

Social Tags Versus LCSH Descriptors: A Comparative Metadata Analysis in the Field of Economics

Kalyan Sundar Samanta[#] and Durga Sankar Rath^{S*}

[#]*Prabhu Jagatbandhu College, Howrah-711 302, India*

^S*Vidyasagar University, Midnapore- 721 102, India*

^{*}*E-mail: dsrath@mail.vidyasagar.ac.in*

ABSTRACT

The concept of 'social tagging' has gained popularity nowadays due to the emergence of web 2.0 technologies. Those technologies led to the practice of associating metadata with digital resources among users through collaboratively or socially for self-information retrieval. Many researchers have opined that social tags can enhance the use of library collections. The present study was predominantly carried out to compare social tags collected from the LibraryThing website with Library of Congress Subject Heading (LCSH) descriptors collected from the Library of Congress Online Catalogue applied for thousand book titles in the field of Economics. The study also aimed to know whether social tags can be applied in the library database or not. The findings elucidate that users mostly use descriptors (47.39 %) as tags than expert's usage of tags (12.77 %) as descriptors. Spearman's correlation suggests that 75 per cent chance where tags and descriptors can be used simultaneously in overlapping terms. The Jaccard similarity coefficient identifies that users and experts use different terminologies to annotate the books. Users and experts use at least one common keyword for major book titles (908). Users mostly sought title based keywords but experts use mostly subject-based terminologies. The study further clarifies that social tags may be incorporated into the library databases but cannot replace LCSHs. The accessibility and usage of documents especially in the field of economics may be enhanced once the notion of social tags is incorporated with the library OPAC.

Keywords: Social tags; Library of Congress Subject Headings (LCSHs); Social tagging; Folksonomy; Collaborative tagging; LibraryThing.

1. INTRODUCTION

As the world is relying more and more on search engines, metadata becomes the central element for organising, identifying and retrieving the relevant information from the heap of information¹. The success of modern search engines completely depends on effective metadata implementation. Metadata is the data about data that gives a complete description of digital objects². Traditionally there was centralisation in the metadata creation, only experts or professionals using a controlled vocabulary at libraries, archives and museums (LAMs) can generate it for the content they manage such as Library of Congress Subject Headings (LCSHs)³. Expert's created metadata is highly technical, specialised and inflexible in nature which predominantly illustrates the content or subject headings of resources. The underlying cost for such metadata generation is huge and more times required for the metadata creation⁴. The rapid growth of digital resources on the web is gradually exposing the limitations of the expert's assigned metadata in the present scenario. The limitations of the expert's assigned metadata generation need further introspection, in view of user-driven metadata approach⁵. The concept of such

metadata is derived from 'social tagging' or 'collaborative tagging'. That has been emerged due to the effectiveness of web 2.0 applications that enable users to collaboratively annotate web resources using keywords, also known as tags⁶. The resources on the web may be different such as for videos (YouTube), music (last.fm), photographs (Flickr), bookmarked web sites (Del.icio.us, CiteULike), and for catalogue records (LibraryThing)⁷. The process of social tagging or collaborative tagging provides a social classification of resources which is known as 'folksonomy', a personal free tagging of information and objects for one's own retrieval⁸ where users and resources are related through tag assignments⁶. Being originated from the uncontrolled vocabulary social tagging allows people to enter non-hierarchical free-form keywords based on user choice and needs⁹, like the same found in LibraryThing, an online service that helps people to catalogue books by tags. LibraryThing has more than thirteen crores of books catalogued by more than twenty-three millions of people with fourteen crores of tags¹⁰ which allows people for own information retrieval and to find the people of similar interest in the digital environment¹¹. Having many advantages, it suffers from disadvantages too. It has semantic ambiguity, synonymous issues, lack of controlled vocabulary¹²⁻¹³ and use of many personal tags ('to-read', 'read', 'read in 2007') for personal use rather than public benefit¹⁴.

Though there is a quality issue of tags, still worldwide social taggers and social tags are increasing¹⁵ years after years. This indicates that the world is tended to rely on social media and social tagging. Many researchers and information professionals realised that social tags can be incorporated into the cataloguing of records. Though tags could not replace the role of controlled vocabulary like LCSH but can be used to enhance the subject accessibility of library collections¹⁶.

