data and its underlying significance. [**Method/Process**] After identifying the target audience for the infographics, data was collected from multiple sources, then cleaned, integrated, and analyzed to extract key themes, which were subsequently presented using appropriate visualization techniques. [**Result/Conclusion**] Practice demonstrates that reasonable visualization helps to better understand the utilization of various library resources, clarify reader demands for different library services, predict future directions for library development, and provide a decision-making basis for library operations.

Keywords: university library; service data; infographic; visual expression
Classification: G251
DOI: 10.13266/j.issn.0252-3116.2018.19.005

1 Introduction

As society continues to develop and knowledge updates accelerate, libraries face increasingly higher demands for their services, having fully entered the era of information-based service delivery. In this era, data represents a key asset for library user service capabilities and market competitiveness. Beyond their own resource data, libraries generate massive amounts of data through daily services, including patron visit statistics, book circulation, portal website traffic, electronic resource downloads, and facility usage patterns. This service data conceals valuable information that, if further integrated and analyzed using appropriate visualization methods, can clarify reader service needs, illuminate resource utilization patterns, predict future library development directions, and address how to keep pace with the times by improving data processing capabilities and expanding new avenues for data analysis, knowledge discovery, and knowledge mining—an important issue requiring continuous consideration from library researchers [1].

With the rapid development of various services, university library service data no longer simply refers to the number of books purchased or physical resources borrowed and returned. Rather, it encompasses all types of raw data generated through daily service operations, closely related to patron behavior and constituting an important component of library assets, including circulation data, consultation service data, various system platform data, and social media data. The role of this service data is becoming increasingly prominent, with related research emerging as a hot topic in the library and information science field. University libraries typically compile these statistics to describe, analyze, and investigate various phenomena related to resources and users, thereby revealing the actual conditions, development trends, and general patterns of resources, services, and user behavior to provide decision-making support for services. This paper uses examples from Shanghai Jiao Tong University Library to illustrate how to analyze and integrate library service data.

2 Overview of Service Data and Infographics

2.1 Overview of University Library Service Data

In the big data era, how to develop and utilize data resources and mine their value has become a concern across all sectors of society. As public service institutions and information dissemination centers, libraries must consider how to effectively leverage data assets. Library service data represents a valuable treasure, truthfully demonstrating the entire development process of libraries while fully reflecting their service levels and development directions [12].

2.2 Overview of Infographics

In today's information explosion, people receive information not only through traditional books, newspapers, and magazines, but also from the internet and social media. Reading habits have changed significantly, with more and more people avoiding 枯燥, 乏味, and lengthy texts when consuming information. As the saying goes, "a picture is worth a thousand words," making the value and efficiency of visual communication particularly prominent in this era.

Infographics are widely considered a category of data visualization [2]. As an important method for information extraction and refinement, infographics—with their concise, intuitive, and vivid characteristics—satisfy people's psychological need to obtain effective information more quickly and at lower cost, entering daily life through various channels and achieving explosive popularity. Research shows that on Facebook, infographics are 30 times more likely to be read than pure text content and are shared 12 times more frequently [3]. This advantage is closely related to the physiological characteristics of the human brain; the fundamental reason people prefer reading infographics is that the brain "prefers graphics and is better at processing them" [4].

The general definition of infographics is "a visual representation of data," though they now have a new definition: large-scale graphic design that combines data visualization, illustration, text, and images to tell a complete story [5]. They are used to quickly and accurately express complex information, expanding human visual perception of patterns and trends [6]. This definition comprises two aspects: first, infographics are images transformed from text and data, meaning they are visual representations of data, information, or knowledge; second, infographics are a method and means for explaining data or complex information. Vision is the strongest form of information input and the most powerful way humans perceive the world. The greatest benefit of visual communication is helping people better remember and understand content. Therefore, infographics serve to clarify, effectively, and aesthetically convey information by integrating content planning and visual expression of complex data, information, and knowledge [7].

Since the 20th century, with the development of Flash, HTML, CSS, and other technologies, infographics have taken on richer forms, evolving from early static

infographics to dynamic and interactive versions. This paper focuses primarily on static infographics. Currently, static infographics for university library service data can be categorized by content as follows: (1) infographics expressing numbers and concepts, used to present massive, 枯燥 statistical figures and difficult-to-understand textual concepts; (2) infographics expressing operations and relationships, used to illustrate operational processes, hierarchies, and interrelationships; (3) infographics expressing time, place, and characters, used to chronologically connect events from one or multiple dimensions into a relatively complete record system for presentation; and (4) tag cloud infographics, a search engine-based visualization method that displays tag popularity.

