

## How Indian libraries tweet? Word frequency and sentiment analysis of library tweets

Swapan Kumar Patra

Tshwane University of Technology, Pretoria, Republic of South Africa,  
Email: skpatra@gmail.com; PatraSK@tut.ac.za

*Received: 12 July 2019, revised: 08 November 2019, accepted: 25 November 2019*

The study attempts to map Twitter activity of selected Indian libraries using word frequency and sentiment analysis. Tweets of 18 libraries' (5 academic libraries, 5 government libraries, 5 school libraries and 3 public libraries) were downloaded during June and July 2019. 'R' software was used for the analysis. The study finds that Indian libraries are less active on Twitter. The word cloud based on the most frequently occurring words from the Tweets observed variations in Tweets depend upon the type of libraries. Sentiment analysis of Tweets showed that most of the libraries' Tweets are positive in nature. The study recommends that Indian libraries should use twitter to promote their collections and services.

**Keywords:** Twitter; Libraries; Library services; Sentiment analysis; Word cloud; R software; Social network sites

### Introduction

There are many Social Networking Sites (SNS) and these sites are pervasive, affecting almost every sphere of life. Library and information systems are not immune to the influence of these social networking sites<sup>1</sup>. Particularly, Twitter and Facebook are engaging library stakeholders in various ways<sup>2</sup>. Hundreds of libraries across the globe are using Twitter to promote their services<sup>3</sup>. Twitter also helps libraries to reach out to their clientele<sup>4</sup>. Many empirical studies have observed that Twitter accounts of academic libraries' act as multifaceted tool for library functions. Among the various services, "News and announcements" posted on Twitter received the highest score. This is followed by "library collections" and "library services". Additionally, Twitter accounts of academic libraries generally post links from their parent institutes. Academic libraries also post other links more often than actual contents. Other results show different patterns of communication and interaction between libraries and their Twitter followers<sup>5</sup>. Moreover, it is observed that library twitter accounts mainly circulate news and other information from their parent institute's accounts<sup>6</sup>.

As for India, there have been hardly any empirical investigation on the use of Twitter by Indian libraries

for propagating their services to their users<sup>7</sup>. This paper is an attempt to investigate, how Twitter is used by different types of Indian libraries for disseminating their services. This study uses keyword clustering by text analysis<sup>8</sup> and sentiment analysis<sup>9</sup> techniques to investigate the nature of various Indian libraries' Tweets<sup>7</sup>. The study explores the following research questions: How Indian libraries tweet on Twitter? Is there any difference in tweets depending on the nature of the library? Do the academic, school, public and government libraries differ in their tweets in terms of topic modelling? Is there any difference in their nature of tweets based on the sentiment analysis?

### Review of literature

Information and communication technology (ICT) has the all-pervasive impact in every sphere of modern life. Libraries are not immune to the effect of ICT in their functioning. The latest web technologies and social networking sites have radically changed the libraries all over the globe. Traditional library services, for example, Current Awareness Services (CAS) and Selective Dissemination of Information (SDI) services have been redefined. Twitter is one of those innovative tools that libraries and librarians are utilizing along with their other existing

communication channels. For content alerts, new arrivals, CAS, SDI and some other services librarians are utilizing Twitter as the new and handy tool for their services.

Although, Twitter is very popular among libraries globally, it is not that really popular among Indian librarians. Twitter in Indian libraries is yet to gain popularity with respect to libraries and librarians worldwide<sup>10</sup>. However, Indian libraries are not far behind in terms of latest technology use. Empirical evidences shows that Indian librarians use the ICT in their libraries quite extensively. For example, a case study of networking services in the university libraries of Karnataka observed that among a sample of 54 universities, 53 are quite active on the Internet and have their own websites. However, very few libraries are connected to Twitter. The study further observed that, the university libraries in the study area are catching up with the social network technologies but many of these services are still in 'germinating phase' and a few are used for very basic purposes<sup>11</sup>.