2. LITERATURE REVIEW

Lu⁴, *et al.* compared social tags assigned by users from LibraryThing website and subject terms assigned by experts according to the LCSHs in order to identify the difference and connections between social tags and expert assigned subject terms and further the feasibility and obstacles of implementing in library systems. The study showed that it is possible to use social tags to improve the accessibility of library collections.

Heymann and Garcia-Molina¹⁷ also conducted the same study to know whether a controlled vocabulary of library keywords called the Library of Congress Subject Headings (LCSH) is different from the vocabulary developed by the users of LibraryThing. The results indicated that 50 per cent of LCSH terms are overlapped with the LibraryThing database.

Yi and Chan¹⁸ investigated the linking of folksonomy and LCSH on the basis of word matching where two-thirds of all tags matched with LC subject headings with additional 10 per cent remaining tags having potential matches.

Rolla¹⁷ also compared LibraryThing tags with LCSH subject headings for a group of books. The study indicated that user tags can enhance subject access to library materials but they cannot entirely replace the controlled vocabulary like LCSHs.

Thomas¹⁹ conducted a quantitative analysis to know whether folksonomies can complement to LCSHs and to which extent folksonomies can replicate the LCSHs. The study showed that social tagging had enhanced access to resources.

Lawson²⁰ compares LCSHs with user tags from Amazon.com and LibraryThing for each title which evaluates the user tags can enhance subject access.

3. RESEARCH QUESTIONS

The present study is aimed to formulate some research questions. These are as follows.

- Do both social tags and LCSH descriptors follow the same vocabulary?
- Do social tags and LCSH descriptors are similar on the basis of usage?
- Do social tags can enhance the subject access like LCSH descriptors?
- Do social tags and LCSH descriptors use keywords from the title of books?
- Do social tags and LCSH descriptors complement each other?

4. METHODOLOGY WITH DATASET

The study was predominantly carried out to evaluate the relationship between social tags and LCSH descriptors where both were applied for each book title in the field of

economics. The present study has used LibraryThing, a cataloguing and social networking site for book lovers (ref. 11) for collecting social tags. Though there are few such databases like Goodreads, bookcrossing and Litsy etc. but LibraryThing database exclusively provides a collection of social tags of each book through a weighted list technically called tag cloud. The tag cloud graphically represents the most active tags currently found in the most popular content of a site. Besides, the database also has a wide range of collection of books from Amazon, the Library of Congress and over a thousand other libraries. Thousand book titles in the English language were extracted randomly from the LibraryThing database (www.librarything.com) subject to the availability of those titles in the Library of Congress Online Catalogue (<https://catalog.loc.gov>). The study also ensures that only those titles were extracted which were catalogued by at least ten members (≥ 10) and had at least three social tags (≥ 3) in the LibraryThing database and simultaneously had at least one subject heading (≥ 1) under MARC field 650 (Topical term) in the Library of Congress database.

The crawling process was conducted in the month of September to December 2018. Each tag has a different frequency in the tag database which defines the number of users used it for annotating the particular title. The more frequency means the tag is more popular or more important for the title. The present study takes into account those tags that have a frequency at least twice or more than twice (≥ 2) in the tag database. Otherside, in the Library of Congress database each bibliographic record is described in MARC format which comprises many MARC fields for representing different pieces of information. The study selects the contents of field 020 for ISBN and field 245 for Titles and field 650 for Subject Added Entry (topical terms). Library of Congress MARC records contain many fields that come under Subject access field (6XX) for subject information but the study selects only the contents of field 650 for Topical term. Field 650 comprises two indicators and many subfields. The present study clarifies that the first indicator (Level of subject) must be from 0 (zero) to 2 (two) range and in the case of second indicator (Thesaurus) must be 0 (zero) i.e., Library of Congress. In the case of subfields only \$a, \$x, \$y, \$z and \$v have been taken into the account. Each subfield holds different subject descriptors which altogether form the subject headings. The descriptors that occur in those subfields only are taken into the account separately. Some titles contain the same descriptors under different subfields in 650 data field but the study only chooses the unique descriptors for each title. The study confines only unique social tags are compared with the unique LCSH descriptors for each book title.

