

Research on the Annual Reading Report of Academic Libraries Based on Personas

Lan Du, Ziqi Wang

Jinan University Library, Guangzhou, China

Email: dulan003@163.com

How to cite this paper: Du, L. and Wang, Z.Q. (2018) Research on the Annual Reading Report of Academic Libraries Based on Personas. *Open Journal of Social Sciences*, 6, 95-105.

<https://doi.org/10.4236/jss.2018.610009>

Received: October 9, 2018

Accepted: October 22, 2018

Published: October 25, 2018

Copyright © 2018 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

This paper expounds relevant theoretical research on personas and the advantages of the annual reading report of academic college library based on personas, and expounds the construction of personas in detail from four aspects: data collection, behavior modeling, construction personas and visual presentation. The annual reading report of the academic college library with personas digs in depth all kinds of reading data in the library, constructs personas model by using a set of tags, objectively understands the reader's reading tendency, and facilitates librarians to carry out accurate reading promotion services.

Keywords

Personas, Reading Report, Academic College Library, Reading Promotion

1. Introduction

In recent years, in the context of the era of big data, libraries have creatively applied technologies such as big data and artificial intelligence to explore the potential needs of readers, established a multi-dimensional interaction model between libraries and readers to provide readers with more personalized and humanized services. At present, more and more libraries publish the annual reading report through their official WeChat public accounts or the library homepages. The author takes the “two-class” (the world first-class universities and first-class disciplines, referred to “double-class”) academic libraries announced in China in September 2017 as an example, investigates and analyzes the 2017 library reading annual reports issued by the 42 academic libraries. According to the survey results, there are 29 academic libraries publishing 2017 library reading reports through their official WeChat public accounts or the library home-

pages, accounting for 69.05% of the total. The annual reading report reflects the library's literature resources utilization and space services in the past year. The library borrows the annual list of books, analyzes the library's reader data, and reveals various types of data in a simple and vivid way. The reader data shows characteristics of multi-sources, massive, isomerization, and rapid change. By constructing the annual reading report of the academic college library based on personas, the depth of the various reading data of the library is mined, and the collection of labels is used to construct the personas. The model can intuitively understand the readers' reading tendency and facilitate the librarians to carry out accurate reading promotion services.

Personas are based on massive data, extracting the overall information related to the users, including inherent attributes, such as users' names, ages, genders, and the users' reading habits. Such information provides the basis for further analyzing the users' behavior habits by big data and for more accurate users targeting and personalized services [1]. The concept of personas was first proposed by Alan Cooper (father of interaction design)—“Personas are a concrete representation of target users” [2]. It means, persona is a virtual representation of the real user. There are more researches on personas from foreign libraries than from China, and the foreign application gradually matures. On July 6th 2018, the authors of this article searched for topics in the categories of “user profile” or “Persona” by the Summon Foreign Language Resource Discovery System. The searched subject area was limited to library and information science. A total of 102 articles were detected. A total of 70 articles are in line with the subject after filtering according to the contents. It was found that the research interests of these articles focused mainly on the theories, concepts, algorithms, techniques, personas modeling, practices and applications. For example, Teixeira C, Pinto J S, Martins J A provide a new approach about how user modelling applications can be integrated into any organization without the cost of reengineering the entire information system already in place [3]. Silvia Rossi, François Ferland, Adriana Tapus find that the current literature introduces a general classification scheme for both the profiling and the behavioral adaptation research topics in terms of physical, cognitive, and social interaction viewpoints [4].

In practical applications, personas first appeared in the field of foreign libraries in the mid-1980s, and were applied in the British National Bibliography and Blaise-line (one of the first online services in Europe) for service optimization [5]. In the American University Library, the annual report of the library is also an important part of library management. The earliest library to release the annual report was the Yale University Library, which has published its annual reports since 1899. The main contents of the American University Library annual report includes: the letter written by the curator to the readers, the mission statements, major achievements, annual hotspots, exhibition activities, financial work, statistics, gifts and donations, etc. [6].