Students are the major consumers of social media. The use pattern of networking sites among the Indian college students show that there is an increasing trend of social media utilization by the university and college students for their information need. College and university students in particular quite aware of the strength of social media. They understand their information needs and how to find precise information of their need<sup>12</sup>. Hence the library and information professionals have the tremendous opportunity to reach these young minds of colleges and universities through SNS. So, the library and information experts may develop professional relationships among the different user groups through the social media to increase information sharing in a real-time basis. This will perhaps ensure libraries are more committed to their users in information need<sup>13</sup>.

With the introduction of social media, people have found various ways through which they can express their thoughts, opinions and beliefs. Hence, the present-day social media is huge sources of unstructured data. If this data is collected and analysed with proper strategy it could result in useful outcomes. In case of libraries and information centres, information from various social media can be collected, analysed and repackaged depending on various needs<sup>14</sup>. Even the analysis of social media information through text mining and sentiment

analysis perhaps cater to the information needs of today's users. Also, the information obtained from the social media can be a good feedback mechanism of the libraries to improve their services. These analyses could be helpful to design strategy in present day libraries. Moreover, these powerful techniques could open various new avenues for marketing of library services<sup>15</sup>.

It can be concluded from the literature review above that in today's world, opinions and reviews accessible in the public domain could be very crucial in formulating strategies for any institute. For a library, it can be useful to find users' opinion in a particular service, library collection and so on<sup>16</sup>. In this context publicly available Twitter data is extremely informative.

There are various ways of collecting the data and doing analysis like topic modelling, word cloud, sentiment analysis of Tweets and so on<sup>16</sup>. Although, Indian Libraries are not that active on social media, this study is an attempt to investigate some selected Indian libraries' activity on Twitter. The study also identifies the top most occurring keywords and analyses the sentiments of the Tweets and finally recommendation have been made based on the findings.

### Objectives of the study

The exponential growth of unstructured data from different social network and blogging sites can be analyzed to draw insights and take necessary actions based on the institute's strategy<sup>17</sup>. The three primary research objectives of this study are:

- To examine how different types of Indian libraries tweet;
- To find if there is difference between tweets among libraries based on the topic modelling produced by different types of libraries; and
- to find if the content of the tweets by libraries tend towards being positive, negative or bear other types of sentiment.

### Methodology

Library tweets were collected from 18 different Indian libraries during June and July 2019. The libraries with more than 100 Tweets were collected for this study. For a balanced sample size, the study

has selected different types of libraries. Among the 18 libraries, 5 are academic libraries, 5 government libraries, 5 school libraries and 3 public libraries. The selected libraries are from different parts of India. Hence, the sample size is perhaps a fair representation of libraries across the country. The number of selected libraries and their tweets are shown in the Table 1. The word frequency and sentiment analysis of the libraries are done based on the total tweets. The tweets are further clubbed and categorised based on the nature of the libraries. For example, separate analysis was done on academic libraries, government libraries, public libraries and school libraries.

The data has been collected from the Twitter using R software. R is a statistical programming language consists of thousands of integrated packages and their various functions<sup>18</sup>. These packages allow to handle numerous data, various statistical analysis of data and have visualization functions. Also, a strong R user and developer community is continually working to enhance and improve R base codes. Many online blogs are available to help on the R packages that are freely available and accessible to interested users<sup>19</sup>.

The tweets from the individual libraries' Twitter handle were collected using the Twitter API. The collected tweets were stored as separate files in Excel sheet. The data collection and processing steps are shown in Fig. 1. The tweets were then grouped based

on the nature of the library— academic, government, public and school libraries. A separate file was prepared merging all tweets from all libraries.