Moreover, the total 20699 social tags and 4144 LCSH descriptors were collected from both databases initially. The datasets were changed into the lower case for processing and identification of duplicates as well. MS Excel was used for the whole data processing and especially the Pivot table was used for counting unique terms. Table 1 indicates that after removing duplicates unique social tags and unique LCSH descriptors are 2983 (average 2.98 per book) and 804 (average 0.80 per book) respectively. The overall study was conducted based on those unique social tags (2983) and unique LCSH descriptors (804).

Table 1. Total terms crawled and unique terms

| | Total terms crawled | Unique terms |
|-------------|---------------------|--------------|
| Social tags | 20699 | 2983 |
| LCSH terms | 4144 | 804 |

(1) Bush (1) **business** (42) Calibre import (1) **capitalism** (28) Case 2 Shell 4 (1) cataloged (1) Central America (1) change (1) Chas. T. Main Inc. (1) **CIA** (1) class war (1) coercion (1) cold hard cash (1) college reading list (1) **colonialism** (3) Columbia (1) Communist Propaganda (1) completed (1) confession (1) Confession of an Economic Hitman (1) Confessions (1) Confessions of an Economic Hit Man (1) confessions of an economic hitman (1) **conspiracy** (19) **conspiracy theory** (2) Conspiracy/Corporate/U.S. Government (1) consultants (1) consulting (1) contemporary (1) Contemporary Commentary (1) corp. evil (1) corporate (1) corporate america (1) corporate faithfulness (1) Corporate globalism (1) corporate greed (1) corporate skullduggery (1) corporate-state collusion (1) corporation (1) **corporations** (19) Corporations - Corrupt practices (1) Corporations Economists United States National Security Agency World Bank Imperialism (1) Corporations--American--Foreign countries (1) Corporations--United States--Corrupt practices (1) **corporatocracy** (6) Corrupções norteamericanas - corrupção (1) **corruption** (24) cover (1) covert (1) crestwood (1) **crime** (2) criminal (1) critique (1) **culture** (9) **current affairs** (2) current affairs. US (1) Current Affairs/Politics/Globalisation (1) **current events** (18) Currently-reading (2) D.88 POLITICA (1) D.POLITICA (1) daemon bibliography (1) Danny (1) dark (1) DCOffline (1) DD (1) **debt** (3) **decalt** (2) **deception** (2) deep politics (1) deep state (1) default (1) **democracy** (3) dev (1) **developing countries** (2) **development** (22) Devassa (1) disaster-capitalism (1) **discard** (2) **dism** (1) Division: Non-fiction (1) dollar hegemony (1) done (1) dpp (1) DR (1) Dutch (1) East Greenbush PL (1) ebk (1) **ebook** (4) Ec leader (1) eco (1) econ business (1) econ_dev_history_postIR (1) Economia Internacional (1) **economic** (3) **economic development** (2) economic development theory (1) **economic exploitation** (2) economic history (1) economic imperialism (1) economic issues (1) **economic policy** (2) economic systems (1) economic theory (1) economic warfare (1) **economics** (28)

Figure 1. Sample social tags in LibraryThing database.

```
650 _0 |a Economists |z United States |v Biography.
650 _0 |a Energy consultants |z United States |v Biography.
610 20 |a Chas. T. Main, Inc.
610 20 |a World Bank |z Developing countries.
650 _0 |a Corporations, American |x Corrupt practices.
650 _0 |a Imperialism |x History |y 20th century.
650 _0 |a Imperialism |x History |y 21st century.
```

Figure 2. Sample subject descriptors in Library of Congress MARC database.

5. DATA ANALYSIS AND RESULTS

5.1 Terminological Overlapping

The overall unique social tags and unique LCSH descriptors were compared to identify the common terms which were used by both users and experts for the entire collection. Table 2 defines that 2983 unique social tags and 804 unique LCSH descriptors were compared and of which 381 terms found overlapped i.e., only 381 terms which are applied by both users and experts for the entire collection. Overlapping terms comprise a very small portion of social tags (12.77 %) which means a major portion of tags (87.23 %) can't be found in LCSH descriptors. Likewise, overlapping terms comprise a near to half of the LCSH descriptors (47.39 %), that means there is near about 50 per cent probability that it can be adopted by users as social tags⁴.

