Journal of Library and Information Studies 19:2 (December 2021) pp.43-76
doi: 10.6182/jlis.202112_19(2).043

Exploring Media Images of Librarians in Newspapers of Mainland China and Taiwan

Fan Xie[1], Kuang-Hua Chen[2]

Abstract

Public opinion shapes the image of organizations. Stereotypes about libraries have harmed their image in conjunction with the quick development of technology. It is necessary to initiate new strategies so that public opinions on libraries could be improved and resources of libraries could fully demonstrate their values. Librarians play an important role in shaping the image of libraries. This study attempts to identify the image of librarians in the media based on two major newspapers in mainland China and Taiwan through text and sentiment analysis. The considered lexical features are high-frequency words, shared words, keyness words, and sentiment words. The results revealed that in mainland China, librarians are regarded as history experts whose key functions are protecting, preserving, and studying ancient materials. In Taiwan, librarians provide various high-quality services for readers. In both regions, individuals agree that patrons and resources are the major concerns of librarians, and librarians are a bridge between patrons and resources. Finally, the sentiment values of words related to librarians are positive in both regions, but only slightly exceed neutral sentiment.

Keywords: Librarian; Media Image; Sentiment Analysis; Text Analysis

1. Introduction

Public opinion is influenced by factors outside of an organization. Media affect the public mindset and generate public opinion (McCombs, Juan, Lopez-Escobar, & Rey, 1997). The power of the media is symbolic and persuasive; thus, media have the potential to affect the opinions of audiences (Van Dijk, 1995). Media images can influence marketing for both organizations and people (Feng, 2010; Grabe, Ward, & Hyde, 2008; Hammett, 2014; Pang, 2018; Wu, 2002; Yang, 2006). News media are a type of media that deliver news to the general public or to a target public. People learn factual information from news media and understand the importance of topics based on whether news media emphasize those topics (McCombs & Reynolds, 2002). Newspapers are a form of news media; thus, newspapers can be analyzed to better understand the media image of an organization and the influence of this image on human behavior. Libraries, as a knowledge source, have long collected valuable content and provided users with various knowledge services (Gimenez & Pedreno, 2002; Liu & Gao, 2009; Zhang & Chen, 2020). However, some people believe that libraries do not require highly educated staff to provide their services and that the tasks in libraries are routine, easy, and mundane (Majid & Haider, 2008). Moreover, libraries have recently faced technological shocks. Previously, libraries were critical information centers, but today internet

[1,2]Department of Library and Information Science, National Taiwan University, Taipei, Taiwan
* Corresponding Author: Kuang-Hua Chen, Email: khchen@ntu.edu.tw

臺灣大學圖書館期刊資料庫

search tools provide similar information-seeking functions. If libraries do not adapt by using marketing to improve their image and by attracting users, they cannot overcome these technological shocks. Library funding and staff might be reduced accordingly. Fortunately, many libraries have recognized the importance of public opinion and have made efforts to change public stereotypes. For example, the American Library Association has attempted to refresh the image of libraries and has held an image transform campaign since 2015 (Xie & Xu, 2018). Xiamen University Library improved its image with graduates by running a marketing event called Tuan Time since 2013; the event won second prize in the International Federation of Library Associations and Institutions International Marketing Award (Cheng, 2016). In Taiwan, National Taiwan University Library has initiated activities to promote its professional image, including a series of activities called "Our Distance to Graduate Students" for helping students write theses or dissertations.

People might ignore the services performed by libraries in daily life. Attempts to transform library images lack central themes. Moreover, some actions may go unnoticed by users because the services provided are inconsistent with their expectations of librarian services. Libraries and librarians must understand mainstream public viewpoints to establish a good image. Notably, the image of librarians is inseparable from the image of libraries. Creating a good image for librarians is beneficial for numerous reasons. Library administrators could implement reasonable strategies based on marketing experiences to create a better relationship with users. Experts in the library community can establish beneficial library and librarian images strategies simultaneously. Accordingly, libraries with positive images could attract more users, and services could better meet user needs. As a result, libraries would have greater funds for training librarians and acquiring resources.

Some scholars have already investigated librarian images based on patron opinions, but these investigations have been performed using questionnaires. By contrast, this study attempts to identify media images of librarians based on reports in major newspapers on both sides of the Taiwan Strait. The purpose of this study is to understand the media images of librarians based on newspapers in these two regions from 1995 to 2019. We identified media images presented in news articles by using text and sentiment analysis. The results of this study could help official organizations, welfare institutions, libraries, and other sectors improve their marketing.

The research objectives of this study are as follows:

1. To explore media images of librarians in mainland China and Taiwan.
2. To compare media images of librarians in mainland China and Taiwan.
3. To examine the sentiment tendencies of news articles related to librarians.

This article is structured as follows. Section 2 presents a review of previous studies and the guidelines for this study. Section 3 presents the research design and the details of the research subjects, data collection, research tools and resources, and research steps. Section 4 reports the results and discusses them from various perspectives and by considering sentiment tendencies. Finally, Section 5 concludes this study.

臺灣大學學術期刊資料庫

2. Literature Review

2.1 Librarian media image

The media image of librarians is the public's general impression of librarians formed by the mass media. Related studies have typically employed content analysis, document analysis, questionnaires, and interviews, but scholars have not reached a consensus on the necessary factors for media image analysis. Edgar (1976) analyzed the public stereotypes of librarians based on newspapers, magazines, and TV programs. The Paper of Qiu (1992) is the earliest study about librarian media images in mainland China. Qiu reported that librarians regard themselves as a tool for connecting books and users, but he suggested that librarians could increase their confidence to change that self-perception. The first study of librarian media images in Taiwan was performed by Wu (2003). He investigated various media, such as movies, novels, and TV advertisements.

Media are categorized as videos and written works in this study. In related papers, we noted that scholars frequently analyze the physical characteristics of librarians, such as style of dress and age. Librarians are sometimes described as conservative, middle-aged women who lack professional skills and may even obstruct the information-seeking process (Black, 1994; Hsieh, 2016). Furthermore, in Japanese video media, male librarians are frequently portrayed as incompetent (Li, 2010). However, librarians do not always fit these stereotypes. In science fiction films, librarians are intelligence experts who assist the protagonist (Wu, 2003). In Chinese kung fu films, librarians may even have dual identities. In *Black Mask*, Jet Li played a weak librarian who was secretly the world's most powerful killer (Xu, 2011). Although the traditional stereotype of librarians has evolved due to these portrayals of librarians, librarians are still typically portrayed as bookish types in the film industry (Helms, 2006). In written media, including novels and newspapers, researchers have investigated the contents of works about librarians. The findings reveal a lack of representation of librarians' skills and duties despite the press rarely reporting negative images of librarians and regarding librarians as experts in knowledge management (Maynard & McKenna, 2005; Robinson, 2006; Shaw, 2010). Although most librarians are female, writers have also created a few male librarians in young adult novels; these novels can improve the image of libraries for young users as well as improve their understanding of librarians (Guo, 2015).

Changes in librarian images can be observed in papers on a broader level. Images of librarians in the media are typically positive, but negative stereotypes also exist. The image of a librarian has ranged from a bespectacled mousy man with little social skills, power, or ambition in the 1800s to a shushing spinster with a bun in the 1900s. Their tasks were solely checking out books and collecting fines (Radford & Radford, 1997). The media images of librarians in the twenty-first century did not evolve in response to advances in information technology (Luthmann, 2007). However, some technology has been included in library scenes, and the appearance of librarians has also changed; fashionable young female librarians have become more common (Wells, 2013). Furthermore, if people create and curate content in digital media, gender stereotypes

臺灣大學學術期刊資料庫

about librarians (i.e., librarian is a feminized job) can be strongly challenged (Singh, Chayko, Inamdar, & Floegel, 2020). However, we are not convinced that the image in Chinese newspapers reinforces conventional stereotypes of librarians, and we hope to identify the roles of and changes in librarians in the two regions. Therefore, we performed a study to investigate changes in librarian images in Chinese news articles.

2.2 Text analysis

Text analysis (also called text mining) is the application of techniques or tools to explore word, phrase, and sentence features in texts. All features are considered together to reveal information to solve problems. Researchers using text analysis in library and information science (LIS) not only analyze texts but have also compared text analysis tools and algorithms (Abbasi & Chen, 2008; Magerman, Van Looy, & Song, 2009; Nasukawa & Nagano, 2001; Tseng, Lin, & Lin, 2007; Wilbur et al., 1999).

Approaches for text analysis can be divided into two categories: lexical features and content features. The first approach involves analyzing the lexical features of the text, such as word frequency, word co-occurrence, and other features (Hsieh & Yang, 2018; Ke & Sie, 2019; Liu, 2013; Short, 2019). Many tools can be used to extract lexical features, such as the R statistical packages (Maceli, 2015), WordSmith Tools (Yang & Ling, 2018), and PamTAT (Al-Daihani & Abrahams, 2016). Maceli (2015) conducted text analysis with the R statistical packages to count the frequency of words in job titles and job descriptions and identified the skills librarians must master. The results revealed that librarians not only require technical skills, but also must be able to manage technology and make decisions. The second approach is to create a coding sheet in accordance with the text content. The content features are based on information from text and requires manual judgments. Scholars classify texts in various fields according to their research questions and by using tools such as Nvivo 9 or SPSS (He, Zha, & Li, 2013). Scholars have discussed software and coding sheets used in LIS research and observed that software is increasingly used in the field (Pan, Yan, Cui, & Hua, 2019).

2.3 Sentiment analysis

Sentiment analysis is also known as opinion mining and is a computational study of people's opinions, sentiments, emotions, appraisals, and attitudes toward entities such as products, services, organizations, individuals, issues, events, and topics (Liu, 2015; Zhang, Wang, & Liu, 2018). At least two approaches to sentiment analysis exist: one is the lexicon-based approach; the other is the machine learning approach.

A sentiment lexicon is indispensable for the lexicon-based approach. Stone, Bales, Namenwirth, and Ogilvie (1962) first performed the sentiment analysis task; later, scholars proposed using different lexicons such as WordNet, WordNet-Affect, SenticNet, MPQA, and SentiWordNet (Habimana, Li, Li, Gu, & Yu, 2020). Researchers either create a sentiment lexicon or use an existing lexicon (Muhammad, 2019), extract sentiment values for words in the text from the sentiment lexicon, and finally sum these values to identify the sentiment of the text.