The study has the following limitations. Firstly, the Tweets were collected at a particular time of the year hence the number of tweets may vary. Secondly, this study is based on only the English language tweets. Many of the Indian libraries tweet in local languages. For example, Delhi Public Library tweets in Hindi, Anna Centenary Library tweets in Tamil and so on. The tweets other than English language have not been included. Thirdly, a few libraries particularly the school libraries did many retweets. These tweets were not captured in the downloaded data. Finally, the study is based on the twitter handle of the respective institute's library twitter handle. This study did not made attempt to find who handles the tweets of the library.

In all out of the total 8,854 tweets and 5,540 tweets were used for the analysis.

## Results

The five academic libraries selected for this study include Indian Institute of Management Library, Ahmedabad (Vikram Sarabhai Library); Indian Institute of Technology Bombay (IITB-Central Library, Mumbai); Library of Mar Athanasios College

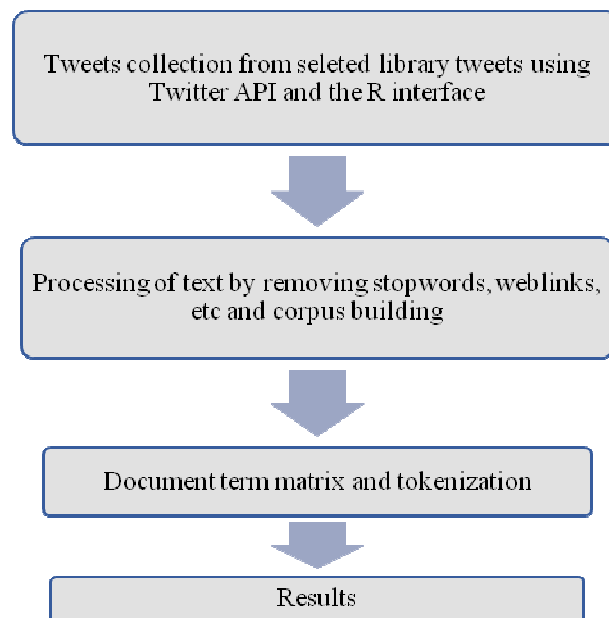


Fig. 1—The methodology flowchart

Table 1—The selected sample of library Twitter used for this study

Sl. no.	Library	Twitter handle	Description	Place	Joined Twitter in	Total tweets	Library type	Tweets collected
1.	Library KV Pattom	@librarykvpatom	Automated School Library of Kendriya Vidyalaya	Trivandrum	November 2007	932	School Library	932
2.	Delhi Public Library	@delhipubliclibrary	Public Library services in Delhi and one of the busiest Public Library in South East Asia	New Delhi	April 2010	831	Public Library	214
3.	National Digital Library of India	@NDLIndia	National digital literacy NDL is a mission project by Ministry of Human Resources and Development along with Indian Institute of Technology Kharagpur	Kharagpur	June 2016	436	Government Library	275
4.	Nehru Memorial Museum & Library	@_NMML_	An autonomous institution under the Ministry of Culture, Government of India	New Delhi	May 2015	626	Government Library	509
5.	Vikram Sarabhai Library	@VSLIIMA	Library of Indian Institute of Management Ahmedabad	Ahmedabad	December 2016	564	Academic Library	233
6.	Library KV Azamgarh	@librarykvazam	School library of Kendriya Vidyalaya, Azamgarh	Azamgarh	November 2015	673	School Library	177
7.	Library KV 2 Calicut	@librarykvtwocla	School library of Kendriya Vidyalaya No.2, Calicut	Calicut	March 2012	182	School Library	182
8.	Library KV Haldwani	@librarykvhld	School library of Kendriya Vidyalaya, Haldwani	Haldwani	February 2010	521	School Library	346
9.	Library KV Kanjikode	@kvklibrary	School library of Kendriya Vidyalaya, Kanjikode	Palakkad	September 2011	1019	School Library	989
10.	Anna Centenary Library	@ACLChennai	State library of Tamil Nadu.	Chennai	February 2017	998	Government library	78