Table 2. Unique terms and overlapping terms

| | Unique terms | Overlapping terms | Per cent |
|------------------|--------------|-------------------|----------|
| Social tags | 2983 | 381 | 12.77 |
| LCSH descriptors | 804 | 381 | 47.39 |

5.2 Rank Correlation of Overlapping Terms

Further, the study wants to know the usage frequency of overlapping terms when used as tags and LCSH descriptors. The overlapping terms were ranked frequency wise (highest to lowest) in both datasets. Spearman's rank correlation was used in the following equation to asses.

Spearman's correlation coefficient:

$$r = 1 - \frac{6 \sum d^2}{n(n^2 - 1)} \quad (1)$$

Spearman's correlation coefficient of the two rankings becomes 0.75 which identifies there is a strong relation between them. It also determines that when an overlapping term is used as LCSHs by experts, there are 75 per cent chances to be used as social tags also by the users.

5.3 Similarity and Distance Measurement Based on Usage

Top frequent social tags & top frequent LCSH descriptors were analysed in order to identify whether there is any similarity and distance exist on the usage level. Top frequent social tags & LCSH descriptors were ranked in both datasets according to their corresponding frequencies and grouped into five different top frequent levels like 100, 200, 300, 400 & 500 to show the similarity and distance in different levels. The following equation was used for the Jaccard similarity index.

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|} \quad (2)$$

[Where A = social tags and B = LCSH descriptors]

In the equation, n is the number of top frequent terms. A is the set of n frequent tags and B is the set of top n frequent LCSH descriptors. Fig. 3 shows the Jaccard index when n varies from 100 to 500 (ref. 19, p. 3). Jaccard similarity index becomes 0.13, 0.18, 0.15, 0.15 and 0.14 respectively for five different levels which indicates a low overlap between tags and descriptors. The study also measures Jaccard distance between them which results in 0.87, 0.82, 0.85, 0.85 and 0.86. That means top frequent social tags used by users and top frequent LCSH descriptors used by experts are different.

The following equation was also used for the Jaccard distance.

$$1 - J(A, B) = \frac{|A \cup B| - |A \cap B|}{|A \cup B|} \quad (3)$$

[Where A = social tags and B = LCSH descriptors]

5.4 Top Twenty Frequent Social Tags & LCSH Descriptors

The study compared the top twenty frequent social tags with LCSH descriptors in both vocabularies. The study tried to measure subject-based terms (particular to this subject)

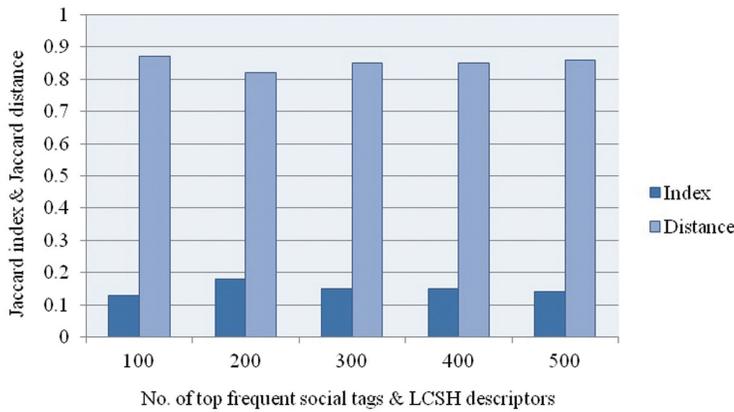


Figure 3. Jaccard similarity and distance.

Table 3. Top twenty frequent social tags and LCSH descriptors with special reference to economics and its allied subject

| Social tags | Frequency | LCSH descriptors | Frequency |
|-------------------------|-----------|----------------------------------|-----------|
| economics | 899 | united states | 296 |
| non-fiction | 838 | history | 239 |
| to-read | 564 | economics | 173 |
| history | 557 | economic conditions | 138 |
| politics | 495 | capitalism | 135 |
| business | 393 | economic policy | 114 |
| finance | 355 | economic history | 86 |
| capitalism | 343 | financial crises | 79 |
| economy | 311 | economic aspects | 62 |
| economic history | 280 | economic development | 48 |
| political economy | 273 | investments | 47 |
| usa | 230 | 21st century | 43 |
| read | 219 | biography | 40 |
| unread | 216 | globalisation | 40 |
| sociology | 201 | great britain | 40 |
| wishlist | 201 | 20th century | 39 |
| globalisation | 169 | finance | 37 |
| philosophy | 169 | international economic relations | 35 |
| political science | 167 | europe | 30 |
| 20th century | 163 | money | 30 |