Through literature research, there are few studies on personas in the domestic

library field in China. In terms of theoretical research, personas and labeling systems are widely used in data statistics, data analysis, and big data. With the development of smart libraries and the popularization of artificial intelligence, in the past two years, personas have gradually become a hot issue in the library research field. Through the CNKI database, “user profile” or “Persona” was used as a keyword for searching and 11 related articles were detected. The publication dates have been centered between 2017-2018. The research and application fields mainly focused on the following two modes: construct a complete reader portrait model or user behavior model through data collection and completion; establish a persona (or user behavior model) for library personalized service or library precision service, focusing on how to use the persona or user behavior model to serve the library or reader. In addition, the author also surveyed the 2016-2018 Annual Project of the National Social Science Fund of China (including key projects, youth projects and general projects) issued by the National Office of Philosophy and Social Science Planning. The research topic on “Persona” has only one hit in 2018. The theme of the project is “Research on Library User Portraits of Multi-source Heterogeneous Data Fusion”. Therefore, current research on personas for the annual reading report has not been well investigated in China.

In terms of practical application, through the massive survey of the 2017 reading reports published by the university libraries, it is found that in the “double top-class” university libraries in China, only the reading reports from Peking University Library [7] and the Sun Yat-sen University Library [8] revealed the readers’ data through personas. The reading reports found different reading tendencies of different readers through their body physical characteristics and found that there was a positive correlation between font size and reader preference.

2. Advantages

Supported by big data technology, artificial intelligence and other technologies, with the goal of intelligent services, the advantages of applying personas to academic library annual reading reports are as follows:

First, reveal the content of the reading report through visual effects. Professor Huanwen Cheng, the director of the library of Sun Yat-sen University, mentioned in a report during the 2018 China Academic Library Development Forum: “Reading promotion is fundamental for the construction of a first-class university.” [9] Academic college library reading report acts as a beacon for academic college reading. The report can help accurately understanding the reading status of the university staff and students as well as help carrying out reading promotion according to the factual data and actively promoting the construction of a reading atmosphere for the scholarly campus. At present, the reader behavior data and library resource data reflected in the annual reading report of China's libraries are relatively planar. The user images appeared in the annual reading report were mostly reflected in the popular books, authors, etc., which

were relatively simple. Further analysis could be from different angles, subdividing user group behavior, fully displaying the library's annual data from user personas, resource personas, etc.

Second, by constructing a persona label system, deep mining and profiling of reading data can be performed. For personas, the higher “pixel” can better reflect the user data. That is to say, the higher the pixel, the easier it is to carry out library accurate services. In the process of constructing the label system of the annual reading report personas, the library should deeply explore the readers’ personal basic data and the readers’ reading behavior data, etc., as well as pay attention to the attachment and association of the readers with the resource data after modeling, and carries out category management, cross-analysis, and iterative analysis for the multi-layer data labels. Only by these means, the library can generate a complete data labeling system to reflect reading report data in a much deeper level.

Third, according to the results of personas in the reading report, it is convenient to carry out personalized, active and personalized reading promotion services. At present, most library reading reports lack the reading-trend analysis of the readers. According to the author's investigation, taking the “double-class” university library as an example, the content of the reading report data displayed by each college is different. The common annual reading reports mainly include the library's admission volume, borrowing volume, resource construction and utilization, space utilization, various types of reader service and reading promotion statistics, book ranking data and various reader star data, WeChat statistics, etc., the data mining of most reading reports stays in revealing the amount of admissions, book borrowing and other content, lack of deep mining of data. The readers’ reading tendency can be deeply analyzed through the construction of various groups of personas. This could help revealing the readers' reading tendency and reading behaviors as well as help analyzing problems emerge during the reading promotion process. Thus, applying personas to academic library annual reading reports can better guide reading, help carrying out targeted reading activity and increase participation rates during the library reading promotion activities.