Contd—

Table 1—The selected sample of library Twitter used for this study

Sl. no.	Library	Twitter handle	Description	Place	Joined Twitter in	Total tweets	—Contd	
							Library type	Tweets collected
11.	Rampur Raza Library	@raza_library	Public library	Rampur	January 2015	587	Public Library	581
12.	Indian Institute of Technology Bombay-Central Library	@centrallibraryb	Central Library of Indian Institute of Technology, Bombay	Mumbai	March 2010	474	Academic Library	475
13.	MACFAST Library	@Macfast_Library	College library of Mar Athanasios College for Advanced Studies	Tiruvalla	March 2011	107	Academic Library	92
14.	M D College Library	@MDCLibrary	Library of Maharshi Dayanand College is Arts Science & Commerce college.	Mumbai	September 2011	165	Academic Library	151
15.	Central Reference Library, Kolkata	@KolkataCentral	National Bibliographic Agency under the Ministry of Culture, Government of India	Kolkata	October 2018	67	Government Library	24
16.	Raja Rammohun Roy Library Foundation	@RrrifKolkata	Autonomous body under the Ministry of Culture, Government of India.	Kolkata	October 2015	392	Government Library	11
17.	Amrita Mysore Library	@ASASLibrary	Library of Vishwavidyapeetham Amrita University, Mysore campus	Mysore	June 2009	158	Academic Library	151
18.	Nehru Centre Library	@NehruCenLib	Nehru Centre Library act as a reference library for researchers, students and the general public.	Mumbai	January 2016	122	Public Library	120
						Total available Tweets	Total Tweets collected	5,540

Source: Data downloaded from Twitter API using R software in the month of June / July 2019

for Advanced Studies, Tiruvalla (MACFAST Library); Library of Maharshi Dayanand College of Arts, Science & Commerce college (M D College Library, Mumbai) and Amrita Vishwa Vidyapeetham University Library, Mysore Campus (Amrita Mysore Library). These libraries together tweeted 1,468 times of which 1,102 were used for the study.

Figure 2 shows the frequently occurring words from the cumulative academic library tweets. The top 10 keywords were collected and plotted. Library (148 times) is the most frequently occurring keyword followed by IIT (106 times), IIM Ahmedabad (88 times). It is observed that, Indian academic libraries mostly Tweets the activities of their parent institute's activities. It is also substantiated by earlier studies that academic libraries generally posted information of their parent institute's vision, mission and so on<sup>2</sup>.

The five government libraries include the National Digital Library of India (NDL India, Kharagpur); Nehru Memorial Museum & Library (NMML, Delhi); Anna Centenary Library (ACL, Chennai); Central Reference Library, Kolkata; and Raja Rammohun Roy Library Foundation (RRLF, Kolkata).

These libraries altogether have 2,519 Tweets and 897 were used for this study. Among them Anna Centenary Library has the maximum number of tweets (998). However, only about 78 Tweets are collected in this study. Rest other Tweets are in Tamil languages. The other libraries have many retweets. So, the smaller number of Tweets were harvested for these types of libraries in the database.

Figure 3 shows the frequently occurring keyword of academic library tweets. NMML has maximum number of tweets so it is apparent that NMML is the most frequently occurring keywords (84 times) in the academic library tweets. Also, there are variations among the keywords, for example India (72 times), born (42 times) and so on. Unlike the academic libraries, the major difference here is that the libraries' tweets consist of the institute's identity. For example, NMML, NDL India and so on are the frequently occurring keywords. Moreover, NMML has 'Library' in its organizational name, but most of their tweets are not related to library activities.