2.3 Current Development of Infographic Expression in Libraries

In 2010, OCLC (Online Computer Library Center) released a comprehensive infographic report titled *How Libraries Stack Up: 2010*, illustrating the role of public libraries in employment assistance, small business development, and family life, reflecting their impact on American economy, society, and culture [9]. In 2013, Canada's National Reading Campaign used infographics as a promotional vehicle, using detailed data and visual representations to demonstrate reading's positive effects on mental health, stress reduction, and emotional intelligence cultivation, creating an intuitive and engaging presentation [10].

In recent years, domestic libraries have also actively explored infographic applications. As early as 2009, Shanghai Jiao Tong University Library began using comics instead of lengthy text to explain borrowing procedures to new students. These comics proved extremely popular, representing an early attempt at service visualization that demonstrated how infographics are more accessible to most people than abstract FAQs or dull instruction manuals. In 2013, Shanghai Library released the *Shanghai Public Library Monthly Report 2012*, the first domestic library reading white paper for the public. In 2014, Xiamen University Library launched a "Taobao-style" infographic titled "Cool Library Data," which was widely shared and praised by users. In 2015, the library held a marketing campaign themed "Time with You," deeply mining patron data to create narrative infographics showing users' first library visits, borrowed books, visit frequency, borrowing volume, and reading preferences during their university years [11]. Shanghai Jiao Tong University Library also releases annual patron reading reports on World Reading Day. Additionally, libraries have begun creating infographics for new services, publishing them on social media platforms like WeChat and Weibo to both promote and instruct users, facilitating sharing among friends and achieving better service promotion results. These examples demonstrate that libraries have embarked on the journey of service and data visualization.

3 Service Data Processing

Data, as a raw product, can be transformed into meaningful information and knowledge through processing, organization, and analysis. As libraries' valuable

asset, service data authentically demonstrates their entire development process while fully reflecting service levels and development directions [12].

3.1 Service Data Collection

When collecting service data, libraries face the primary question of data sources: what data to collect and how to collect it. Data sources mainly include: (1) self-managed system databases, such as patron borrowing data, which are relatively easy to collect and involve various data types (e.g., SQL Server, MySQL, Oracle, Access), requiring export to fixed-format files or import into unified databases; (2) vendor-hosted system databases, where some purchased systems run on company servers using their databases (e.g., APP usage statistics), accessible through company-provided APIs; (3) web data, such as website clicks, visit duration, and user online behavior, requiring web tools and programs for capture and collection; and (4) derived data unavailable from systems, such as legacy paper records requiring digitization, or librarian workload data typically obtained through manual entry or form imports based on work hours and frequency.

3.2 Service Data Cleaning

While pursuing data completeness comes at a cost, massive data increases can lead to inaccurate results and collection of erroneous data. Different sources, collection methods, and data schemas result in metadata containing duplicates, errors, missing values, inconsistencies, irrelevant data, and anomalies. This "dirty data" affects analytical efficiency and must be removed before analysis—a crucial and indispensable process known as data cleaning. Though no universally accepted definition exists, data cleaning generally refers to any processing that helps resolve data quality issues [13]. The cleaning process typically involves five stages [Figure 1: see original paper]: (1) acquiring metadata and analyzing existing problems; (2) defining cleaning rules and determining the cleaning plan; (3) selecting data samples to verify the plan's correctness and efficiency, returning to rule definition if validation fails; (4) performing data cleaning; and (5) replacing original data with cleaned data.

Data cleaning aims to detect and correct both individual data impurities and inter-data inconsistencies, improving data quality [14]. Each collected dataset undergoes usability checks, primarily filtering out disordered and irrelevant data, deleting erroneous values, supplementing missing information, and modifying inconsistent values to retain valuable, operationally strong data as preparation for subsequent integration.

3.3 Service Data Integration and Analysis

To visualize data, one must understand what it expresses—the relationship between data and what it represents is key to visualization, comprehensive analysis, and deep understanding [15]. To fully display data and relationships, inte-

gration is necessary. Single data tables often cannot reflect the full picture of a behavior or phenomenon, requiring combination with data from other sources. This integration involves organizing, consolidating, merging, and completing data integrity.

Data collection, cleaning, and integration are prerequisites for analysis, whose fundamental goal is extracting useful knowledge from data according to needs and applying it to specific domains. Even after integration, service data remains voluminous and structurally complex, with hidden meanings requiring further analysis by librarians. For example, analyzing borrowing data across undergraduate grade levels can reveal how student focus areas change throughout their four-year academic journey. Data analysis can thus be considered a form of mining for data information connotations.

3.4 Automated Implementation of Service Data Processing

Given the complexity and volume of daily library operations and service data, automated processing has important practical significance for resource management, data utilization, and efficiency improvement. Automated processing mainly includes intelligent data collection, automatic integration, and automatic statistical analysis.