3. Construction

To introduce personas that are widely used in the commercial field into the annual reading report of the library and present the reading tendency of the readers by visual means, it is necessary to use the technical equipment, such as big data and artificial intelligence, to capture the readers’ reading data and construct a persona model. By these means, the library can accurately understand the readers’ needs for information, provide more accurate reading promotion services and provide different service strategies for readers under different models. The construction of personas model mainly includes data collection, construction of label system and data model, iterative analysis of label system, generation

of portraits, etc. The users' portraits are used to construct the reading tendency of typical groups of colleges. The process is shown in **Figure 1**.

3.1. Data Collection

The authors conducted an internet survey on the 2017 annual reading report published by Chinese university libraries. From the common annual reading report, data such as the libraries' admission volumes, borrowing volumes, resource constructions and utilizations, space utilizations, various types of reader service volumes (such as training lectures, subject services, etc.), reading promotion statistics, book ranking data, various reader ranking data and WeChat statistics etc., were collected. Books' borrowing volume, admissions volume and rankings constitute the main contents of the reading reports of the colleges and universities. According to the readers' needs from the reading report, combined with the characteristics of the users' portraits, the relevant data for constructing the user portraits of the library annual reading report is preliminarily compiled (as shown in **Figure 2**). These data can be accessed through the library integrated management system and access control. The system and the library business statistics platform and other multi-channel acquisitions, according to the various types of data required, in the corresponding statistical platform to obtain data.

3.2. Behavior Modeling

User image data can generally be grouped into two dimensions, namely static image data and dynamic image data. In the static portrait data of the annual reading report of the library, one is resource metadata (including resource attribute information) and related data of resource utilization, such as popular books, publishers, authors, borrowing methods, etc.; second is, reader attribute

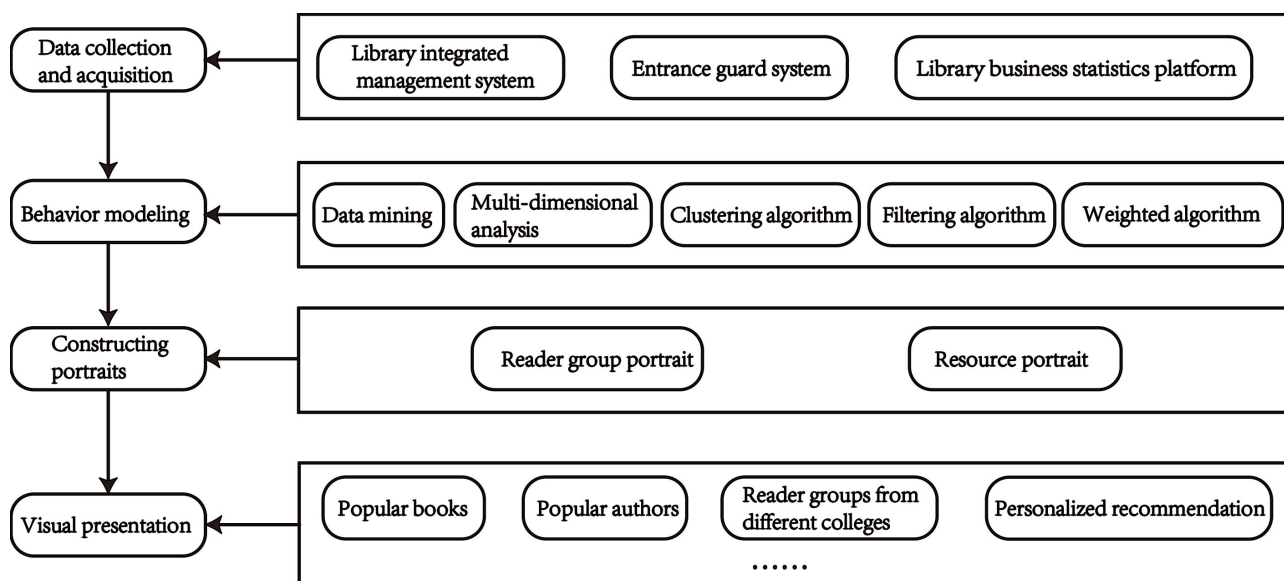


Figure 1. Construction flowchart of the user portrait of the library reading report.