The machine learning approach uses a learning algorithm trained with labeled or unlabeled

臺灣大學學術期刊資料庫

sentiment data to identify features that can be used for subsequent tasks (Kuo & Zhang, 2013). The machine learning approach could be supervised or unsupervised. Turney used a pioneering unsupervised learning method (Turney, 2002). He selected two seed words (poor and excellent) and computed the associations of phrases with the seed words. The semantic orientation of phrases could thus be identified. The sentiment of a text was calculated as the average semantic orientation of all examined phrases in the paper (Zagibalov & Carroll, 2008).

Sentiment analysis is useful in many research areas, including public opinion management and control, business decisions, opinion retrieval, information prediction, and emotion management. The main application in LIS is public opinion management with the goal of improving library services. Zhang (2012) used sentiment analysis to extract positive or negative user opinions from social networks and discovered that users are not satisfied with library collections and services. Choi (2019) analyzed online children's book reviews, and the final results revealed a weak correlation between a book's sentiment value and its rating. Moreover, online reviews could identify critical topics to improve the library catalog. Cheng and Wu (2021) collected over 70,000 online user reviews to examine user perspectives on the digital reading of preschool children and discovered a significant shift in the attitudes of Chinese users. Aside from that, sentiment analysis can be useful in surveying larger public events. To examine the public response to COVID-19—related outreach efforts of the Ministries of Health in Singapore, the United States, and England, Raamkumar, Tan, and Wee (2020) used the DeepMoji algorithm to identify emotions in Facebook comments. They observed that most of comments responding to departmental posts were angry. Golder and Macy (2011) used Linguistic Inquiry and Word Count to analyze data from millions of public Twitter messages and identified individual-level diurnal and seasonal mood rhythms in cultures worldwide.

After examining related researches and approaches, we used the following guidelines to perform this study.

1. Most of the previous studies used questionnaires to investigate media images. This study used news articles instead.
2. A large temporal range (25 years) was used for news article analysis. A larger temporal range facilitates observations of transitions between different periods.
3. The lexicon-based approach was used due to its effectiveness in uncovering evidence in small corpora.
4. The lexical items used in this study were high-frequency words, shared words, keyness words, and sentiment words.

3. Research Design

3.1 Research subjects

This study investigated media images of librarians based on newspapers in mainland China and Taiwan. The newspapers used in this study were the People's Daily and China Times published in mainland China and Taiwan, respectively.

1. The People's Daily is one of the most popular state newspapers in mainland China. In 1992, it was named one of the top 10 newspapers in the world by UNESCO and has substantial

臺灣大學學術期刊資料庫

influence both domestically and overseas (People's Daily, n.d.).

2. The China Times is one of the top four newspapers in Taiwan and was founded in 1950 by Yu Jizhong. The China Times was the first Taiwanese newspaper to be published digitally in 1995 (China Times, n.d.).

The available news articles of the People's Daily database and China Times are not equally distributed across publication years; the largest overlapping temporal range (from 1995 to 2019) was used in this study. News articles related to librarian image were retrieved from the databases of the People's Daily and the China Times by using three keywords: "圖書館(library) and 館員(staff)," "圖書館員(librarian)," and "圖書管理員(library administrator)." Then we deleted duplicate news articles in the three retrieval results. The number of related news articles in the People's Daily and the China Times databases from 1995 to 2019 were 249 and 662, respectively. To investigate trends, the 25-year range was divided into five periods of five years each. Table 1 presents the number of relevant news articles in the five periods.

3.2 Research tools

To analyze the text collections (i.e., the collections of news articles), the following platforms, tools, and resources are used in this study:

1. NLPIR, Natural Language Processing & Information Retrieval Sharing Platform, a natural language processing system. Its functions include Chinese word segmentation, keyword extraction, and part of speech tagging (Zhang & Shang, 2019). NLPIR has a good reputation from 300,000 institutions with 400,000 recorded users worldwide including Huawei, CITIC Trust, and Tsinghua University (NLPIR1, 2021). New articles in simplified Chinese were segmented by NLPIR.
2. CORPRO, a Chinese corpus-based tool for textual data analysis (Chueh & Chen, 2016). Simplified Chinese and traditional Chinese sometimes do not have a perfect one-to-one mapping; for example, the characters "台" and "臺" are semantically different in traditional Chinese but are both represented by the single character "台" in Simplified Chinese. Therefore, to avoid mismatches between traditional and simplified Chinese characters, CORPRO was adopted to analyze the traditional Chinese text.
3. SPSS 25, an IBM software for statistical analysis, data mining, value prediction, and decision-support. SPSS was used to analyze sentiment data in this study.
4. BosonNLP is a commercial natural language processing platform that includes various functions for sentiment analysis, text information classification, named entity

Table 1. The Retrieval Results in the Five Analyzed Periods

	T1 (1995–1999)	T2 (2000–2004)	T3 (2005–2009)	T4 (2010–2014)	T5 (2015–2019)	Total (1995–2019)
People's Daily	23	38	64	67	57	249
China Times	103	190	213	102	54	662

臺灣大學學術期刊資料庫

recognition, opinion extraction, text clustering, and keyword extraction (Boson, n.d.).

5. HowNet: HowNet is an online common-sense knowledgebase with interconceptual relationships and interattribute relationships of concepts and their connotations in Chinese lexicons with their English equivalents (Dong, Dong, & Hao, 2010).

3.3 Research steps

Step 1: Identify high-frequency words. The goal of this step is to identify high-frequency words to better understand people's impressions of librarians in the two regions. Important keywords of librarian media images in mainland China and Taiwan were identified. After segmenting words in news articles, words were ranked based on collection frequency (CF). The CF is the total frequency of a word in the corpus. Words with higher frequency have greater importance in the corpus (Salton & Buckley, 1988). CF was used instead of term frequency-inverse document frequency (TF-IDF) because the length of the document does not affect CF values. Therefore, normalizing word frequency to decrease the effects of document length is unnecessary. Texts in traditional Chinese were processed by CORPRO and simplified Chinese texts were processed by NLPIR. Stop word removal was performed before counting the frequency of words in newspapers. We included the stop words @, ◆, and ■, etc. into the stop word lexicons of CORPRO and NLPIR. The top 20 words were selected for further analysis. This number could be changed; however, if the number of keywords was fewer than 20, we observed the gaps between each word and selected words with a gap beyond the significant gap.

Step 2: Select shared words. In this step, words shared between the two regions were selected for comparisons. After segmenting the news articles, a list of words in each region was generated. Query functions in Microsoft Excel were used to search for same words in both regions for the five periods. Shared words were ranked by CF in accordance with the results of the first step. Because shared words have CF values for both the simplified and the traditional Chinese corpora, the CF values must be normalized. Each CF was divided by the total number of words in the corresponding corpus. Shared words were ranked by the sum of the two normalized CF values. Four categories of shared words were identified according to the number of periods in which the shared words appeared: high duration, middle duration, low duration, and no duration.

Step 3: Extract keyness words. A keyword is defined as a word that occurs with unusual frequency in a given text (Scott, 1997). Keyness is a feature of keywords and is expressed as a statistical probability (WordSmith Tools Help, n.d.). We compared keyness words for the two regions in this step by using CF to extract keyness words. Keyness words were used to compare librarian media images in the five periods of 1995 to 1999 (T_1), 2000 to 2004 (T_2), 2005 to 2009 (T_3), 2010 to 2014 (T_4), and 2015 to 2019 (T_5).

For keyness identification, the CF of words in news articles were calculated for each region during the same periods. A line chart was then created to display the differences in CF values between the two regions; clear differences indicated possible keyness words. We used the mainland China corpus as the reference corpus against the Taiwan corpus. Keyness statistical

臺灣大學學術期刊資料庫

approaches include the chi-square test, log-likelihood (LL) ratio, TF-IDF, deviation from randomness, Kullback-Leibler divergence, and information gain (Oakes, 2008), etc. Chi-square and LL are frequently used in keyness research and produce similar keyword rankings (Chujo & Utiyama, 2006). However, LL values in the two keyword lists should not be compared if the sizes of the collections are different (Pojanapunya & Todd, 2018). Because the two studied collections had different sizes (249 and 662), the chi-square test in CORPRO was performed to determine whether the selected words in the Taiwan collection were significantly different ($p < .05$) compared with those in the mainland China corpus. Significantly different words were identified as the keyness words for that period. Both positive and negative keyness words were identified. Positive and negative keyness words are those that occur more and less frequently, respectively, than expected by chance in the reference corpus (Scott, 2004). Thus, positive keyness words are Taiwan's keyness words and negative keyness words are mainland China's keyness words.

In addition, we categorized keyness words with different themes according to the topics of the two newspapers, such as people, organizations, action, time, location, thing, or other, for further analysis and comparison. The changes and trends of keyness words are depicted in the appropriate tables.

Step 4: Calculate sentiment values. The sentiment value of each word was used to determine the sentiments of the general public toward librarians. Sentences with the word "圖書館員(librarian)" were considered in the analysis to ensure the accuracy of the final results. Sentiment words reflect the public satisfaction and opinions about librarians. The words identified in Step 1 were mapped with words in a sentiment dictionary BosonNLP (Boson, n.d.) and assigned designated sentiment values. If BosonNLP could not provide needed information, HowNet (Dong & Dong, 2003) was used instead.

The following presents the process for Step 4. First, sentences containing "圖書館員(librarian)" were considered as the targets. Each news article may have more than one qualified sentence. All qualified sentences in news articles were concatenated as long sentences. Second, we normalized the frequency of each word. The frequency normalization for each word is defined as follows:

$$F_k = \frac{f_k}{f_{total}} \quad (1)$$

where f_k is the frequency of word k, f_{total} is the number of all words, and F_k is the normalized frequency of word k.

Third, the sentiment value for each word was also normalized. The value for each word was normalized using the maximal value in BosonNLP as follows:

$$V_k = \frac{v_k}{v_{max}} \quad (2)$$

where v_k is the sentiment value of word k, v_{max} is the maximal sentiment value found in BosonNLP, and V_k is the normalized sentiment value for word k.

$$SV_k = F_k \times V_k \quad (3)$$

where SV_k is the total sentiment value of word k in the qualified sentence.

Fourth, the sentiment value of a sentence (the "long" sentence; i.e., the news article) is

臺灣大學學術期刊資料庫

the sum of the weighted sentiment value of all words as follows:

$$SV_s = \sum_{k=1}^{n} SV_k \quad (4)$$

where n is the number of words in the "long" sentence s and SVs is the sentiment value of sentence s.