and non-subject terms (not particular to this subject but allied subject) out of top frequent terms in both datasets. Table 3 shows that social tag vocabulary contains eight subject-based

terms (08), eight non-subject terms (08) and four personal tags (04) e.g., ‘to-read’, ‘read’, ‘unread’, and ‘wishlist’ whereas LCSH vocabulary contains thirteen subject-based terms (13), seven non-subject terms (07). Table 3 also shows that out of both datasets only seven terms (07) are common. It means the seven terms e.g., ‘economics’, ‘history’, ‘finance’, ‘capitalism’, ‘economic history’, ‘globalisation’, ‘20th century’ are used by both users and experts.

Bold terms in Table 3 clarifies the five (05) subject-based common terms which are used by both users and experts in both datasets. The frequency analysis clarifies that the term ‘economics’ has the highest frequency (899) in social tag vocabulary whereas the highest third (173) in LCSH vocabulary. It means the term ‘economics’ is used in 899 titles out of 1000 titles by social taggers whereas used in 173 titles by experts. Likewise, the term ‘finance’ is used in 355 titles in tag vocabulary and 37 titles in LCSH vocabulary. The term ‘capitalism’ is used in 343 titles in tag vocabulary and 135 titles in LCSH vocabulary. The term ‘economic history’ is used in 280 titles in tag vocabulary and also used in 86 titles in LCSH vocabulary. The term ‘globalisation’ is used in 169 titles in tag vocabulary whereas 40 titles in LCSH vocabulary. That means users mostly use common subject-based terminologies rather than experts.

5.5 Title Wise Comparison of Social Tags with LCSH Descriptors

In this portion, social tags were compared with the LCSH descriptors for each book. The study clarifies that out of 1000 books there are 908 (90.8 %) books that have at least one matching between social tags and LCSH descriptors. That means users and experts adopt at least one common term for annotating the books and the rest 92 (9.2 %) books where users and experts use different terminologies. Figure 4 indicates the magnitude of matching per books. The study deeply describes the matching range from 0 to 100 per cent where it is found that 100 per cent matching for major books i.e., 247(24.7 %). That means all the descriptors were used by social taggers as tags for those 247 books. The other matching percentages are 90 per cent for 1(0.1 %) book, 80 per cent for 39(3.9 %) books, 70 per cent for 50(5 %) books, 60 per cent for 117(11.7 %) books, 50 per cent for 128(12.8 %) books, 40 per cent for 51(5.1 %) books, 30 per cent for 103(10.3 %) books, 20 per cent for 118(11.8 %) books, 10 per cent for 54(5.4 %) books and 0 per cent matching for 92(9.2 %) books.

5.6 Social Tags Compared with LCSH Subdivisions

The study compares social tags with each subfield under MARC field 650 (Subject added entry-topical term) for each book to know the most used subfields by experts for annotating books and to know the subfields where at least one social tags appeared. The subfields taken into the consideration under Field 650 are \$a – topical or geographic name entry element; \$x- topical subdivision; \$y – chronological subdivisions; \$z – geographic subdivision; \$v – form division.

Table 4 explores that the MARC subfield \$a (1000) used for all the records and others are as follows \$x (543), \$z (356), \$y (226) and \$v (62) by the experts. The comparison also

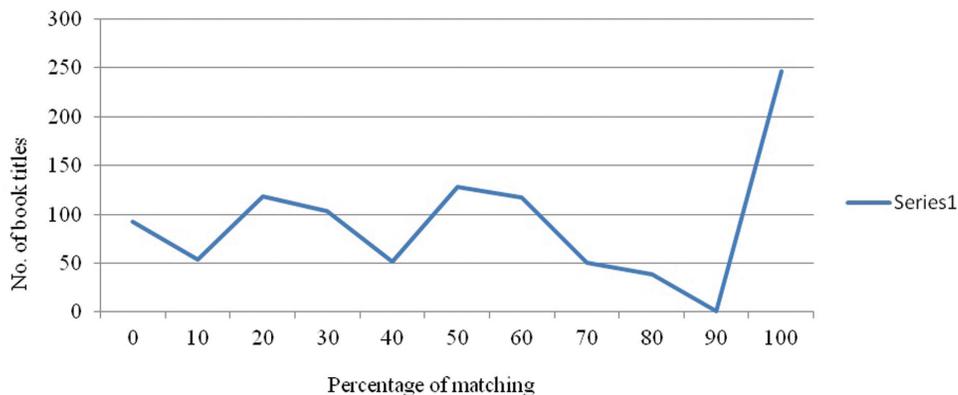


Figure 4. Individual titles wise matching of social tags with LCSH descriptors.