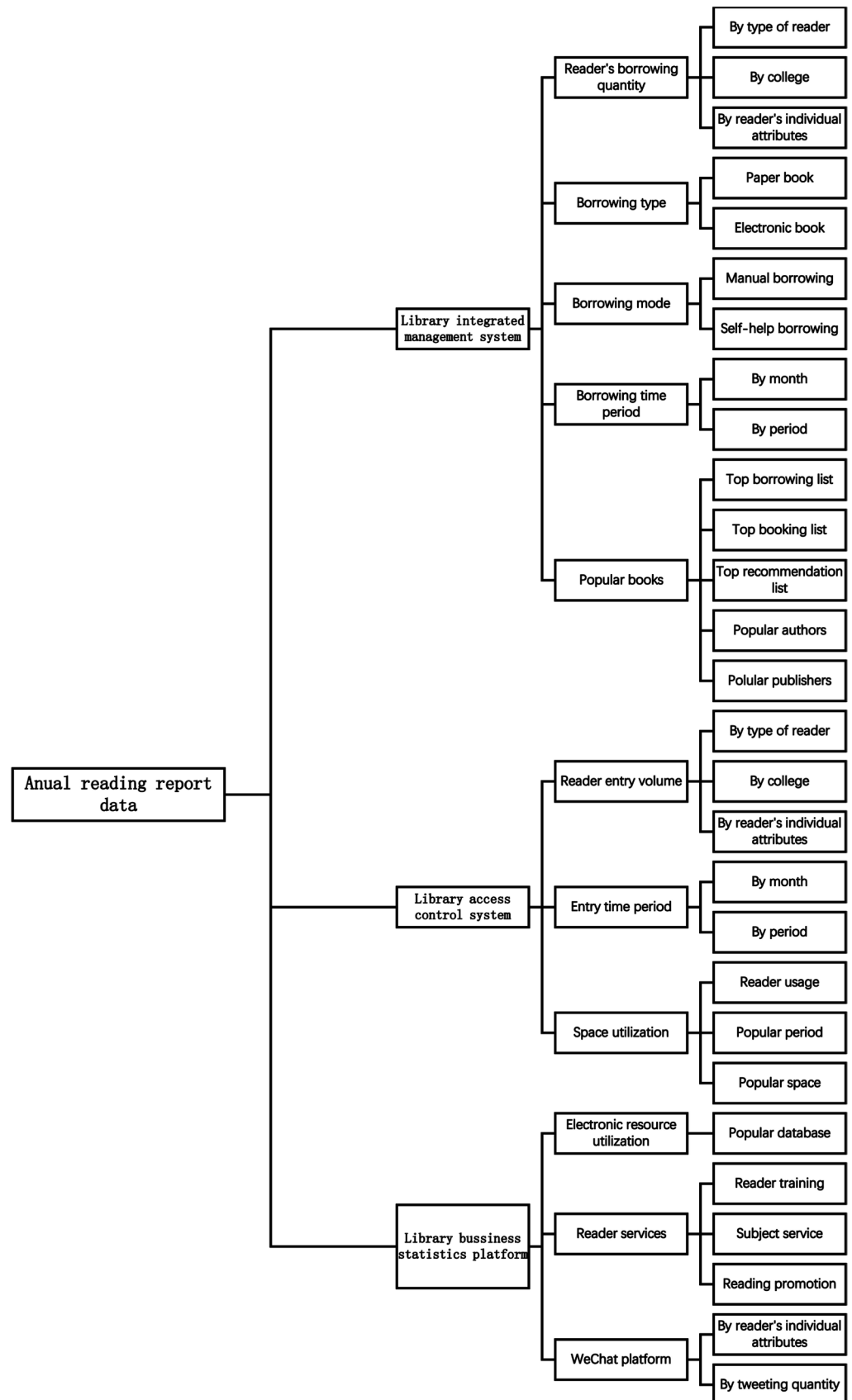


Figure 2. Constructing the user portraits of the annual reading report of the library.

data, such as name, gender, age, etc. The dynamic portrait data of the annual reading report of the library is the readers' behavior data, such as the amount of borrowing, the amount of admission, the amount of WeChat subscription, the number of site hits, and the time period of the reader's use of the library, as shown in **Figure 3**.