Based on the aforementioned processing, the sentiment distribution of news articles on both sides of the Taiwan Strait could be determined and compared. The independent chi-square test was used to determine whether word usage differs significantly between the regions. The sample t test was used to test whether the sentiment distribution is differs significantly. Cluster analysis was used to investigate the sentiment in both regions. Finally, further analyses could be performed accordingly to understand the characteristics of both sides.

4. Results and Discussion

In accordance with the literature review 2.1, we portrayed changes in media images on the basis of five factors: appearance, character, behavior, competence, and sentiment. "Appearance" indicates the external appearance of librarians. These words usually are typically adjectives or nouns, such as "scholar" or "plain." "Character" words are words describing personality, such as "kind" or "enthusiastic." Notably, some character words only indirectly reveal the media images of librarians because the analyzed objects are all news, and the news may contain figures other than librarians. "Behavior" words are linked to words about actions and events in texts and indicate the actions of librarians mentioned in the news. "Competence" words denote the professional skills of librarians mentioned in the news articles. These words are typically connected to actions or events, such as "information retrieval" or "database." "Sentiment" words indicate moods of librarians mentioned in the news.

4.1 High-frequency words

4.1.1 High-frequency words in mainland China

We selected the top 20 words in the mainland China corpus and ranked them by CF. CF was determined for each period rather than for the entire corpus. To identify connections between high-frequency words, we categorized these words into four usage duration levels.

High-duration words receive substantial attention from the general public, middle-duration words are less frequently used to describe librarians, and low-duration words are only used in two periods, and no-duration words were only widely used in one period. Notably, other periods may also include these no-duration words but at lower frequencies. High frequency was a necessary criterion for duration words.

(1) High-duration words

High-duration words had high frequency in all five periods. "圖書館(library)," "國家(country)," "人(people)," "中國(China)," "讀者(reader)," and "文化(culture)" were the most commonly used words across the entire study duration. This result was consistent with a description of librarians: the successful librarian is an individual who is functioning in a job which provides satisfaction both to himself and to the recipient of his services (Watkins, 1968). That is, librarians are those who work for people or readers in a library and protect the culture of their country. "中國(China)" had high frequency because news about librarians

臺灣大學圖書館期刊資料庫

was collected from this region, although China was not a word used to retrieve news articles. “圖書館(library)” was another high-duration word because this study was focused on libraries. “文化(culture)” was also identified in relation to the development of cultural industry. For example, a Hubei delegate suggested that the government should increase investment in the culture industry. High-duration words in mainland China reveal that librarianship is a project related to the culture that has developed over the past 25 years.

(2) Middle-duration words

A word appearing in three or four periods was a middle-duration word. “書(book),” “工作(work),” “古籍(ancient book),” “資訊(information),” “公共(public),” “大學(university),” “社會(social),” and “服務(service)” were the middle-duration words in mainland China. These words indicate that “librarian” is a job that provides services to readers with books or information. This result is similar to that of Nelson and Irwin (2014) who indicated that librarians based their identity on claimed expertise regarding how information should be organized and obtained. The words “書(book)” and “古籍(ancient book)” also indicate that librarians protect ancient books with high academic value. In a news article, a librarian of the French National Library provided a writer with historical documents about Chinese in France. Middle-duration words indicated changes in the work of librarians. From 2005, Chinese newspapers often used the phrase “古籍(ancient book)”; this may be related to a restoration project for Qing dynasty documents from 2004 (Note 1). Furthermore, since 2005, “公共(public)” was the most frequently used word in texts and is related to the introduction of American-style public library services during that period. From 2010 to 2014, numerous news articles described methods of increasing book usage in library communities. In different periods, different requirements for libraries were identified, and librarians needed to develop new skills and overcome technological challenges. “大學(university)” typically indicates the name of a university rather than its library in text, because “大學(university)” is typically used to introduce experts from an institution. An expert may be a librarian discussing library-related issues or an individual had intersections with librarians. For example, Peking University’s librarian Chu Xiongfei mentioned that he met Professor Meng Erdong working in the Department of Ancient Books almost every day.

(3) Low-duration words

Low-duration words are high-frequency words appearing in only two periods. “北京(Beijing),” “先生(sir),” “文獻(document),” “發展(development),” “作者(author),” “研究(research),” “閱讀(read),” and “活動(event)” are low-duration, high-frequency words. “Low duration” does not indicate that these words were not used in other periods. For example, “北京(Beijing)” appears in all five periods, but was only used with high frequency in 1995–1999 and 2005–2009. In 2010–2014, “閱讀(read)” was used to indicate that librarians in public were responding to the challenges of digital books and that librarians in university libraries should assist students in improving their humanistic literacy. In 2015–2019, “閱讀(read)” indicated that librarians needed to attract public attention to implement the draft of “Reading for All (《全民閱讀促進條例》),” which was released by the Chinese

臺灣大學學術期刊資料庫

government in 2017. Librarians also needed to overcome the challenges of digital reading and redesign systems for ancient documents to increase usage. The result corresponds with advertisements for librarian jobs that ranked "public service" as a librarian's secondary duty (Gerolimos, Malliari, & Iakovidis, 2015). Therefore, reading and document-related services were the primary responsibilities of librarians since 2010.

(4) No-duration words

No-duration words only appeared with high frequency in one period. Two phenomena were observed for no-duration words. First, people mentioned in the news were changed from the other fields' experts to the librarians. Newspapers published librarians' related experts in 1995–1999. "陳寧(Chen Ning)" was a young computer scientist with poliomyelitis. Librarians in Hunan Provincial Library were impressed by his working schedule and performance. "凡爾納(Jules Verne)" was a fiction writer. However, news articles more frequently mentioned librarians in 2000–2004. "毛澤東(Mao Zedong)" was a librarian at Peking University before becoming the leader of the Communist Party of China. "章士釗(Zhang Shizhao)" was a famous politician of the Republic of China, and Mao Zedong frequently turned to him for help and suggestions. Second, the high CF words changed from words that were people-related to words that were more service-related over time. Those words indicated that librarian services are a part of communities, and communication is a critical part of these services. "社區(community)" and "村落(village)" are key areas for library development and "交流(exchange)" and "合作(cooperation)" among libraries are also important activities; for example, a Chinese and British library forum was held in Chengdu 2019.

High-frequency words in mainland China reveal that the image of a librarian is a famous male experts (not limited to LIS professionals). Reader services provided by these professionals include ancient book preservation and reading promotions. Although various technologies are used in LIS, librarians and libraries are still considered cultural rather than scientific entities.

4.1.2 High-frequency words in Taiwan

High-frequency words in Taiwan were examined to identify high-duration words, middle-duration words, low-duration words, and no-duration words.

(1) High-duration words

The high-duration words in the Taiwanese newspaper were "圖書館(library)," "書(book)," "民眾(populace)," "館員(librarian)," "報導(report)," and "臺灣(Taiwan)." These words indicate that librarians frequently work with books and readers. This result also confirmed previous studies: finding information for people was the most frequent task undertaken by librarians in public libraries, academic libraries, and special libraries (Harrison & Beraquet, 2010; Vassilakaki & Moniarou-Papaconstantinou, 2016; Wang, Tang, & Knight, 2010). "報導(report)" introduced individuals who wrote or were responsible for the news articles. The reporter of a news article may also be a librarian.

(2) Middle-duration words

Middle-duration words are high-frequency words appearing in 3 or 4 periods. "活動(activity)," "閱讀(reading)," "新聞(news)," "工作(work)," "服務(service)," "自己(self)," "文化(culture)," "讀者(reader)," and "美國(America)"

臺灣大學學術期刊資料庫

were the middle-duration words. “活動(activity),” “讀者(reader),” and “美國(America)” were high-frequency words in four periods: 1995–1999, 2000–2004, 2005–2009, and 2010–2014. Newspapers in Taiwan began to discuss the library reading problem earlier than those in mainland China; “閱讀 (reading)” became a high-frequency word in Taiwan between 2000 to 2004, whereas it only became a high-frequency word in mainland China between 2010 and 2014. Book exchanges, book recommendations, child-parent reading, and events that help readers understand libraries were the main activities organized by librarians in Taiwan. Moreover, news articles reported the activities of numerous public libraries in these four periods but of few university libraries. “美國(America)” is a country that produces numerous recommended books and scientific data; many outstanding people also study or live there. For example, “楊永斌(Yang Yongbin)” suggested that good books result in a good life. He is an honorary professor at the National Yunlin University of Science and Technology. “工作(work),” “服務(service),” and “自己(self)” were used to indicate that librarians were not focused on monetary compensation but instead prioritized a sense of duty through public service.

(3) Low-duration words

“小朋友(little children),” “學生(students),” “孩子(children),” “提供(support),” “時間(time),” “借閱(borrow),” and “本書(the book)” are low-duration words. These low-duration words indicate a clearer picture of changes in perceptions of librarians over time. Most low-duration words were identified in the periods of 2005–2009 and 2015–2019; the results revealed that these two periods were important for librarian development. Moreover, the low-duration words revealed that the major patrons of libraries in Taiwan are children and the most important services provided by librarians are tasks related to circulation. Some librarians changed the operating hours of the library to respond to the needs of student and adult readers. This flexibility indicates that librarians put their readers first.

(4) No-duration words

No-duration words from 1995 to 1999 were “文化中心(culture center),” “琉球(Ryukyu),” “記者(journalist, usually connected with the news reporter’s name in texts),” “鄉鎮(villages and towns),” “使用(use),” “台北(Taipei),” and “人員(staff)”; those from 2000 to 2004 were “人(people),” “資料(material),” “問題(issue),” “文化局(department of culture affairs),” “公共圖書館(public library),” and “資訊(information).” No no-duration words were identified from 2005 to 2009. The no-duration words were “書店(bookstore),” “志工(volunteers),” “出版(publish),” “學校(school)” in 2010–2014 and were “分館(branch library),” “新北市(New Taipei City),” “邦德(Julian Bond),” “學童(schoolchildren),” and “體驗(experience)” in 2015–2019. The most frequently mentioned words were related to cultural organizations or to people.

In comparison to the no-duration words for mainland China that were related to the preservation of ancient books, these words in Taiwan were related to reader services, in particular, services for child readers. Furthermore, Taiwanese librarians have wider networks with other organizations than mainland Chinese librarians do. These organizations are bookstores, schools, and local cultural departments. However, the word identification revealed that mainland

臺灣大學學術期刊資料庫

Chinese librarians have become more open than in previous periods and are also expanding their networks. Stakeholders in networks facilitate the provision of user services by libraries and the initiation of professional training courses.