Table 4. Number of book titles which have atleast one tag appearing in LCSH subfields

| LCSH subfields | \$a | \$x | \$y | \$z | \$v |
|---|----------------|----------------|----------------|----------------|--------------|
| No. of records with this field | 1000 (100%) | 543 (54.3%) | 226 (22.6%) | 356 (35.6%) | 62 (6.2%) |
| No. of records which have atleast one tag appearing in this field | 839 | 291 | 63 | 211 | 36 |
| Per cent | 83.9 | 53.59 | 27.88 | 59.27 | 58.06 |

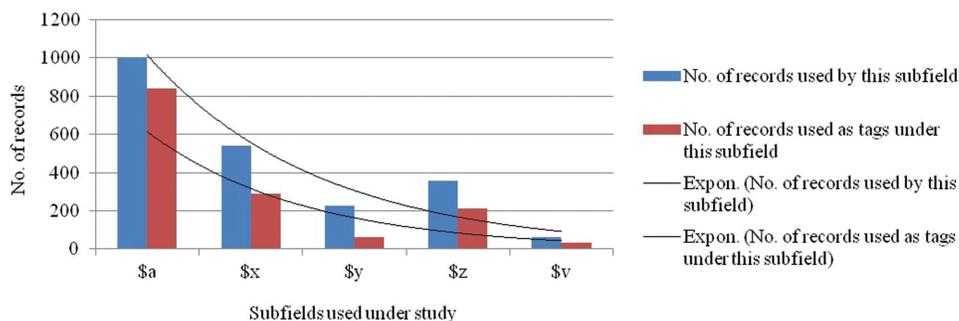


Figure 5. Usage of subfields by experts and users comparatively.

Table 5. Title wise occurrence of social tags and LCSH descriptors

| | Total terms | Unique terms |
|-------------|-------------|--------------|
| Social tags | 1264 | 444 |
| LCSH terms | 603 | 175 |

clarifies that \$a contains major book titles (839) followed by \$x (291), \$z (211), \$y (63) and \$v (36) which have at least one tag appeared in this subfield. That means in the case of subfield \$a, there is at least one term matched with the social tags of each record or title. The contents of subfield \$a are most popular to the users and others are as follows \$x, \$z, \$y and \$v. Figure 5 indicates the usage of subfields by experts and users comparatively, where a downline is found. That means \$a contains the most titles where as \$v contains the least.

5.7 Social Tags and LCSH Descriptors Compared with Each Book Title

Titles of documents are used as an effective information retrieval tool over a prolonged period. Title based searches not only retrieve desired documents but discover those documents which are not retrieval by subject-based search²¹. Moreover, the majority of searched documents can be retrievable through the keywords used in the titles²². The present study also compares social tags with LCSH descriptors which are applied to each book title. Table 5 indicates that 444 unique social tags and 175 unique LCSH descriptors appeared on the titles which are 14.88 per cent of total unique tag vocabulary & 21.77 per cent of total unique LCSH descriptor vocabulary respectively. Also, the study focuses that 806 (80.6 %) books have one tag appeared in their titles and 499(49.9 %) books have one term appeared in their titles. That means users focus mostly on title keywords rather than experts for cataloguing of books.

6. CONCLUSIONS

The overall comparison between social tags and LCSH descriptors draws many results regarding the effectiveness and usability of social tags into the library database.

The vocabulary overlapping clarifies that tag vocabulary is large than the LCSHs database. Only 381

unique terms which are a small portion of social tags (12.77 %) and near about half of LCSH descriptors (47.39 %) used by both users and experts for the entire collection. That means users mostly use descriptors as tags for describing resources, but experts use very little social tags as descriptors. In the case of overlapping terms, Spearman’s rank correlation suggests that when the term is used as social tags, there is a 75 per cent chance to be used as descriptors. The study also clarifies that there is a vocabulary difference between both datasets.