Through the deep excavation of the library reader data, extraction of the static portrait data and dynamic portrait data, construction of models by clustering algorithm, filtering algorithm, feedback algorithm and weighted algorithm etc., libraries can conduct iterative analysis of the label system. By further data filtering, elimination of redundant data and analysis of user tags, the user behavioral modeling can be completed.

3.3. Build Persona

The reader persona in the annual reading report of the library is a persona of the reader group. It is necessary to analyze the reader group model through multi-dimensional analysis, and according to the reader behavior data, all readers are grouped by associated clustering to form different types of reader group personas or resource personas. The persona of a reader group was subdivided and was used as an example as in **Figure 4**.

After completing the model building of the reader group, the personas are constructed according to the weights of the reader behavior tags. The construction, improvement and application of personas are inseparable from the support of algorithms and technologies, such as classification algorithms, filtering algorithms, feedback algorithms, weighting algorithms, etc. [10]. The common algorithm in the classification of persona weights is the TF-IDF algorithm, which is a

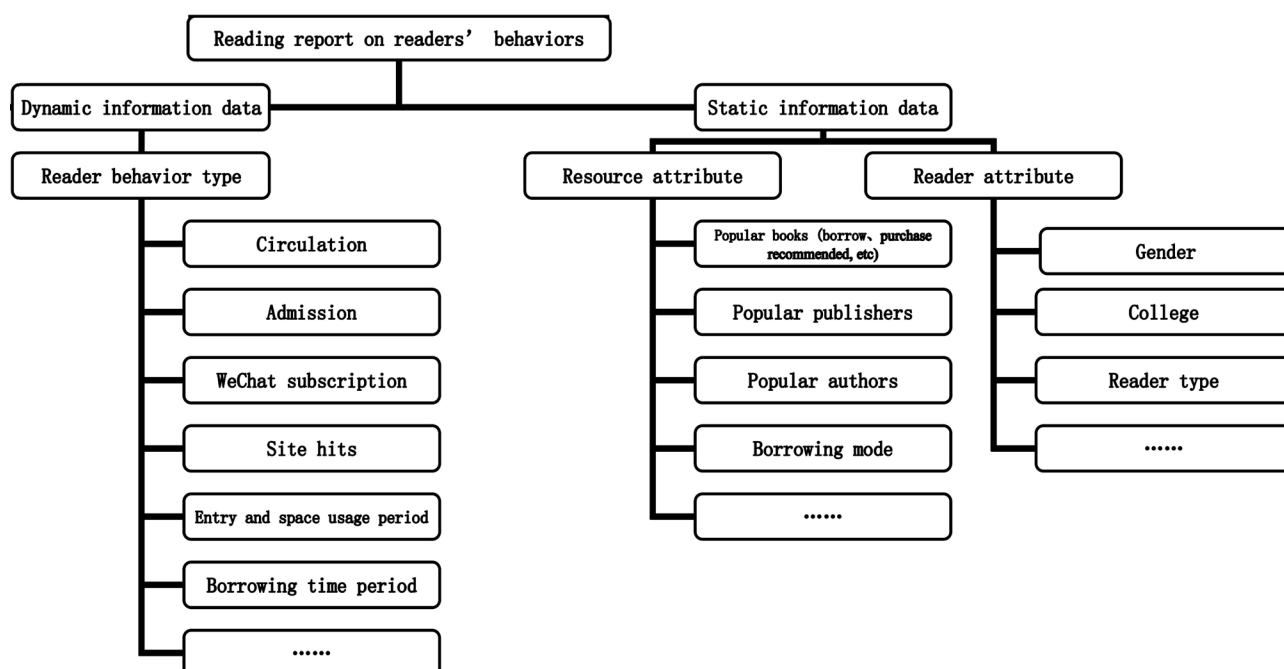


Figure 3. Reader behavior modeling in the library annual report.