4.2 Data analysis for shared words

The term "shared words" in this study indicates words that appear in both collections. To compare shared words, the same categorization scheme for high-frequency words was again used. Thus, the evolution of topics appearing in news articles on both sides of Taiwan Strait can be discussed temporally.

(1) High-duration shared words

High-duration shared words are words identified with high frequency in both mainland China and Taiwan for all five periods. These words include "圖書館 (library)," "書(book)," "文化(culture)," "服務(service)," "活動(activity)," "中國(China)," "讀者(reader)," and "工作(work)." Clearly, activities for readers are considered the most critical tasks on both sides of the Taiwan Strait. These results also agree with the results of the previous analysis for high-frequency words.

(2) Middle-duration shared words

Middle-duration shared words are words shared in three or four periods. These words may have also had high frequency in other periods; the selected shared words met two criteria: they had high frequency in both regions and were in the top 20 shared words. According to the criteria, "人(people)," "國家(country)," "資訊(information)," "館員(librarian)," "大學(university)," "學生(student)," "美國(America)," and "閱讀(read)" were the middle-duration shared words. "資訊 (information)" usually co-occurs with "館員(librarian)" and "大學(university)"; thus, universities are institutions where librarians perform informational services. Moreover, words identified for both mainland China and Taiwan revealed concerns about library patrons and universities issues.

(3) Low-duration shared words

Low-duration shared words are words appearing in two periods for both mainland China and Taiwan. "先生(Mr)," "問題(issue)," "社會(society)," "古籍(ancient book)," "孩子(children)," and "歷史(history)" are the low-duration shared words. The low-duration shared words generally indicate that children and the relationships with their parents were major concerns for libraries over a relatively long period. Additionally, historical issues and rare books were important topics in both regions; these documents included old photos, old posters, and oral histories. "先生(Mr)" is typically used to refer to well-known people. From 1995 to 2004, "美國(America)" and American libraries were frequently mentioned in news articles in both mainland China and Taiwan. Numerous services or concepts originating in America have been adopted on both sides of the Taiwan Strait. Because words related to librarian appearance did not appear in shared words of high-, middle-, or low-duration, no information was gathered on this topic.

(4) No-duration shared words

No-duration shared words appeared in one period only, but were shared between the two regions. The no-duration shared words changed from those related to preservation of historical materials to resource-sharing between libraries. "北京(Beijing)," "作者(author)," "生活(life)," and "記者(journalist)" were the top 20 shared words from 1995 to 1999.

臺灣大學學術期刊資料庫

In mainland Chinese newspapers, “北京(Beijing)” is the city from which mainland Chinese librarians or other experts come from; in Taiwan, the term “北京(Beijing)” is typically connected with a theft of ancient books from Peking University. “記者(journalist)” indicates a role at the newspaper for mainland China, but tends to appear in conjunction with a reporter’s name in Taiwan. “毛澤東(Mao Zedong),” “藏書(book collection),” “資源(resource),” and “研究(research)” were the no-duration shared words in the period from 2000 to 2004. Mao Zedong is typically mentioned in connection with libraries in both regions because he was a librarian at Peking University. The term “藏書(book collection)” typically appears in conjunction with a number of volumes (e.g., Guanmiao Township Library has a collection of 35,000 volumes). “資源(resource)” was related to librarians’ suggestions for resource-sharing.

In the period from 2005 to 2009, “時間(time)” and “文獻(document)” were used in both regions, “時間(time)” typically referred to the opening and closing times of a library, and “文獻(document)” indicated historical documents in the context of the news articles; these documents are a major concern of libraries.

From 2010 to 2014, “公共(public),” “發展(development),” and “書店(bookstore)” were in the top 20 shared words. These terms indicate that public libraries have a key role in the development of services by libraries and librarians. “書店(bookstore)” refers to stores selling old or second-hand books in Taiwan and new books in mainland China. The relationship between bookstores and libraries is complex; it is not always competitive but may be mutually beneficial.

“傳統(tradition),” “交流(exchange),” “成為(become),” “合作(cooperation),” “提供(support),” and “民眾(people)” were the top 20 shared words from 2015 to 2019. This five-year period had high commonality between the regions. The increase in shared words indicated that word usage became more similar between mainland China and Taiwan; this may be due to easier communication between the two regions in the internet era. “傳統(tradition)” is mentioned in mainland China when describing rural libraries that preserve traditional treasures. “交流(exchange)” is mentioned in the context of communication and collaboration with foreign organizations. “合作(cooperation)” usually appears in a similar context to “交流(exchange).”

4.3 Data analysis for keyness words

The keyness words are those for which the word frequency in one corpus is significantly higher than that in the other corpus. The collection of news articles from mainland China was regarded as the reference corpus; that of Taiwan was the target corpus. If the p value of a word (chi-square test) was less than 0.05, the word was regarded as a keyness word. However, more than 300 words were identified as keyness words based on this criterion; therefore, a new criterion of a p value of $\leq .001$ was adopted. Moreover, the frequency must be zero in one of the two corpora. Positive and negative keyness words were selected based on the new criteria. Because only two corpora were used in this study, positive keyness words represent the characteristics of Taiwan and negative keyness words are representative of mainland China. Seven categories for keyness words were identified: people, organization, action, time, location, thing, and other.

臺灣大學學術期刊資料庫

4.3.1 Keyness words in 1995 to 1999

As presented in Table 2, "location" category words were keyness words for Taiwan from 1995 to 1999. For example, Taiwan obtained a selection of precious Ryukyu documents; thus, Taiwanese and Ryukyu librarians cooperated to annotate, compile, and publish them. The second category was "people"; the words in this category indicate that libraries focus more on child readers than other readers. The final category was "organization"; the results reveal that people consider libraries to be a cultural center. The keyness words were derived from the comparisons of two corpora, and we conclude that in 1995–1999, Taiwanese librarians focused more on users and township library services than mainland Chinese librarians did.

The keyness words for mainland China are presented in Table 3. "風生水起(feng sheng shui qi)" was the title of a TV drama that stimulated discussions on advertising and securities. Numerous "電視劇(TV dramas)" in 1995–2000 were similar to "風生水起(feng sheng shui qi)" in that they drew public attention to various issues. Similarly, "中醫(Chinese medicine)" is connected to "本草綱目(*Compendium of Materia Medica*)"; a researcher, "劉山永(Liu Shanyong)" spent 10 years compiling the book in accordance with his father's last will." Finally, the swastika symbol (卐) was discussed because a company printed this symbol on its publications, but the authors disagreed with this action due to its association with Nazism. Mainland China had more common categories overall, indicating significant differences in language usage between mainland China and Taiwan, especially in the words used to describe things and actions. The keyness words related to the "thing" category indicated different focuses of library work between the two regions. Librarians in mainland China primarily performed the tasks of "編撰(compilation)" and "校勘(collation)" of ancient books; for example, Xia Renhu, a librarian at the Central Museum of History, donated books he had helped compile. For the "people" category, differences could be due to different writing styles. The keyness words

Table 2. Keyness Words of Taiwan from 1995 to 1999

Category & percentage	Keyness word	CF in TW
Organization (103/557 = 0.185)	文化中心(culture center)	103
People (224/557 = 0.402)	民眾(populace)	107
	小朋友(children)	63
	圖書館員(librarian)	54
Location (230/557 = 0.413)	琉球(Ryukyu)	72
	鄉鎮(villages and towns)	58
	臺北(Taipei)	50
	臺灣(Taiwan)	50

臺灣大學學術期刊資料庫

Table 3. Keyness Words of Mainland China from 1995 to 1999

Category & percentage	Keyness word	CF in MC
Others (12/296 = 0.041)	貧困(poor)	12
Action (33/296 = 0.111)	科研(research)	12
	編撰(compile)	11
	校勘(collation)	10
Organization (35/296 = 0.118)	北圖(National Library of China)	21
	銀行(bank)	14
Location (39/296 = 0.132)	這裏(here)	19
	北京市(Beijing)	11
	長春市(Changchun)	9
Thing (67/296 = 0.226)	電視劇(TV drama)	20
	卐 (Hakenkreuz, Nazi symbol)	15
	本草綱目(*Compendium of Materia Medica*)	12
	風生水起(feng sheng shui qi)	11
	中醫(Chinese medicine)	9
People (110/296 = 0.372)	陳寧(Cheng Ning)	56
	凡爾納(Jules Gabriel Verne)	19
	劉山永(Liu Shanyong)	14
	少兒(children)	11
	英雄(hero)	10

related to the "people" category in mainland China are typically related to famous people; those in Taiwan are related to users. Notably, keyness words for "location" in the two regions were mostly associated with their northern areas. Taipei is a northern city in Taiwan, and both Beijing and Changchun are in the north of mainland China. Moreover, the headquarters of the two newspapers are both in these northern regions, and communications technologies were relatively underdeveloped in 1995–1999.

4.3.2 Keyness words in 2000 to 2004

As presented in Table 4, librarians in Taipei County were frequently mentioned in news articles due to discussions regarding a new library building. The word "時論(public opinion)" indicated that the librarian-related news includes up-to-date opinions of the general public. The "公所(administrative office)" is the government office of a village or town. The most keyness words in the Taiwan collection belong to the "people" category. "蘿拉(Laura Welch Bush),"

臺灣大學學術期刊資料庫

Table 4. Keyness Words of Taiwan from 2000 to 2004

Category & percentage	Keyness word	CF in TW
Thing (56/649 = 0.086)	時論(public opinion)	56
Location (133/649 = 0.205)	台灣(Taiwan)	97
	北縣(Taipei County)	36
Organization (182/649 = 0.280)	公共圖書館(public library)	90
	國家圖書館(National Central Library)	47
	公所(administrative office)	45
People (278/649 = 0.428)	圖書館員(librarian)	152
	志工(volunteers)	48
	蘿拉(Laura)	39
	布希(Bush)	39

the wife of the 43rd President of the United States, was a librarian and obtained substantial funding for libraries after becoming the first lady. Libraries and librarians in the north part of Taiwan were highly reported over this five-year period.

Table 5 indicates that the major category in mainland China was also "people"; as for 1995–1999, these were famous individuals. "楊昌濟(Yang Changji)," an educator, supported the New Culture Movement (新文化運動) and was the teacher of Mao Zedong. "屈萬里(Qu Wanli)" was a bibliographer and the librarian of Shandong Provincial Library. "敦煌(Dunhuang)" refers to ancient books that are difficult to preserve. This organizational category contains the least keyness words in mainland China [i.e., "上海圖書館(Shanghai Library)" and "連隊(troops)]." Shanghai Library is a major public library that has offered an online reading consultation service since 2000. "連隊(troop)" appears in articles describing a donation of books by the librarians of Beijing Chaoyang district to construct a small troop library.