The similarity between top frequent social tags and top frequent LCSH descriptors is very low ($J = 0.13, 0.18, 0.15, 0.15$ and 0.14) in different top frequent word levels (100, 200, 300, 400, 500) which indicates that users and experts use different terminologies for annotating the books. Terminological comparison clarifies that experts mainly use subject-based terminologies than users. Users use some personal tags such as ‘read’, ‘to-read’, ‘unread’ etc. The study also finds that out of the top twenty frequent terms, seven terms (07) are common of which five terms (05) are subject-based

terminologies. Frequency analysis clarifies that users mostly use subject-based terminologies in most titles of the study than experts.

More than ninety percent (90.8 %) of book records where users and experts use at least one common term for annotating the books and the rest (9.2 %) where is no match that means both use different terminologies out of thousand book titles. Out of match records, 100 per cent matching for major books (247) and there are also different matching levels. Moreover, there are more than 50 per cent of book records (582) out of thousand which have fifty to a hundred percent matching. The study also clarifies that subfield \$a is used by the experts for describing all the titles under the Library of Congress MARC database. The comparison of each subfield with social tags for each record determines that social tags majorly matched with the contents of the subfield \$a (at least one tag matched) for 839 titles, then \$x for 291 titles, \$z for 211 titles, \$y for 63 titles and \$v for 36 titles. That means the contents of subfield \$a are equally popular to the experts as well as to the social taggers. The similarity between the contents of subfields and social tags for each record concludes that social tags can be used to enhance the subject access of library collection as alternative metadata.

The comparison of both dataset in respect of each title clarifies that more than 80 per cent (80.6 %) of books where at least one tag appeared in titles and near 50 per cent of books (49.9 %) where at least one descriptor appeared in titles which summarises that mainly users sought title keywords for the description rather than the experts, those use mainly subject based keywords for subject headings.

Moreover, the study indicates that the inclusion of social tags can improve the overall library experience of users²³, If the social tags contain more subject-based keywords, it could enhance the subject access of library collections; but cannot replace the controlled vocabulary like LCSHs¹⁷.