statistical method used to evaluate the importance of a word to a file set or one of the files in a corpus. The importance of the word increases proportionally with the number of times it appears in the file, but at the same time it decreases inversely with the frequency it appears in the corpus [11]. Here we use $W(P, T)$ to indicate the number of times a tag T is used to mark the user P , and $TF(P, T)$ to indicate the proportion of the tag number in all tags of the user P , as follows:

$$TF(P, T) = \frac{W(P, T)}{\sum W(P, T_i)}$$

Note: $T_i \in$ all tags of the reader or resource.

The corresponding $IDF(P, T)$ indicates the scarcity of the tag T in all tags, that is, the probability of occurrence of this tag. The formula is as follows:

$$IDF(P, T) = \frac{\sum W(P_j, T_i)}{\sum W(P_j, T)}$$

Note: $P_j \in$ all tags of the reader or resource.

Then, according to the $TF * IDF$, the weight value of the reader group label can be obtained. Finally, a persona of a specific group can be outlined, and the associated data is presented in the form of a persona, as shown in Figure 5.

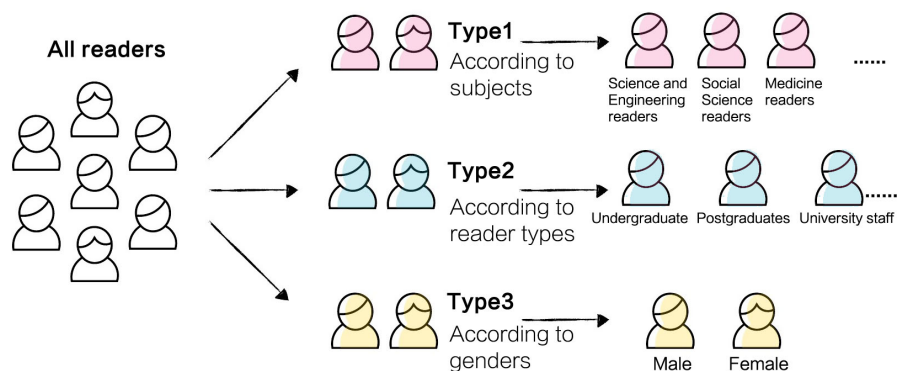


Figure 4. Reader group subdivision model.

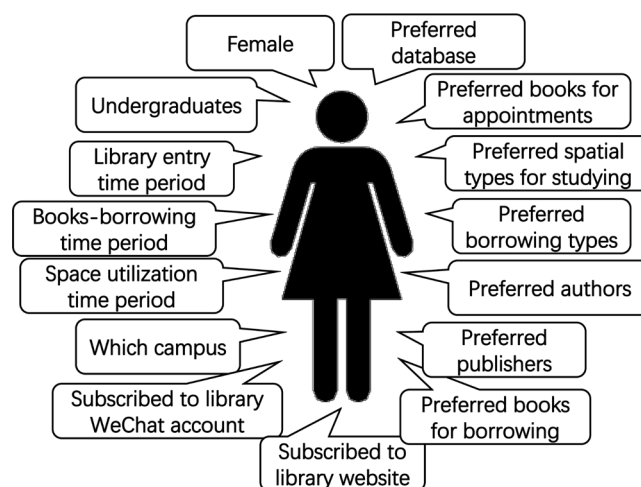


Figure 5. Library reading report: user portrait label.