In comparison with Taiwan, keyness words for mainland China focused on the preservation of historical materials. For example, to protect and develop "藏書樓(collection houses)," library curators and experts held a seminar in Zhengjiang. As presented in Tables 4 and 5, numerous keyness words describe people in both collections; however, differences are still apparent between the two regions. The words used in Taiwan are more general, whereas more detailed descriptions of people appear in news articles from mainland China such as an interview with "辛希孟(Xin Ximeng)," a professor of LIS. "志工(volunteers)" often appears in the Taiwan collection; the mainland China collection includes the similar word "志願者(volunteers)" instead. The results reveal that Taiwanese libraries frequently invited volunteers to collaborate with librarians; this practice was uncommon in mainland China between 2000 and 2004.

臺灣大學學術期刊資料庫

Table 5. Keyness Words of Mainland China from 2000 to 2004

Category & percentage	Keyness word	CF in MC
Organization (33/411 = 0.083)	上海圖書館(Shanghai Library)	22
	連隊(troops)	11
Action (58/411 = 0.141)	遺書(ancient book)	31
	修復(restore)	14
	開展(carry out)	13
Location (74/411 = 0.180)	敦煌(Dunhuang)	39
	大連(Dalian)	13
	廣州(Guangzhou)	11
	舟山(Zhoushan)	11
Thing (77/411 = 0.187)	藏書樓(collection house)	50
	西遊記(*Journey to the West*)	14
	行業(industry)	13
People (169/411 = 0.411)	章士釗(Zhang Shizhao)	48
	楊昌濟(Yang Changji)	31
	辛希孟(Xin Ximeng)	23
	屈萬里(Qu Wanli)	21
	葉星球(Ye Xinqiu)	19
	巴金(Ba Jin)	15
	戰士(warrior)	12

4.3.3 Keyness words in 2005 to 2009

Keyness words for Taiwan and mainland China in the period from 2005 to 2009 are presented in Tables 6 and 7, respectively. Here, the term "cat" was assigned to the "people" category because the mentioned cat was also regarded as a librarian at the Spencer Public Library in Iowa, USA. The keyness words were spread evenly among the five categories from 2005 to 2009. However, many words (28%) were connected to books, suggesting a strong relationship between books and librarians. Librarians in Taiwan recommend picture books to children, held activities about them, and even created a picture book library in Taipei (紙風車繪本館, The House of Paperwindmill's Children Books); the mainland China collection did not include picture books from 1995 to 2019. In fact, the earliest LIS paper mentioning children picture books in mainland China was published in 2010. Thus, if mainland Chinese librarians intend to develop picture book activities, they may benefit from the experience of Taiwanese librarians.

臺灣大學學術期刊資料庫

Table 6. Keyness Words of Taiwan from 2005 to 2009

Category & percentage	Keyness word	CF in TW
Organization (131/1141 = 0.115)	分館(branch library)	61
	公共圖書館(public library)	36
	國小(elementary school)	34
Action (172/1141 = 0.151)	看書(read books)	54
	都會(will)	42
	開卷(open a book)	42
	不會(will not)	34
Thing (256/1141 = 0.224)	本書(the book)	86
	一本(a volume)	61
	好書(good book)	39
	優格(yogurt)	36
Location (274/1141 = 0.240)	台灣(Taiwan)	130
	台北(Taipei)	62
	北縣(Taipei County)	42
	鄉鎮(villages and towns)	40
People (308/1141 = 0.270)	圖書館員(librarian)	150
	志工(volunteer)	80
	貓(cat)	40
	書人(book man)	38

In Table 7, the category "thing" is the largest category for news articles from mainland China, but the keyness words were different from those in Taiwan. The keyness words for Taiwan were related to books and user services. By contrast, the keyness words for mainland China are related to the protection and exhibition of historical materials. Shanghai Library completed the "*Catalogue of Chinese National Genealogy*"(《中國家譜總目》) in 2007 and presented many rare tablets in 2005. Famous people, including library experts and librarians, are more frequently described in mainland Chinese news articles. In Taiwanese news articles, librarians typically act as supporting reporters.

4.3.4 Keyness words in 2010 to 2014

In Table 8, "楠" is a word from a person's name: 肯楠(Kennan). "新共和(*The New Republic*)" is an American magazine. "艾倫(Paul Allen)" is a Microsoft co-founder who was born into a family of librarians. These keyness words indicate that reading promotion events

臺灣大學學術期刊資料庫

Table 7. Keyness Words of Mainland China from 2005 to 2009

Category & percentage	Keyness word	CF in MC
Time (38/842 = 0.045)	五四(May Fourth)	38
Action (88/842 = 0.105)	再造(reforge)	62
	勤工儉學(work-study)	26
Location (86/842 = 0.102)	網上(online)	53
	紅樓(Red House)	33
Organization (101/842 = 0.120)	北大(Peking University)	73
	圖書館(library)	28
People (212/842 = 0.252)	孟二冬(Meng Erdong)	65
	袁熙坤(Yuan Xikun)	42
	李鐵強(Li Tieqiang)	36
	金一南(Jin Yinan)	35
	李大釗(Li Dazhao)	34
Thing (317/842 = 0.377)	善本(rare book)	118
	家譜(genealogy)	49
	中華(Chinese)	36
	科舉(imperial examination)	33
	戰略(strategy)	28
	書(book)	27
	碑帖(tablet)	26

and book recommendations are the main focus of Taiwanese librarians. For the period from 2010 to 2014, the category with the most words was "thing"; those keyness words were related to book introductions. For example, to encourage people to face challenges, a new book about hearing-impaired athletes was introduced and some of them became librarians after retrying. The category with the least words was "time," and the word "一年(annually)" was used as a unit period to evaluate the performance of librarians. The term "電子書(electric book)" indicates that Taiwanese librarians began to discuss challenges of digital books and the future of libraries during this period. Subsequently, digital books were included in library collections. Taiwan's keyness words center on the services of recommendations and promotions of both physical and digital books.

Table 9 presents the keyness words for mainland China. The "thing" category also had the most words. "家譜(genealogy)" was related to a book written by librarians at Shanghai

臺灣大學學術期刊資料庫

Table 8. Keyness Words of Taiwan from 2010 to 2014

Category & percentage	Keyness word	CF in TW
Time (25/611 = 0.041)	一年(annually)	25
Organization (48/611 = 0.079)	國小(elementary school)	25
	分館(branch library)	23
Action (75/611 = 0.123)	看書(read books)	28
	都會(will)	25
	開卷(open a book)	22
Location (107/611 = 0.175)	台灣(Taiwan)	71
	台北(Taipei)	36
People (171/611 = 0.280)	圖書館員(librarian)	70
	志工(volunteer)	56
	艾倫(Paul Allen)	25
	楠(Kennan)	20
Thing (185/611 = 0.308)	台(platform)	39
	電子書(electric book)	36
	卅(thirty)	25
	聽障(hearing impairment)	25
	一本(a volume)	22
	新共和(*The New Republic*)	19
	氣味(smell)	19

Library about genealogies; Shanghai Library has the largest collection of Chinese genealogies. “立法(legislate)” is related to the expansion of librarians’ duties in mainland China since the legislative decision service department of the National Library began providing information services to the government in 1999. “讀書會(Book club)” refers to the regular workshops for reading services in Beijing community libraries; similar activities were common for many years in Taiwan. “中國館(China National Pavilion)” is the China National Pavilion at the World Expo in Shanghai. Notably, mainland China began to open more library resources to the general public in the 2010s. A news article from 2011 reported that public libraries in Beijing initiated a resource-sharing project that enabled patrons to borrow and return books with shared cards at 170 public libraries. Therefore, librarians in mainland China were confronted with the new challenge of developing digital sources and branch libraries for building a public cultural service system.

臺灣大學學術期刊資料庫

Table 9. Keyness Words of Mainland China from 2010 to 2014

Category & percentage	Keyness word	CF in MC
Others (37/704 = 0.053)	基本(basic)	37
Location (54/704 = 0.077)	深圳(Shenzhen)	29
	陝西(Shanxi)	25
Organization (85/704 = 0.121)	高校(university)	85
Action (126/704 = 0.179)	立法(legislate)	45
	開展(carry out)	43
	共用(share with)	38
People (181/704 = 0.257)	人民(people)	39
	王鵬(Wang Peng)	32
	斯特林堡(August Strindberg)	30
	人才(talent)	29
	書記(secretary)	27
	戴逸(Dai Yi)	24
Thing (221/704 = 0.314)	家譜(genealogy)	50
	中華(Chinese)	43
	中國館(China National Pavilion)	35
	讀書會(book club)	35
	體系(system)	32
	平臺(platform)	26

4.3.5 Keyness words in 2015 to 2019

In Table 10, “李哲洋” is Li Zhe-Yang, a Taiwan musician. “那麼(like that)” indicates a comparison describing degree. “邦德(Julian Bond)” is an African American politician; his mother is a librarian. The keyness words in the corpus reveal that librarians in Taiwan focus on events promoting reading for children and vulnerable groups such as homework tutoring for schoolchildren, parent-child do-it-yourself activities, schoolchildren staying overnight in the library, and visually impaired reading clubs. The “people” category had many elements in this period and includes numerous famous people. “親子(parent-child)” indicates a difference in the targets of librarian reading services between Taiwan and mainland China. Parent-child reading clubs and book exchanges increase interactions between patrons and librarians. The locations include “新北市/新北(New Taipei City),” “桃園(Taoyuan)” and “基隆(Keelung).” Notably, librarians at National Taiwan Library are working

臺灣大學學術期刊資料庫

Table 10. Keyness Words of Taiwan from 2015 to 2019

Category & percentage	Keyness word	CF in TW
Others (12/455 = 0.026)	那麼(like that)	12
Action (17/455 = 0.037)	夜宿(stay overnight)	17
Thing (54/455 = 0.119)	本書(the book)	20
	視障(visually impairment)	20
	繪本(picture book)	14
Organization (73/455 = 0.160)	分館(branch library)	56
	新北市立圖書館(New Taipei City Library)	17
People (130/455 = 0.286)	邦德(Julian Bond)	33
	學童(schoolchildren)	29
	圖書館員(librarian)	25
	親子(parent-child)	16
	李哲洋(Li Zhe-Yang)	15
	維奇(Svetlana Alexievich)	12
Location (169/455 = 0.371)	新北市/新北(New Taipei City)	64
	台灣(Taiwan)	46
	桃園(Taoyuan)	17
	基隆(Keelung)	16
	台北(Taipei)	14
	飛驒市(Hida City)	12

on books for the "視障(visually impairment)" and an app for digital audiobooks.