Various service data are widely distributed across systems, databases, and networks in multiple types, channels, and formats, characterized by large volumes, dispersed storage, complex types and formats, heterogeneous resource organization, and complex granularity levels. Intelligent collection uses programmed interface calls to regularly fetch data from various systems and sources without manual extraction, with configurable intervals (e.g., weekly or monthly) and preliminary filtering of unnecessary and invalid data for initial cleaning.

Extracted metadata often suffers from inconsistent descriptive syntax, data conflicts, and incompleteness, requiring automated programs for integration. Automatic integration can be understood as information extraction—querying stored data, matching required data fields and entity relationships, achieving semantic associations, extracting information, and storing it centrally in unified formats.

Based on integrated service data, automatic statistical analysis provides users with unified application interfaces according to programmed routines, enabling data summarization, statistics, queries, and report generation to facilitate access to various statistical information.

4 Empirical Cases of Service Data Infographic Expression

While anyone could discover patterns or insights by spending hours examining tables or databases, people are generally unwilling to invest such time in details. They don't want raw data—they want results. Most users believe that "seeing is believing," making data visualization and infographics serve their needs.

Concise and refined infographics help librarians and users quickly understand large amounts of service data and its specific meanings, assisting libraries in determining future service development directions and strategies. They also spark interest, attracting more librarians to participate in creative conceptualization and problem-solving. This section presents cases from Shanghai Jiao Tong University Library.

4.1 Infographic Production

The workflow for visualizing service data through infographics is shown in [Figure 2: see original paper]:

4.1.1 Determine Objectives

First, identify the target audience as library professionals and patrons interested in library services, with the goal of expressing service conditions and effectiveness.

4.1.2 Collect Data and Materials

Library service data is categorized into business, resource, and social media types, with data sources and content described in . Different collection methods are applied according to source characteristics.

4.1.3 Process Data

Data cleaning and integration are performed. Using patron hold request data as an example, metadata from the library integrated management system for January 2017 was extracted. The original table was extensive; shows partial fields and records. The raw table contained unclear field names, unnecessary fields, empty data, and incomprehensible content. Cleaning involved translating field names, removing unnecessary fields like LETTER_{STATUS} (notification status) and LETTER_{DATE} (notification date), deleting empty records, adding record number fields for easier statistics, and correcting STATUS (request status), PICKUP_{LOCATION} (pickup location), and FILTER_{{{SUB}}_{{LIBRARY}}} (holding library) fields for better comprehension. The cleaned data is shown in .

For service data analysis, the content in is insufficient—patron information and bibliographic information may also be needed. Using patron ID to find patron information and request number to locate bibliographic information, three tables (patron hold requests, patron information, and bibliographic information) can be integrated (possibly requiring more tables), as shown in .

4.1.4 Extract Themes

Integrated data is analyzed and statistics generated using tools like Excel to extract themes based on objectives. These statistical results inform decision-making for book purchasing, themed exhibitions, and departmental outreach. Using hold request data as an example, themes extracted include total request volume, cross-campus requests, most requested books, and most requested cross-campus books.

4.1.5 Design and Implement Infographics

Infographics are designed and implemented based on extracted themes. [Figure 3: see original paper] clearly shows the overall hold request situation, available pickup locations, most popular requested books, and cross-campus request patterns. Designing effective infographics is often more challenging than anticipated, but certain strategies can help:

1. **Determine overall style and color schemes.**
2. **Ensure precision**—the most critical design element. For example, pie chart sections must be drawn strictly proportionally, summing to exactly 100%.
3. **Choose fresh topics** that interest users and provide new, unexpected information. Shanghai Jiao Tong University Library serves millions of patrons annually, yet most are unaware of the library's annual consumption of water, electricity, and even toilet paper. [Figure 5: see original paper] visualizes this data using familiar campus landmarks for comparison, making it both novel and environmentally consciousness-raising.
4. **Maximize visualization** for quick readability. [Figure 6: see original paper] displays homepage visit data graphically, making overall patterns easier to grasp than tabular formats. Monthly statistics are shown in .
5. **Minimize text and highlight key points.** Infographics should convey main messages within seconds. [Figure 7: see original paper] effectively communicates that interlibrary loan uses various transportation methods to deliver unavailable books while showing success rates.
6. **Borrow successful precedents.** Proven visualization schemes like word clouds effectively display keywords. [Figure 8: see original paper] shows 2017 high-frequency search terms from the library's discovery system (Siyuan Exploration) as a tag cloud.
7. **Add 趣味性** through comics, classic quotes, or trendy expressions, especially for younger audiences [16].