Although there are many famous public libraries in India, they either do not have Twitter handles or they

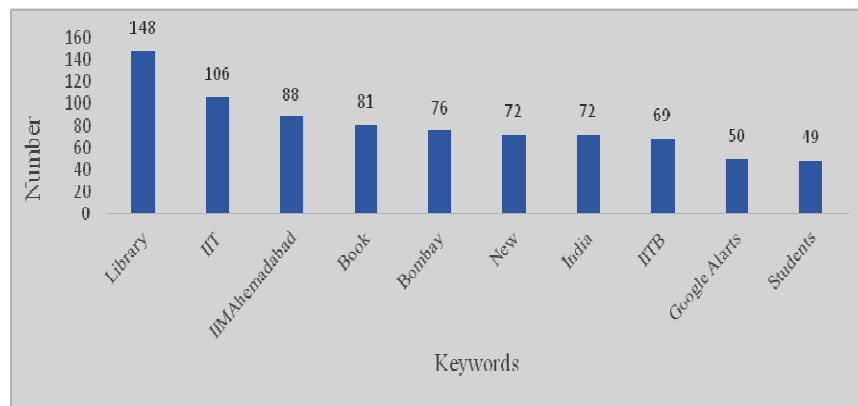


Fig. 2—Most frequent words of academic library tweets

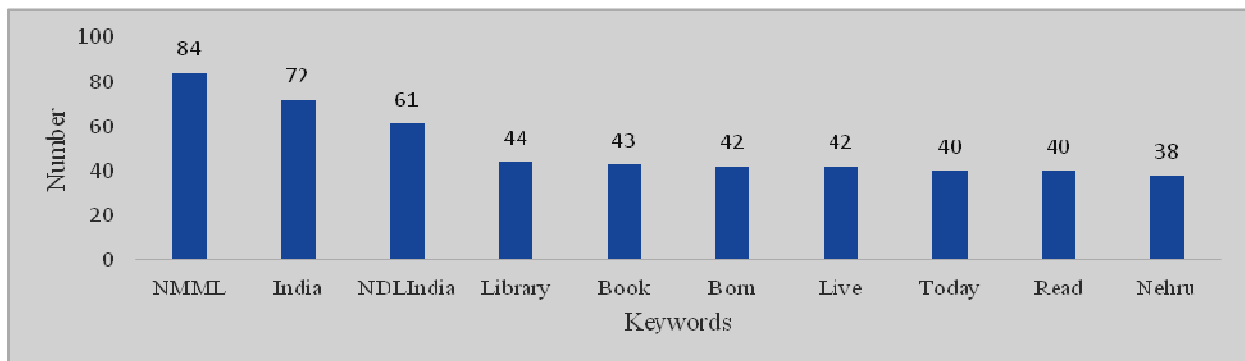


Fig. 3—Most frequent words of government library Tweets

sparsely tweet from their twitter handles. There were are only three public libraries with about 100 tweets. These libraries are Delhi Public Library, Delhi; Rampur Raza Library, Rampur, Uttar Pradesh; and Nehru Centre Library, Mumbai. There are altogether 1,540 Tweets from these libraries. Among them, 915 tweets are collected through the R interface for this analysis. The most frequently occurring keywords are shown in Figure 4. In these tweets, library is the most frequent keywords (182 times) followed by new (154 times).

There are five school libraries in the selected sample. These libraries are Kendriya Vidyalaya Library Pattom, Trivandrum, Kerala; Kendriya Vidyalaya Library, Azamgarh, Uttar Pradesh; Kendriya Vidyalaya Library, Calicut, Kerala; Kendriya Vidyalaya Library, Haldwani, Nainital, Uttarakhand; and Kendriya Vidyalaya Library Kanjikode, Kerala.

It is observed from the Table 1 that school libraries had 3,327 tweets and among them 2,626 Tweets were collected using R interface. Figure 5 shows that the

most frequently occurring keywords from the school library tweets are Books (535 times) followed by library (347 times).

**Sentiment analysis**

Sentiment analysis is a useful tool to automatically measure the various types of emotions in the text<sup>9</sup>. The library tweets were categorized into 10 different types of sentiments. These sentiments are: positive, trust, anticipation, joy, negative, surprise, fear, anger, sadness and disgust. The sentiment analysis of words from the tweets show that most of the library tweets fall into positive sentiments. From the public libraries, among the 1,401 words 624 are positive words (about 45 percent). Maximum number of negative sentiments are observed form the school libraries tweets 649 words (10 percent). Figure 6 shows the sentiment analysis of library tweets.