"同志(comrade)" is a word used to refer to a person in mainland China. "田青(Tian Qing)," is the librarian of the Central Research Institute of Culture and History and is also an expert on Buddhist music. "遺產(heritage)" typically refers to intangible cultural heritage in the corpus in Table 11. In the period from 2015 to 2019 in mainland China, the protection of traditional cultural treasures was the primary focus of librarians. News articles refer to librarians as the guardians of this intangible cultural heritage. However, protecting ancient books is often difficult because of the huge number of books involved and the difficulties developing appropriate procedures to protect these books. The 80th anniversary of the "Long March" of the Communist Party of China was in 2015. Libraries in mainland China initiated activities about the long venture of "紅軍(the Red Army)." Moreover, 2018 was also the 120th anniversary of Peking

臺灣大學學術期刊資料庫

Table 11. Keyness Words of Mainland China from 2015 to 2019

Category & percentage	Keyness word	CF in MC
Location (38/1108 = 0.034)	海外(oversea)	38
Others (41/1108 = 0.037)	協同(cooperate)	41
Organization (89/1108 = 0.080)	北京大學(Peking University)	48
	紅軍(the Red Army)	41
Action (223/1108 = 0.201)	兼職(do a part-time job)	53
	開展(carry out)	48
	開發(develop)	44
	實現(reach)	40
	打工(work)	38
People (262/1108 = 0.236)	同志(comrade)	71
	村民(villager)	54
	父親(father)	52
	田青(Tian Qing)	48
	志願者(volunteer)	37
Thing (455/1108 = 0.411)	古籍(ancient book)	180
	村落(village)	150
	項目(project)	49
	民族(nation)	39
	遺產(heritage)	37

University. Numerous important figures in the modern Chinese revolution worked for "北京大學 (Peking University)." For example, Mao Zedong and Li Dazhao both worked at Peking University as librarians.

Taiwan's keyness words indicate that librarians were more focused on parent-child relationships and visually impaired people than librarians in mainland China were in 2015–2019; mainland Chinese librarians focused on protecting village cultures and library cooperation in various areas. Overall, Taiwanese librarians focus on service quality, and mainland Chinese librarians focus on cultural development and resource-sharing.

4.4 Data analysis for sentiment values

In this section, we analyzed sentiment tendencies for librarians and compared the results between mainland China and Taiwan. The sentiment values for sentences reflect the sentiment tendencies for librarians. Measurements of sentence sentiment are discussed in Section

臺灣大學學術期刊資料庫

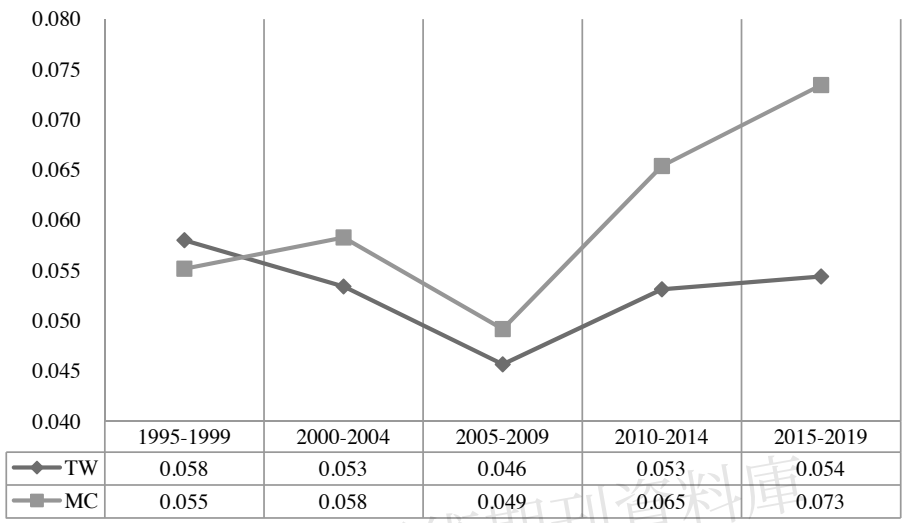
3. We normalized all of the sentiment values for sentences between -1 and 1 by using the maximal values. Figure 1 presents the changes in sentiment values over the five periods.

The sentiment values for both mainland China and Taiwan were positive from 1995 to 2019; the values were slightly higher than the neutral value. For mainland China, the maximal sentiment value of 0.072 was in 2015–2019 and the minimal sentiment value was 0.055 in 1995–1999. Sentiment in Taiwan was slightly higher in mainland China in the first five years but was lower in the following periods. Sentiment in both regions was at a minimum in 2005–2009. For mainland China, approximately 9% of reports were negative in 2005–2009; for the other periods, 4%, 7.9%, 3%, and 3% of reports were negative in 1995–1999, 2000–2004, 2010–2014, and 2015–2019, respectively. For Taiwan, approximately 10% of reports were negative in 2005–2009; for the other periods, 5%, 6%, 9%, and 4% of reports were negative in 1995–1999, 2000–2004, 2010–2014, and 2015–2019, respectively. The values for mainland China change more compared with those for Taiwan.

An independent sample *t* test and one-way analysis of variance (ANOVA) were used to determine the significance of the results. The independent sample *t* test revealed that the trends of sentiment values of mainland China and Taiwan differed significantly (p = .003). Next, we investigated which periods were significantly different by using one-way ANOVA; significant differences were observed for 2005–2009, 2010–2014, and 2015–2019 in mainland China and 2005–2009 in Taiwan.

5. Conclusions

As previously described, the results were analyzed using five factors: appearance, character, behavior, competence, and sentiment. However,



	1995-1999	2000-2004	2005-2009	2010-2014	2015-2019
TW	0.058	0.053	0.046	0.053	0.054
MC	0.055	0.058	0.049	0.065	0.073

Figure 1. Sentiment Values in Each Period

臺灣大學圖書館期刊資料庫

words related to character were rarely used in newspapers; this result may be due to attempts by newspapers to maintain a neutral tone. Most words were related to the behavior and competence of librarians. Although some words were related to physical representations and attitudes (such as female librarian or simple job) and could reflect stereotypes of librarians, news articles typically avoid using these words.

5.1 Similarities and differences between the two regions' media images

After manually examining the words and their contexts, similarities and differences between media images of librarians in mainland China and Taiwan were identified by using shared words and keyness words. Librarianship is an important part of cultural development, and librarians of two regions had similar goals for read services. Because librarians typically appear alongside "books," "readers," and "services" in news articles and because those words could be related to appearance or behavior in news articles from both regions, the librarian profession involves handling interactions between information and users. Moreover, reading services are important information dissemination channels. However, other roles were also identified. People mentioned in news articles from mainland China were librarians or liberal arts scholars, whereas people reported in news articles from Taiwan had more diverse backgrounds (e.g., the first lady of the United States). Some librarians in mainland China are also authors, such as authors of books discussing ancient materials. By contrast, librarians in Taiwan usually play the role of information brokers by recommending books to patrons, especially to children. Thus, librarians in Taiwan organize various reading activities. In mainland China, exhibitions of historical materials are more common. Overall, librarians in both mainland China and Taiwan are connected to both books and people. However, librarians in mainland China have an image as academics and historians, whereas librarians in Taiwan have an image of providing high-quality services for readers.

5.2 Changes in each region's media image

The evolution of media images of librarians could be determined by observing high-frequency words and keyness words. For locations reported in news articles, librarians in northern Taiwan (Keelung and Taipei) are mentioned frequently, whereas news articles of mainland China mentioned various locations equally after 1999. However, Beijing was mentioned more frequently between 1995 and 1999. Children have been the most important patrons of libraries in Taiwan since 1995, and the visually impaired have become another important group over the last decade. The high-frequency words revealed that volunteers have played an increasingly important role in Taiwan's library system since 2005. Library services in mainland China have not evolved as clearly. However, students in universities at mainland China frequently appear in news articles from 2005 to 2009 because many libraries provided part-time jobs for students with financial difficulties. Reading services have always been important at Taiwan libraries, and the books mentioned in news articles have changed from picture books (2005 to 2009) and digital books (2010 to 2014) to braille books (2015 to 2019). By contrast, ancient books have been a focus

臺灣大學學術期刊資料庫

for libraries in mainland China with key topics ranging from the restoration of the Dunhuang documents (2000 to 2004) and Chinese genealogy (2005 to 2014) to develop creative industries of ancient books (2015 to 2019).

In brief, we can observe changes in behavior and competence rather than those in appearance or character in news articles. That is, news articles typically reveal the evolution of tasks rather than of the external bearing of librarians. Librarians in both regions focus on reading services for children. However, in the last ten years, reading services for the visually impaired have become more common in Taiwan. In mainland China, the focus has shifted from protecting ancient materials to creatively promoting them.

5.3 Indication of sentiment changes

The sentiment values for librarians on both sides of Taiwan Strait were slightly positive. However, some differences were noted. The sentiment value for news articles related to librarians in mainland China is higher than that in Taiwan for all periods except for the period from 1995 to 1999. The least positive sentiment value for both regions was for the period from 2005 to 2009. Sentiment values in mainland China increased gradually, but those in Taiwan remained almost the same.

A few words were sometimes used to introduce other negative sentimental words; these included "都(all)," "那(that)," "會(will)." We observed that these words were rarely used in news articles from mainland China. Moreover, words such as "這(this)," "再(again)," and "還(also)" were rarely used in Taiwanese news articles. Thus, news articles related to libraries and librarians from both sides of Taiwan Strait tended to have positive sentiment.

5.4 Limitation and recommendations for future research

Although the results of this study are encouraging, the results should be applied with caution. Note that only one major newspaper in each region was used for analysis. Better conclusions could be obtained based on news articles published by multiple news agencies.

A model for analyzing the media image of organizations was formed in this research. The media images of organizations can be identified based on words with four factors with three dimensions: the temporal axis, spatial axis, and source axis. A media image analysis system could also be designed based on this model to enable organizations to quickly respond to public relations issues. We also had difficulty analyzing related tests. Initially, we intended to analyze more newspapers in Chinese-speaking areas, such as Hong Kong, Macao, and Singapore. However, collecting these newspapers is currently difficult. A framework and method of word classification also require further exploration to enhance the analysis depth. Another possible issue is the choice of dictionary for sentiment analysis. The BosonNLP dictionary was used to assign a sentiment value to each word in this study. However, BosonNLP was developed for simplified Chinese; traditional Chinese words may not be found in BosonNLP due to differences in word usage. Future work should develop a common sentiment dictionary for both simplified Chinese and traditional Chinese.