4.1.7 Verify Infographic Content

Accurate, unambiguous expression is fundamental. Ensure all data and information are authentic and accurate before publication.

4.2 Infographic Publication

Completed infographics can be shared through multiple channels including blogs, social networks, WeChat, and print media. Shanghai Jiao Tong University Library primarily uses WeChat and Weibo due to easy content sharing and multimedia presentation. As of December 2017, the library's official WeChat account had 7,684 followers and its Weibo account had 9,849 followers, providing substantial audiences for infographic dissemination. Visualizing service data through infographics helps patrons better understand library services and has received widespread praise.

5 Conclusion and Outlook

As a valuable library asset, service data can fully express the intrinsic value of library services and guide future development directions through integration and analysis. Infographics, as powerful tools for effective information communication in the internet and big data era, will undoubtedly evolve into new media forms.

Visualizing and disseminating library service data through infographics offers many advantages but also has limitations. If infographic creators cannot correctly understand the data they aim to communicate, they may mislead users. Data errors and misinterpretation can reduce user trust and negatively impact library branding. Therefore, how to maximize strengths while minimizing weaknesses to effectively visualize service data remains a long-term research topic for librarians.

References

[1] Gao Shan. Library user service innovation in the big data era[J]. *Shanghai University Library and Information Work Research*, 2014, 24(4): 45-49.
[2] SHEN H P. Exploration on the data visualization based on the case of infographics[J]. *Advanced Materials Research*, 2014, 909: 375-378.
[3] ZARRELLA D. How to get more likes, comments and shares on facebook[EB/OL]. [2018-02-02]. http://danzarrella.com/infographic-how-to-get-more-likes-comments-and-shares-on-facebook.html.
[4] KRUM R. *Cool Infographics: Effective Communication with Data Visualization and Design*[M]. New Jersey: Wiley, 2014: 15.
[5] KRUM R. *Visual Communication*[M]. Translated by TANG Qin. Beijing: Publishing House of Electronics Industry, 2014: 6.
[6] Infographics-Wikipedia[EB/OL]. [2018-02-12]. http://en.wikipedia.org/wiki/Infographic.
[7] Huang Jing. Practice and reflection on infographic application in libraries[J]. *Information and Documentation Services*, 2014(4): 90-95.
[8] Zhang Bei, Jin Yongcheng. Research on infographics under social media background[J]. *Journalism & Communication*, 2015(19): 19-21.
[9] Yang Jia, Zhao Liang. Review of infographic applications in libraries[C]//Shanghai Library. *Smart City and Library Services—Proceedings of the 6th Shanghai International Library Forum*. Shanghai: Shanghai Scientific and Technological Literature Press, 2012: 124-129.
[10] National Reading Campaign. Infographics: Readers Save the World[EB/OL]. [2018-02-23]. http://www.nationalreadingcampaign.ca/reading-the-facts-the-infographic/.
[11] Wu Meng, Hong Yue. Research on library content marketing practice and development[J]. *Library & Information*, 2016(2): 53-57.
[12] Sun Hui. Research on library data sources and collection in big data environment[J]. *Intelligent Computer and Applications*, 2017, 7(5): 140-141, 145.
[13] Ye Ou, Zhang Jing, Li Junhuai. Review of Chinese data cleaning research[J]. *Computer Engineering and Applications*, 2012, 48(14): 121-129.

[14] Jiang Xun, Liu Xiwen. Research on data cleaning for knowledge services in big data environment[J]. *Library & Information*, 2013(5): 16-21.
[15] Zhou Su, Wang Wen. *Big Data and Its Visualization*[M]. Beijing: China Railway Publishing House, 2016: 23.
[16] Hou Zhijiang. Application and practice of infographics in library services[J]. *Library Development*, 2015(7): 74-77, 86.

Author Contributions

Qian Yin: Designed the paper framework, wrote and revised the manuscript;
Qu Jianfeng: Provided guidance and revision for the paper.

Infographic Expression and Empirical Research on Service Data of University Libraries

Qian Yin, Qu Jianfeng
Shanghai Jiao Tong University Library, Shanghai 200240

Abstract: [Purpose/significance] This paper aims to integrate and analyze the service data produced in daily work and service of the university library, and use infographic to visualize these service data, so that users can understand the service data and the meaning behind the data more clearly and easily. [Method/process] After determining the audience of the infographic, data is collected from various ways, then cleaned, integrated, and analyzed, the theme to be expressed is extracted, and the infographic is drawn in an appropriate way. [Result/conclusion] Practice shows that reasonable visualization helps to understand the utilization of various resources in the library more clearly, clarify readers' demand for various services of the library, predict the direction of future work of the library, and provide decision-making basis for the work of the library.

Keywords: university library; service data; infographic; visual expression

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