**Conclusion**

Twitter is one of the most powerful social network today. This networking site is very popular among

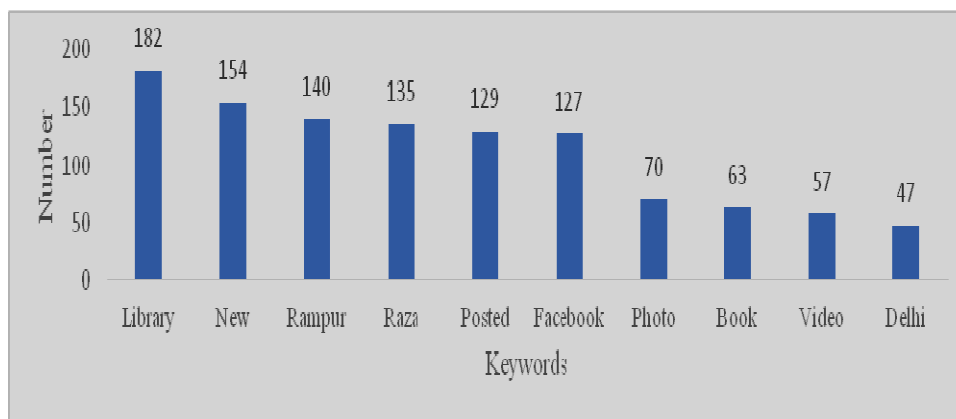


Fig.—4: Most frequent words of public library Tweets

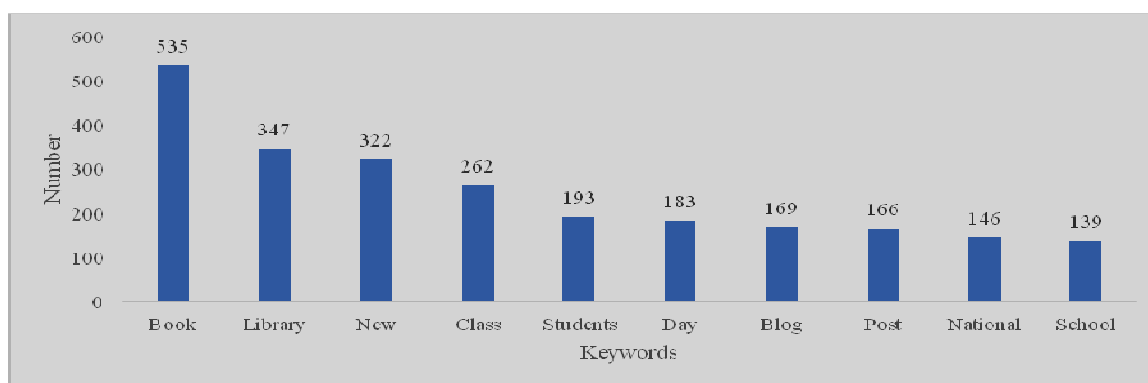


Fig. 5—Most frequent words of school library tweets

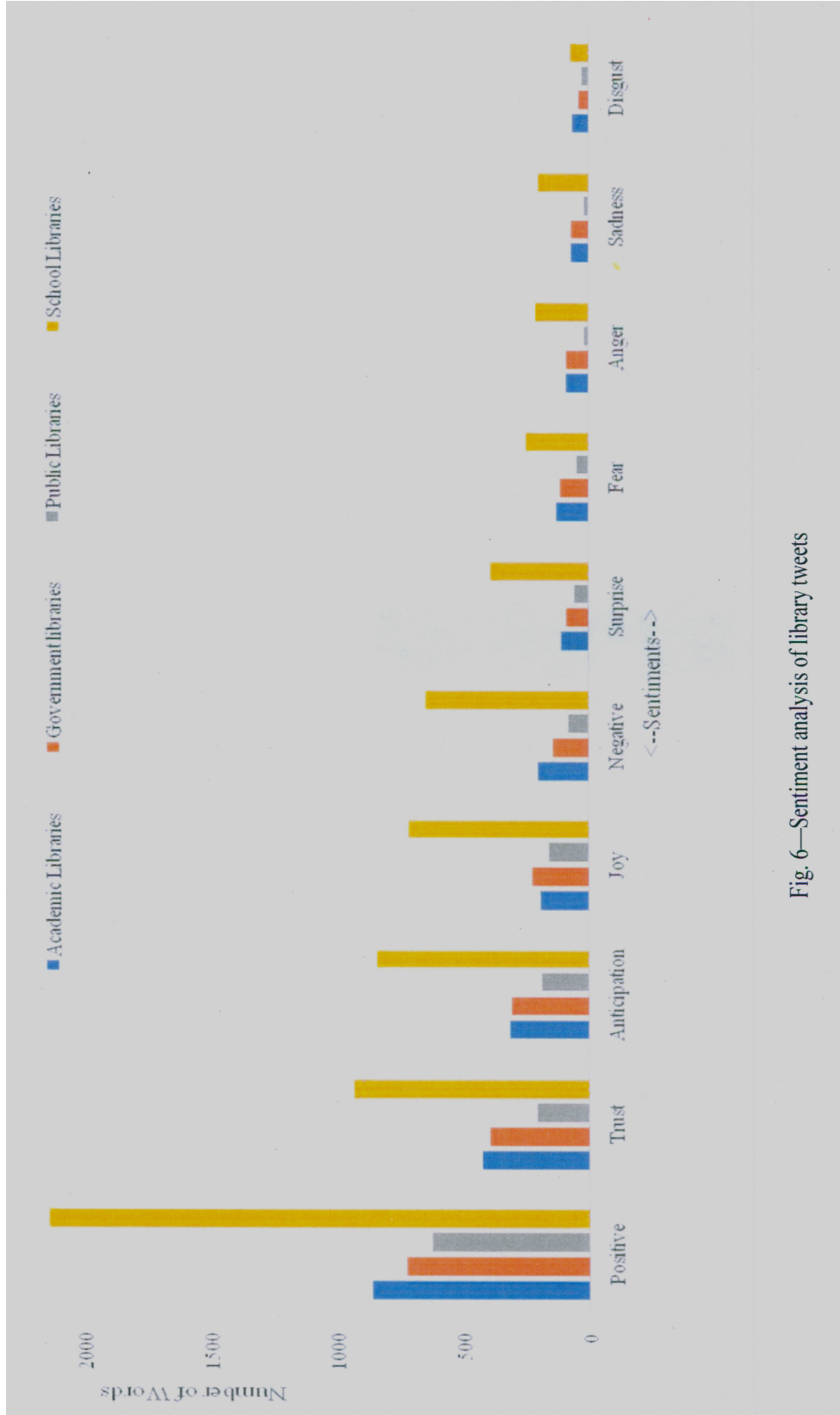


Fig. 6—Sentiment analysis of library tweets



various user groups. People use this microblogging site for various purpose. This powerful social media enables libraries also to share information about the libraries and their activities. Globally, libraries from developed countries are quite active in Twitter. At the time of this study, it was found that many famous libraries across the globe used Twitter extensively. For example, in the months of June and July 2019, Library of Congress, US (@librarycongress) had 17.1 K tweets, National Library of Medicine, US (@nlm\_news) had 11.3 K tweets, The British Library (@britishlibrary) had 13 K Tweets and National Library of Scotland (@natlibscot) had about 28.2 K Tweets.

Indian libraries are not very active on Twitter. It is time that libraries and information centres (academic, public, school libraries and so on) exploit the potential of Twitter. Library activities in other social media like Facebook and other social sites may give a different picture.