臺灣大學學術期刊資料庫

Notes

Note 1 Please visit http://www.nlc.cn/pcab/ztzl/gjbhzcfg/gfxwj/201511/t20151104_107984.htm for detailed information.

References

Abbasi, A., & Chen, H. (2008). CyberGate: A design framework and system for text analysis of computer-mediated communication. *Mis Quarterly, 32*(4), 811-837. doi: 10.2307/25148873

Al-Daihani, S. M., & Abrahams, A. (2016). A text mining analysis of academic libraries' tweets. *The Journal of Academic Librarianship, 42*(2), 135-143. doi: 10.1016/j.acalib.2015.12.014

Black, S. (1994). Checking out the librarian: The depiction of librarians and libraries on prime time television. *Popular Culture in Libraries, 1*(4), 35-62. doi: 10.1300/J117v01n04_02

Boson. (n.d.). *Product—API*. Retrieve from http://static.bosonnlp.com/product (in Chinese)

Cheng, X. L. (2016). Empirical research on the marketing case of Xiamen University Library. *Library & Information, 5*, 62-69. doi: 10.11968/tsyqb.1003-6938.2016095 (in Chinese)

Cheng, W. T., & Wu, J. H. (2021). Topic model and sentiment analysis of user opinions on preschool children's digital reading. *Library Development, 5*, 1-22. Retrieved from https://kns.cnki.net/kcms/detail/23.1331.G2.20210218.1528.022.html (in Chinese)

China Times. (n.d.). *About*. Retrieved from https://static.chinatimes.com/pages/customer-about.htm?chdtv (in Chinese)

Choi, Y. (2019). Finding "just right" books for children: Analyzing sentiments in online book reviews. *The Electronic Library, 37*(3), 563-576. doi: 10.1108/EL-01-2019-0018

Chueh, H. C., & Chen, K. H. (2016). CORPRO: A Chinese language corpus tool and a case study of media representation of organic agriculture. In J. Hsiang (Ed.), *Digital humanities and craft: Technological change* (pp. 285-313). Taipei: National Taiwan University Press. (in Chinese)

Chujo, K., & Utiyama, M. (2006). Selecting level-specific specialized vocabulary using statistical measures. *System, 34*(2), 255-269. doi: 10.1016/j.system.2005.12.003

Dong, Z. D., & Dong, Q. (2003). HowNet-a hybrid language and knowledge resource. In C. Zong (Ed.), *2003 International Conference on Natural Language Processing and Knowledge Engineering* (pp. 820-824). Beijing, China: IEEE. doi: 10.1109/NLPKE.2003.1276017

Dong, Z., Dong, Q., & Hao, C. (2010). HowNet and its computation of meaning. In *Proceedings of the 23rd International Conference on Computational Linguistics: Demonstrations* (pp. 53-56). Stroudsburg, PA: Association for Computational Linguistics. Retrieved from https://aclanthology.org/C10-3014.pdf

Edgar, N. L. (1976). Environment: Image of librarianship in the media. In S. L. Jackson,

臺灣大學學術期刊資料庫

E. B. Herling, & E. J. Josey (Eds.), *A century of service: Librarianship in the United States and Canada* (pp. 303-320). Chicago, IL: American Library Association. Retrieved from https://search.proquest.com/docview/57144508?accountid=14229

Feng, X. T. (2010). The construction and demonstration of the only child's negative image in media. *Sociological Studies, 3*, 177-198. (in Chinese)

Gerolimos, M., Malliari, A., & Iakovidis, P. (2015). Skills in the market: An analysis of skills and qualifications for American librarians. *Library Review, 64*(1/2), 21-35. doi: 10.1108/LR-06-2014-0063

Gimenez, P. G., & Pedreno, M. H. (2002). The social image of libraries in the print and digital media. *Anales De Documentacion, 5*, 177-196. Retrieved from https://search.proquest.com/docview/57701081?accountid=14229

Golder, S. A., & Macy, M. W. (2011). Diurnal and seasonal mood vary with work, sleep, and daylength across diverse cultures. *Science, 333*(6051), 1878-1881. doi: 10.1126/science.1202775

Grabe, S., Ward, L. M., & Hyde, J. S. (2008). The role of the media in body image concerns among women: A meta-analysis of experimental and correlational studies. *Psychological Bulletin, 134*(3), 460-476. doi: 10.1037/0033-2909.134.3.460

Guo, N. E. (2015). [Xiao yi wai guo qing shao nian er tong wen xue zhong tu shu guan yuan ke ban yin xiang zhi cheng yin]. *Journal of Language & Literature Studies, 17*, 85-86. (in Chinese)

Habimana, O., Li, Y., Li, R., Gu, X., & Yu, G. (2020). Sentiment analysis using deep learning approaches: An overview. *Science China Information Sciences, 63*(1), 111102:1-111102:36. doi: 10.1007/s11432-018-9941-6

Hammett, D. (2014). Tourism images and British media representations of South Africa. *Tourism Geographies, 16*(2), 221-236. doi: 10.1080/14616688.2012.762688

Harrison, J., & Beraquet, V. (2010). Clinical librarians, a new tribe in the UK: Roles and responsibilities. *Health Information & Libraries Journal, 27*(2), 123-132. doi: 10.1111/j.1471-1842.2009.00862.x

He, W., Zha, S., & Li, L. (2013). Social media competitive analysis and text mining: A case study in the pizza industry. *International Journal of Information Management, 33*(3), 464-472. doi: 10.1016/j.ijinfomgt.2013.01.001

Helms, B. L. (2006). *Reel librarians: The stereotype and technology* (Master's thesis, University of North Carolina at Chapel Hill). Retrieved from https://cdr.lib.unc.edu/concern/masters_papers/3r074z398

Hsieh, J. L., & Yang, B. C. (2018). A content analysis and comparison of typhoon news in early and recent periods based on the text-mining approach. *Journal of Educational Media & Library Sciences, 55*(3), 285-318. doi: 10.6120/JoEMLS.201811_55(3). 0022.RS.BM

Hsieh, T. Y. (2016). *Differences between movie and real world librarian image* (Master's

臺灣大學學術期刊資料庫

thesis, National Chun Hsing University, Taichung, Taiwan). Retrieved from https://www.airitilibrary.com/Publication/alDetailedMesh1?DocID=U0005-2107201614333400 (in Chinese)

Ke, H. R., & Sie, S. H. (2019). Exploring the research trends of Library and Information Science, 2006-2015. *Journal of Library & Information Science, 45*(1), 65-96. doi: 10.6245/JLIS.201904_45(1).0003 (in Chinese)

Kuo, J. J., & Zhang, Y. J. (2013). A study on library users' satisfaction evaluation using sentimental analysis. *Journal of Library & Information Science Research, 7*(2), 153-197.

Li, N, (2010). The characteristics of "male librarians" in Japan. *Library Journal, 29*(2), 66-67+41. doi: 10.13663/j.cnki.lj.2010.02.020 (in Chinese)

Liu, S. (2013). [Ao Men bo cai ye xing xiang wen ben fen xi]. *Competitive Intelligence, 9*(4), 22-27. (in Chinese)

Liu, B. (2015). *Sentiment analysis: Mining opinions, sentiments, and emotions*. New York, NY: Cambridge University Press.

Liu, Z. H., & Gao, D. (2009). The media image of libraries' in past 60 years: Statistical investigation on reports on libraries in People's Daily. *Library Tribune, 29*(4), 1-5. (in Chinese)

Luthmann, A. (2007). Librarians, professionalism and image: Stereotype and reality. *Library review, 56*(9), 773-780. doi: 10.1108/00242530710831211

Maceli, M. (2015). What technology skills do developers need? A text analysis of job listings in Library and Information Science (LIS) from jobs.code4lib.org. *Information Technology & Libraries, 34*(3), 8-21. doi: 10.6017/ital.v34i3.5893

Magerman, T., Van Looy, B., & Song, X. (2009). Exploring the feasibility and accuracy of Latent Semantic Analysis based text mining techniques to detect similarity between patent documents and scientific publications. *Scientometrics, 82*(2), 289-306. doi: 10.1007/s11192-009-0046-6

Majid, S., & Haider, A. (2008). Image problem even haunts hi-tech libraries: Stereotypes associated with library and information professionals in Singapore. *Aslib Proceedings, 60*(3), 229-241. doi: 10.1108/00012530810879105

Maynard, S., & McKenna, F. (2005). Mother Goose, Spud Murphy and the Librarian Knights: Representations of librarians and their libraries in modern children's fiction. *Journal of Librarianship & Information Science, 37*(3), 119-129. doi: 10.1177/0961000605057475

McCombs, M., & Reynolds, A. (2002). News influence on our pictures of the world. In J. Bryant, D. Zillmann, & M. B. Oliver (Eds.), *Media effects: Advances in theory and research* (pp. 1-18). doi: 10.4324/9781410602428

McCombs, M., Juan, P. L., Lopez-Escobar, E., & Rey, F. (1997). Candidate images in Spanish elections: Second-level agenda-setting effects. *Journalism & Mass Communication Quarterly, 74*(4), 703-717. Retrieved from

臺灣大學學術期刊資料庫

https://search.proquest.com/docview/216927209?accountid=14229

Muhammad, S. H. (2019). An overview of sentiment analysis approaches. In R. Rua, V. Silva, S. Muhammad, & F. Duarte (Eds.), *MAPiS 2019—First MAP-i Seminar, Proceedings* (pp. 65-70). Aveiro, Portugal: UA Editora, Universidade de Aveiro.

Nasukawa, T., & Nagano, T. (2001). Text analysis and knowledge mining system. *IBM Systems Journal, 40*(4), 967-984. doi: 10.1147/sj.404.0967

Nelson, A. J., & Irwin, J. (2014). "Defining What We Do—All Over Again": Occupational identity, technological change, and the librarian/Internet-search relationship. *Academy of Management Journal, 57*(3), 892-928. doi: 10.5465/amj.2012.0201

NLPIR1. (2021). [NLPIR da shu ju yu yi zhi neng fen xi ping tai]. Retrieved from http://www.nlpir.org/wordpress/2021/04/12/%E5%A4%A7%E6%95%B0%E6%8D%AE%E8%AF%AD%E4%B9%89%E6%99%BA%E8%83%BD%E5%88%86%E6%9E%90%E5%B9%B3%E5%8F%B0/ (in Chinese)

Oakes, M. P. (2008). Measures from information retrieval to find the words which are characteristic of a corpus. In B. Lewandowska-Tomaszczyk (Ed.), *Corpus linguistics, computer tools, and applications—State of the art: PALC 2007* (pp. 127-138). Berlin, Germany: Peter Lang Verlag.