3.4. Visual Presentation and Precise Service

At present, the personas embodied in the annual reading report of the library are mainly to reveal the readers' reading behaviors, such as the persona model presented through deep analysis and visualization, which is beneficial to the library to carry out the reading promotion and wisdom recommendation services for a certain group of readers. Specifically, the user group image results of the reader group can be subdivided, and the active readers, ordinary readers and potential readers can be deeply explored. The reading behavior and reading tendency of the specific reader group can be analyzed, and existing problems in the reading promotion process can be found. Based on the precise description and description of the individual or group of users, combined with knowledge discovery technologies, the library can establish multiple associations among users with similar knowledge needs, interest preferences, usage habits and activity levels, forming a portrait-based analysis. User relationship maps to reveal deep values of value [12]. As shown in **Figure 6**, based on various types of reading data of various groups of readers, knowledge map technology can be used to analyze the results based on the formed portraits to establish knowledge associations and interactions among users, and to find similar information needs of various reader groups and potential needs to reveal deep reading tendencies, so that libraries

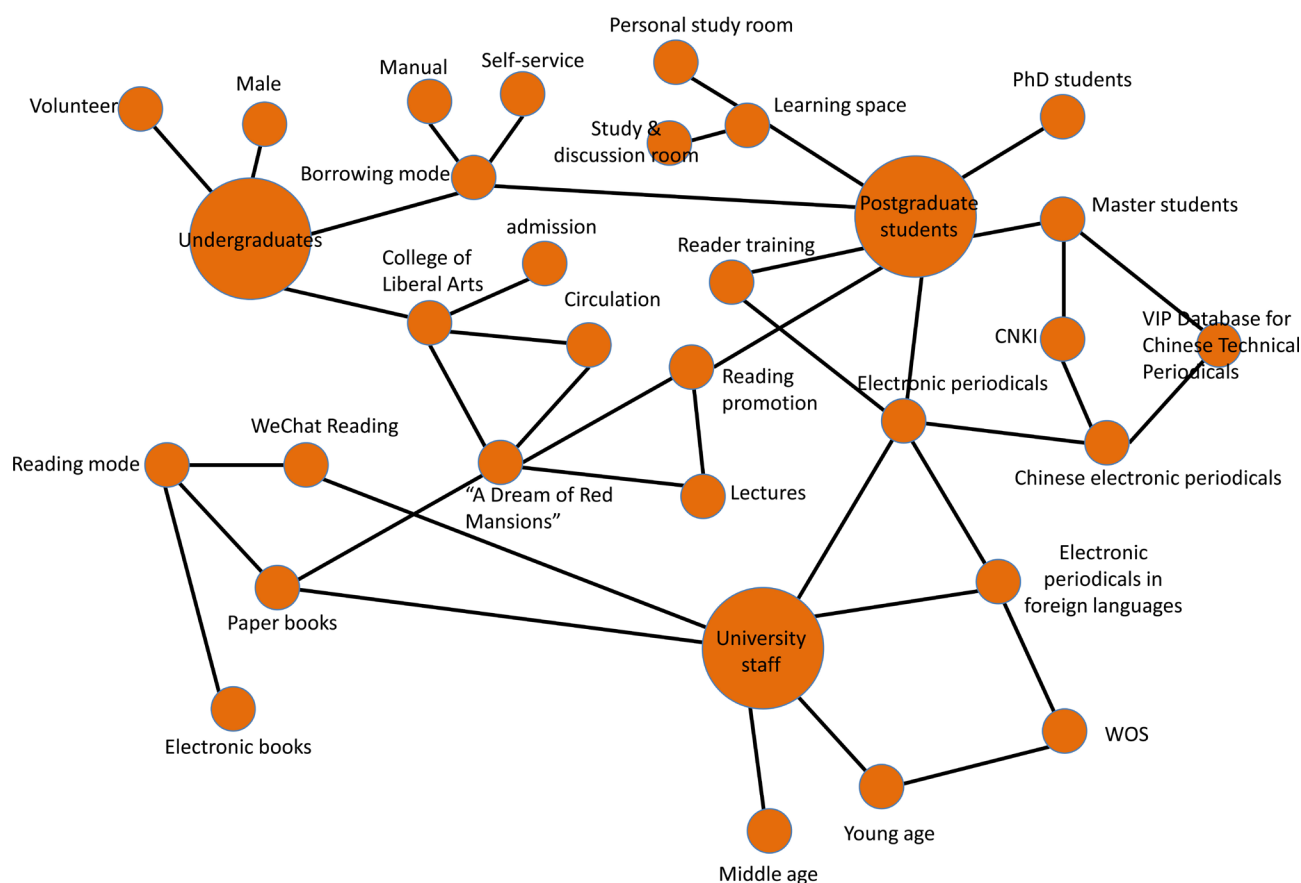


Figure 6. Knowledge relevance in the user portraits of the library reading reports.

can carry out various reading promotion services more accurately.