### Acknowledgement

The South African Research Chairs Initiative (SARChI) - Innovation Studies, and Tshwane University of Technology, Pretoria, South Africa.

### References

- Aharony N, Twitter use in libraries: an exploratory analysis, *Journal of Web Librarianship*, 4(4) (2010) 333-350.
- Palmer S, Characterizing university library use of social media: a case study of Twitter and Facebook from Australia, *Journal of Academic Librarianship*, 40(6) (2014), 611-619.
- Vassilakaki E, and Garoufallou E, The impact of Twitter on libraries: a critical review of the literature. *Electronic Library*, 33(4) (2015) 795-809.
- Sewell R R, Who is following us? Data mining a library's Twitter followers, *Library Hi Tech*, 31(1) (2013) 160-170.
- Al-Daihani S M and AlAwadhi S A, Exploring academic libraries' use of Twitter: a content analysis, *Electronic Library*, 33(6) (2015) 1002-1015.
- Shulman J, Yep J and Tome D, Leveraging the power of a Twitter network for library promotion, *Journal of Academic Librarianship*, 41(2) (2015) 178-185.
- Zou H B, Chen H M and Dey S, Understanding library user engagement strategies through large-scale twitter analysis. In *2015 IEEE First International Conference on Big Data Computing Service and Applications (Big data service 2015)*, 361-370.
- Welbersa K, Atteveldt W V and Benoit K, Text Analysis in R, *Communication Methods and Measures*, 11(4) (2017) 245-265.
- Thelwall M, Buckley K and Paltoglou G, Sentiment in Twitter events, *Journal of the American Society for Information Science and Technology*, 62(2) (2011) 406-418.
- Munigal A, Use of microblogs in India: A study of Twitter usage by librarians and in libraries, *Journal of Library Administration*, 54(7) (2014) 590-608.
- Deepthi, Tadasad P G and Patil S, Use of online social networking services in university libraries: A study of university libraries of Karnataka, India, *DESIDOC Journal of Library and Information Technology*, 37(4) (2017)249-258.
- Balamurugan T and Thanuskodi S, Use of social networking sites among the college students in Tamil Nadu, India, *Library Philosophy and Practice*, (2301) (2019) Available at: <https://digitalcommons.unl.edu/libphilprac/2301/>.
- Chakravarty R and Kiran A, Social networking in libraries: A case study of post graduate students and research scholars of Panjab university library, India, In *S. Thanuskodi (Ed.), Challenges of Academic Library Management in Developing Countries (Hersey, PA, USA: IGI Global), 2013, p. 112-130*.
- Mishra P, Rajnish R and Kumar P, Sentiment analysis of Twitter data: case study on digital India. In *2016 International Conference on Information Technology (InCITe) - Next Generation It Summit on the Theme - Internet of Things: Connect Your Worlds.IEEE*, Noida, India 6-7 October 2016, p 148-153.
- Lamba M and Madhusudhan M, Application of sentiment analysis in libraries to provide temporal information service: a case study on various facets of productivity, *Social Network Analysis and Mining*, 8 (1) (2018) 63.
- Mamgain N, Mehta E, Mittal A and Bhatt G, Sentiment analysis of top colleges in india using Twitter data. In *2016 International Conference on Computational Techniques in Information and Communication Technologies, IEEE*, New Delhi, India (11-13 March 2016).
- Barnwal A K, Choudhary G K, Swarnim R, Kedia A, Goswami S and Das A K, Application of Twitter in Health Care Sector for India. In *2016 3rd International Conference on Recent Advances in Information Technology (RAIT)*, IEEE, Dhanbad, India, 3-5 March, 2016.p. 172-176.
- Luke D A, A User's Guide to Network Analysis in R. (Switzerland: Springer International Publishing), 2015.
- Kwartler T, Text Mining in Practice with R. (John Wiley & Sons Ltd), 2017.