Pan, X., Yan, E., Cui, M., & Hua, W. (2019). How important is software to library and information science research? A content analysis of full-text publications. *Journal of Informetrics, 13*(1), 397-406. doi: 10.1016/j.joi.2019.02.002

Pang, X. L. (2018). [Dian jing zhi bo ping tai de mei jie xing xiang yan jiu—Yi "Hu Ya Zhi Bo" wei li]. *Journal of News Research, 9*(23), 49-50. (in Chinese)

People's Daily. (n.d.). *Introduction*. Retrieved from http://www.people.com.cn/GB/50142/104580/index.html (in Chinese)

Pojanapunya, P., & Todd, R. W. (2018). Log-likelihood and odds ratio: Keyness statistics for different purposes of keyword analysis. *Corpus Linguistics & Linguistic Theory, 14*(1), 133-167. doi: 10.1515/cllt-2015-0030

Qiu, J. H. (1992). [Shi lun xin xing tu shu guan yuan xing xiang de su zao]. *Research on Library Science, 6*, 30-31. (in Chinese)

Raamkumar, A. S., Tan, S. G., & Wee, H. L. (2020). Measuring the outreach efforts of public health authorities and the public response on Facebook during the COVID-19 pandemic in early 2020: Cross-country comparison. *Journal of Medical Internet Research, 22*(5), e19334. doi: 10.2196/19334

Radford, M. L., & Radford, G. P. (1997). Power, knowledge, and fear: Feminism, foucault, and the stereotype of the female librarian. *The Library Quarterly, 67*(3), 250-266. doi: 10.1086/629951

Robinson, L. T. (2006). Curmudgeons and dragons? A content analysis of the Australian print media's portrayal of the information profession 2000 to 2004. *Library & Information Science Research*

臺灣大學學術期刊資料庫

Electronic Journal, 16(2). doi: 10.32655/LIBRES.2006.2.1

Salton, G., & Buckley, C. (1988). Term-weighting approaches in automatic text retrieval. *Information Processing & Management, 24*(5), 513-523. doi: 10.1016/0306-4573(88)90021-0

Scott, M. (1997). PC analysis of key words—and key key words. *System, 25*(2), 233-245. doi: 10.1016/S0346-251X(97)00011-0

Scott, M. (2004). *Oxford WordSmith Tools: Version 4.0*. Oxford, UK: Oxford University Press.

Shaw, L. (2010). Representations of librarianship in the UK press. *Aslib Proceedings: New Information Perspectives, 62*(6), 554-569. doi: 10.1108/00012531011089676

Short, M. (2019). Text mining and subject analysis for fiction; or, using machine learning and information extraction to assign subject headings to dime novels. *Cataloging & Classification Quarterly, 57*(5), 315-336. doi: 10.1080/01639374.2019.1653413

Singh, V. K., Chayko, M., Inamdar, R., & Floegel, D. (2020). Female librarians and male computer programmers? Gender bias in occupational images on digital media platforms. *Journal of the Association for Information Science & Technology, 71*(11), 1281-1294. doi: 10.1002/asi.24335

Stone, P. J., Bales, R. F., Namenwirth, J. Z., & Ogilvie, D. M. (1962). The general inquirer: A computer system for content analysis and retrieval based on the sentence as a unit of information. *Behavioral Science, 7*(4), 484-498. doi: 10.1002/bs.3830070412

Tseng, Y. H., Lin, C. J., & Lin, Y. I. (2007). Text mining techniques for patent analysis. *Information Processing & Management, 43*(5), 1216-1247. doi: 10.1016/j.ipm.2006.11.011

Turney, P. D. (2002). *Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews*. Retrieved from https://arxiv.org/abs/cs/0212032

Van Dijk, T. A. (1995). Power and the news media. In D. Paletz (Ed.), *Political communication and action* (pp. 9-36). Cresskill, NJ: Hampton Press.

Vassilakaki, E., & Moniarou-Papaconstantinou, V. (2016). How public library users perceive the information professional: Is the image transforming? *New Library World, 117*(7/8), 449-463. doi: 10.1108/NLW-11-2015-0090

Wang, H., Tang, Y., & Knight, C. (2010). Contemporary development of academic reference librarianship in the United States: A 44-year content analysis. *The Journal of Academic Librarianship, 36*(6), 489-494. doi: 10.1016/j.acalib.2010.08.004

Watkins, C. (1968). Role of the librarian. *Bulletin of the Medical Library Association, 56*(1), 36-40.

Wells, J. A. (2013). The female librarian in film: Has the image changed in 60 years? *School of Information Student Research Journal, 3*(2). doi: 10.31979/2575-2499.030202

Wilbur, W. J., Hazard, G. F., Jr., Divita, G., Mork, J. G., Aronson, A. R., & Browne, A. C. (1999). Analysis of biomedical text for chemical names: A comparison of three methods. In N. M. Lorenzi (Ed.), *Proceedings of the AMIA*

臺灣大學學術期刊資料庫

Symposium (pp. 176-180). Philadelphia, PA: Hanley & Belfus.

WordSmith Tools Help. (n.d.). *Definition of keyness*. Retrieved from https://lexically.net/downloads/version8/HTML/keyness_definition.html?q=key-ness+defined

Wu, C. Y. (2002). *The construction of the image of the female politician* (Unpublished master's thesis). Department of Communication Management, Nanhua University, Chiayi, Taiwan. (in Chinese)

Wu, S. C. (2003). Discussing the librarian and libraries' image problem by analyzing the content of mass media. *Journal of Educational Media & Library Sciences, 40*(4), 549-565. (in Chinese)

Xie, F., & Xu, C. M. (2018). Marketing methods from the Libraries Transform Campaign in the United States. *Library Development, 3*, 81-86. (in Chinese)

Xu, J. L. (2011). Analysis and reflection of images of librarians in Chinese movie and TV works. *Library Development, 2*, 107-110. (in Chinese)

Yang, C. H. (2006). *[Guo Min Dang dang zhu xi hou xuan ren zhi mei ti xing xiang fen xi—Yi 2005 nian Wang Jin Ping yu Ma Ying Jiu xuan ju wei li]* (Master's thesis, National Chengchi University, Taipei, Taiwan). Retrieved from http://thesis.lib.nccu.edu.tw/record/#G0922520141 (in Chinese)

Yang, T. J., & Ling, M. (2018). Comparison of English titles' structure features of Chinese and international journal articles in Library& Information Science. *Journal of Intelligence, 37*(1), 76-81. doi: 10.3969/j.issn.1002-1965.201703398 (in Chinese)

Zagibalov, T., & Carroll, J. A. (2008). Automatic seed word selection for unsupervised sentiment classification of Chinese text. In D. Scott & H. Uszkoreit (Eds.), *Proceedings of the 22nd International Conference on Computational Linguistics (Coling 2008)* (pp. 1073-1080). Manchester, UK: Coling 2008 Organizing Committee. doi: 10.3115/1599081.1599216

Zhang, Y. U. (2012). *A study on library users' satisfaction evaluation using sentimental analysis* (Unpublished master's thesis). Graduate Institute of Library & Information Science, National Chung Hsing University, Taichung, Taiwan. (in Chinese)

Zhang, H. P., & Shang, J. Y. (2019). NLPIR-Parser: An intelligent semantic analysis toolkit for big data. *Corpus Linguistics, 6*(1), 87-104. (in Chinese)

Zhang, L., Wang, S., & Liu, B. (2018). Deep learning for sentiment analysis: A survey. *Wiley Interdisciplinary Reviews: Data Mining & Knowledge Discovery, 8*(4), e1253. doi: 10.1002/widm.1253

Zhang, W. L., & Chen, Y. H. (2020). The construction of the media image of library—Based on the coverage of People's Daily from 2009 to 2018. *Library Tribune, 40*(1), 96-104. (in Chinese)

(Received: 2020/10/27; Accepted: 2021/9/30)

臺灣大學學術期刊資料庫

圖書資訊學刊　第19卷第2期　（2021.12）　頁43-76
doi: 10.6182/jlis.202112_19(2).043

海峽兩岸報紙中圖書館館員媒體形象之探討

Exploring Media Images of Librarians in Newspapers of Mainland China and Taiwan

謝　凡[1]　陳光華[2]

Fan Xie[1], Kuang-Hua Chen[2]

摘　要

公眾形象反應機構形象塑造，面臨科技衝擊的圖書館需要改變民眾對其之印象。圖書館員是圖書館形象塑造中的重要環節。本研究透過文本分析和情緒分析的研究方法，從詞彙的角度勾勒圖書館員在兩地重要報紙中體現出的形象。這些詞彙特徵包含兩地報紙文本之高頻詞、共現詞、顯著詞、情緒詞的特徵。研究發現中國大陸的圖書館員是以歷史類專家的形象出現，臺灣的圖書館員則是以提供優質讀者服務為目標的形象出現。兩地館員工作內容偏重不同，大陸館員在於保護傳統文化，而臺灣館員則在於高效推動館藏資源的使用；兩地都認同圖書館員是館藏資源和讀者之間的重要橋樑。最後，本研究結果顯示圖書館員的報導情緒在兩地都是積極正向的，雖然僅稍高於中性情緒。

關鍵字：圖書館員、媒體形象、情感分析、文本分析

[1,2]國立臺灣大學圖書資訊學系

Department of Library and Information Science, National Taiwan University, Taipei, Taiwan

* 通訊作者Corresponding Author: 陳光華Kuang-Hua Chen, E-mail: khchen@ntu.edu.tw

註：本中文摘要由作者提供。

以APA格式引用本文：Xie, F., & Chen, K.-H. (2021). Exploring media images of librarians in newspapers of Mainland China and Taiwan. *Journal of Library and Information Studies, 19*(2), 43-76. doi: 10.6182/jlis.202112_19(2).043

以Chicago格式引用本文：Fan Xie and Kuang-Hua Chen. “Exploring media images of librarians in newspapers of Mainland China and Taiwan.” *Journal of Library and Information Studies* 19, no. 2 (2021): 43-76. doi: 10.6182/jlis.202112_19(2).043

臺灣大學學術期刊資料庫