4. Conclusion

The study of the annual reading report of the library has guiding significance for many of the work of the library, because only by grasping the objective information needs of the readers can the document service work be done in a targeted manner. With the advent of the era of big data, the National 13th Five-Year Plan outlines the implementation of the National Big Data Strategy. Personas, as one of the tools for achieving accurate services in the era of big data, can help libraries to judge, analyze, and predict readers' information, demands, and mine the potential information needs of readers, so that the library can provide the readers with more personalized, proactive and humanized intelligent services. This paper puts forward the introduction of personas into the reading report of the university libraries, and draws on the development experience of computer portraits in the field of computers, e-commerce and foreign libraries, and puts forward the design ideas of annual reading reports based on user portraits in the Chinese university libraries. In the field of library in China, it provides constructive suggestions for the mining, integration and classification of library data from the perspective of personas.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

Project

Guangdong Provincial Library Association Project (GDTK1809): Research on the annual reading report of academic libraries based on personas.

References

- [1] Yang, F. (2018) Library Big Data Practice Based on User Portrait Analysis—A Case Study of National Library of China Big Data Project. *Library Forum*, 1-9.
- [2] Alan Cooper (2018). https://en.wikipedia.org/wiki/Alan_Cooper
- [3] Teixeira, C., Pinto, J.S. and Martins, J.A. (2008) User Profiles in Organizational Environments. *Campus—Wide Information Systems*, **25**, 128-144.
- [4] Rossi, S., Ferland, F. and Tapus, A. (2017) User Profiling and Behavioral Adaptation for HRI: A Survey. *Pattern Recognition Letters*, **99**, 3-12.
<https://doi.org/10.1016/j.patrec.2017.06.002>
- [5] Bishop, J. and Lewis, P.R. (1985) BLAISE-LINE and the British National Bibliography: Profiles of Users and Uses. *Journal of Librarianship & Information Science*, **17**, 119-136.
- [6] Zhang, X. and Wang, R.K. (2015) Research and Enlightenment of the Annual Report of American Academic College Library. *Journal of Library Science*, **11**, 135-139.
- [7] Peking University Library.

- <https://www.lib.pku.edu.cn/portal/sites/default/files/zhanlan/readingreport2017.pdf>
- [8] Sun Yat-sen University Library.
https://mp.weixin.qq.com/s?__biz=MzA3NDA2NzIyMA==&mid=2651366524&idx=1&sn=89c4d09cee8a070d1d4471c6758135c6&chksm=84f97807b38ef11154afe92081d23e939a1b400cb4cf2fbe19e48dc6916e207b0baf083383b9&mpshare=1&scene=23&srcid=0505QzKVmGTqpXUH1iIaeYnE#rd
- [9] Construction of China's First-Class University Libraries from an International Perspective. <https://v.qq.com/x/page/n0709fxhgm.html?start=1955>
- [10] Chen, H.X. and Shao, B. (2017) Research Status and Enlightenment of User Profile in Foreign Library Fields. *Research on Library Science*, 16-20.
- [11] Zeng, L. and Zhang, Y. and Qiu, R.G. (2007) Adaptive User Profiling in Enhancing RSS-Based Information Services. *IEEE International Conference on Service Operations and Logistics, and Informatics*, Philadelphia, PA, 27-29 August 2007, 1-5.
<https://doi.org/10.1109/SOLI.2007.4383915>
- [12] Liu, S. (2017) The Persona in Digital Library Knowledge Discovery System—Taking Tianjin Library as an Example. *Library Theory and Practice*, No. 6, 103-106.