REFERENCES

- Lu, Caimei.; Park, Jung-ran.; Hu, Xiaohua & Song, Il-Yeol. Metadata effectiveness: A comparison between user-created social tags and author-provided metadata. *In* 2010 43rd Hawaii International Conference: System Sciences, 5-8 January 2010, Honolulu, Hawaii. 2010. <https://ieeexplore.ieee.org/abstract/document/5428516> (accessed on 2 January 2019) doi: 10.1109/HICSS.2010.273.
- Ramesh, P.; Vivekavardhan, J. & Bharathi, K. Metadata diversity, interoperability and resource discovery issues and challenges, *DESIDOC J. Libr. Inf. Technol.*, 2015, **35**(3), 193-199. doi: 10.14429/djlit.35.3.8074.
- Smith-Yoshimura, Karen. Social Metadata for Libraries, Archives, and Museums: Executive Summary, Dublin, Ohio: OCLC Research, 2011 <http://www.oclc.org/research/publications/library/2012/2012-02.pdf> (Accessed on 10 January 2019)
- Lu, C.; Park, J.R. & Hu, X. User tags versus expert-assigned subject terms: A comparison of LibraryThing tags and Library of Congress subject headings. *J. Inf. Sci.*, 2010, **36**(6), 763-779. doi: 10.1177/0165551510386173.
- Park, J.R. & Lu, C. Application of semi-automatic metadata generation in libraries: Types, tools, and techniques. *Libr. Inf. Sci. Res.*, 2009, **31**(4), 225-231. doi: 10.1016/j.lisr.2009.05.002.
- Godoy, D.; Rodriguez, G. & Scavuzzo, F. Leveraging semantic similarity for folksonomy-based recommendation. *IEEE Internet Comput.*, 2014, **18**(1), 48-55. doi: 10.1109/MIC.2013.26.
- Voorbij, H. The value of LibraryThing tags for academic libraries. Online information review., 2012, **36**(2), 196-217. doi: 10.1108/14684521211229039.
- Wal, T.V. Folksonomy. <http://vanderwal.net/folksonomy.html> (accessed on 12 January 2019).
- Furnas, G.W.; Landauer, T.K.; Gomez, L.M.; & Dumais, S.T. The vocabulary problem in human-system communication. *Commun. ACM.*, 1987, **30**(11), 964-971. doi: 10.1145/32206.32212.
- Librarything. <https://www.librarything.com/zeitgeist> (accessed on 12 January 2019).
- Hoe-Lian Goh, D.; Chua, A.; Sian Lee, C. & Razikin, K. Resource discovery through social tagging: A classification and content analytic approach. *Online Inf. Rev.*, 2009, **33**(3), 568-583. doi: 10.1108/14684520910969961.
- Suchanek, Fabian M.; Vojnovic, Milan. & Gunawardena, Dinan. Social tags: Meaning and suggestions. *In* Proceedings of the 17th ACM conference: Information and knowledge management, 26-30 October 2008, California, USA. 2008. pp. 223-232. doi: 10.1145/1458082.1458114.
- Macgregor, G. & McCulloch, E. Collaborative tagging as a knowledge organisation and resource discovery tool., *Libr. Rev.*, 2006, **55**(5), 291-300. doi: 10.1108/00242530610667558.
- Golder, S.A. & Huberman, B.A. Usage patterns of collaborative tagging systems. *J. Inf. Sci.*, 2006, **32**(2), 198-208. doi: 10.1177/0165551506062337.
- Statista (July, 2017). Number of social media users worldwide from 2010 to 2021 (in billions) <https://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/> (Accessed on 25 January 2019)
- Rolla, P.J. User tags versus subject headings. *Libr. Resour. Tech. Serv.*, 2011, **53**(3), 174-184. doi: 10.5860/lrts.53n3.174.
- Heymann, Paul & Garcia-Molina, Hector. Contrasting controlled vocabulary and tagging: Do Experts choose the right names to label the wrong things? *In* Second ACM International Conference: Web Search and Data Mining (WSDM 2009), 9-13 February 2009, Barcelona, Spain. 2009. <http://ilpubs.stanford.edu:8090/955/1/cvuv-lbrp.pdf> (Accessed on 23 January 2019).
- Yi, K. & Mai Chan, L. Linking folksonomy to Library of Congress subject headings: An exploratory study. *J. Doc.*, 2009, **65**(6), 872-900.

doi: 10.1108/00220410910998906.

19. Thomas, M.; Caudle, D.M. & Schmitz, C.M. To tag or not to tag? *Libr. Hi. Tech.*, 2009, **27**(3), 411-434.
doi: 10.1108/07378830910988540.
20. Lawson, K.G. Mining social tagging data for enhanced subject access for readers and researchers. *J. Acad. Libr.*, 2009, **35**(6), 574-582.
doi: 10.1016/j.acalib.2009.08.020.
21. Aanonson, J. A comparison of keyword subject searching on six British university OPACs. *Online Rev.*, 1987, **11**(5), 303-313.
doi: 10.1108/eb024248.
22. Bottle, R.T. & Preibish, C.I. The proposed KWIC index for psychology: An experimental test of its effectiveness. *J. Am. Soc. Inf. Sci.*, 1970, **21**(6), 427-428.

CONTRIBUTORS

Mr Kalyan Sundar Samanta is currently working as Librarian in Prabhu Jagatbandhu College, Howrah, West Bengal. He is also pursuing PhD from Dept. of Library & Information Science, Vidyasagar University, Midnapore, West Bengal. He has obtained MLISc, and BLISc, both from Vidyasagar University. He has few articles published in Journals and Conference Proceedings.

His contribution to the current study is that he carried out the literature review, data collection & analysis and prepared the initial draft of the study.

Dr Durga Sankar Rath is working as Professor in the Department of Library and Information Science, Vidyasagar University, Midnapore, West Bengal since December 2012. He has obtained PhD & M.Com from Vidyasagar University and ADIS from Indian Statistical Institute. He has more than twenty articles published in National and International Journals and authored four books. His area of interest is Computer application in LIS, library classification and application of artificial intelligence in LIS.

His contribution to this study is the conceptualisation, improvement in contents and preparation of the final draft of the study.