

Controlled Terms Versus Uncontrolled Terms in Resource Description: A Comparative Study Based on Social Science Books

Kalyan Sundar Samanta^{1*} and Durga Sankar Rath²

¹*Prabhu Jagatbandhu College, Howrah-711 302, India*

²*Department of Library and Information Science, Vidyasagar University, Midnapore-721 102, India*

*E-mail: *kalyansundarsamanta@gmail.com*

ABSTRACT

The paper comparatively investigates the relation between controlled vocabularies assigned by the experts in Library of Congress and tags assigned by users in Library Thing database in three subjects, Economics, History and Sociology under Social Science domain. Based on Term matching (S= 14.80 %, E= 12.77 % and H= 8.06 %) and Jaccard similarity coefficient (E= 0.15, S= 0.15 and H= 0.11), we found little matching between both vocabularies. We also found that experts mostly use double-word and multi-word specific topical terms (S= 73.14 %, E= 72.89 % and H= 61.05 %), whereas social taggers mostly use single-word general non-topical terms (E= 54.88 %, H= 54.21 % and S= 48.55%) and little topical and few personal terms. While comparison with LCSH subfield, we found that experts prefer topical terms for all subjects, whereas, taggers only prefer it for Economics and geographic subdivision terms for History and Sociology, but they don't prefer chronological terms for tagging. Even, experts prefer little title-based terms (H= 196 terms, S= 195 terms and E= 175 terms) but taggers mostly prefer title-based terms (H= 673 terms, S= 564 terms and E= 444 terms) in three subjects. However, the study concludes that both vocabularies are different, but libraries can exploit those uncontrolled vocabularies and can introduce 'hybrid metadata ecology' which combines controlled vocabularies, classification and folksonomies for better subject access and retrieval of social science documents.

Keywords: Social tagging; Social tagging in social science; Library of congress subject headings (LCSHs); Controlled and uncontrolled vocabulary

1. INTRODUCTION

Since a long time, finding a document by its subject is a common search practice in library catalogues and different bibliographic databases. Though subject-based search is challenging but still users prefer along with author and title-based searches¹⁻³. Subject-based searches led users to retrieve documents by its subject. With the rapid growth of literature, it is essential for subject indexers or cataloguers to enhance the subject access of library materials so that subject-based retrieval could be possible⁴. Enhancement of subject access is possible through the efficient exploitation of subject metadata that derives from subject cataloguing or subject indexing which provides a direct approach to find a document or group of documents based on subject⁵⁻¹⁰. To make the cataloguing process effective, subject cataloguing first determines the content of documents and then allows cataloguers or indexers to use controlled vocabularies which is a list of standard terms as subject descriptors like list of subject headings used in document description in order to ensure uniformity, universality and discoverability of library records in library catalogues and other bibliographic

databases¹¹⁻¹³. Some researchers still identified that subject access has been given the least priority than optimal in library catalogues and commercial bibliographic databases which results to several retrieval failures¹⁴. The failures may come from both sides, either search query is not formulated properly by users or the indexed documents are not defined with adequate subject terms¹⁵⁻¹⁶. To alleviate this problem, cataloguers must prefer hybrid approach of both controlled and uncontrolled (natural languages) vocabularies simultaneously like modern search engines (Google, Yahoo and Bing etc.) prefer hybrid approach for document indexing which led them to enhance the subject access by matching indexed documents with diversified user queries¹⁷. In that context, a parallel concept as social tagging (a process of social cataloguing) has emerged with the appearance of web 2.0 applications which allows natural languages in the form of keywords for document indexing. Social tagging allows users to assign keywords as tags and other kinds of metadata as per their need that helps to retrieve those resources in future (Sample Social tags mentioned in Fig. 1)^{7,18-19}. Social tagging differs from traditional subject cataloguing in many ways but the core philosophy is that any users can assign tags to any web resources using any keywords (uncontrolled terms) but, in case of subject cataloguing,

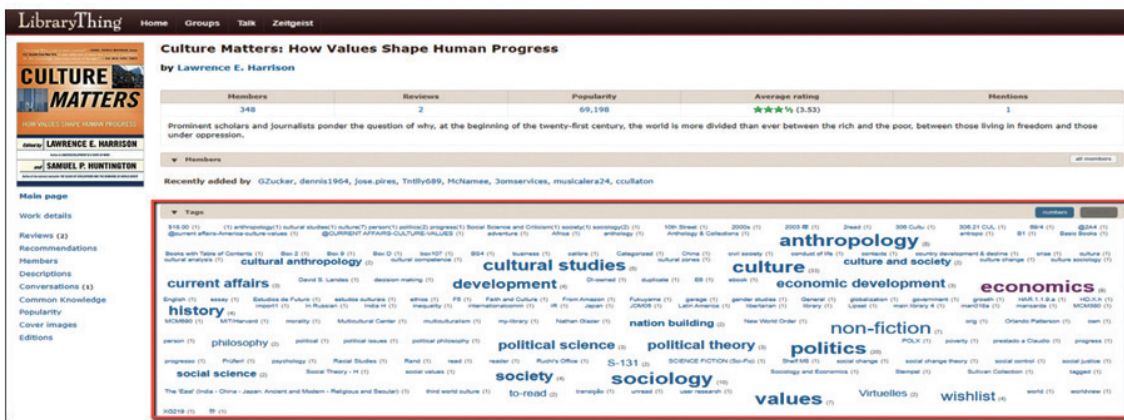


Figure 1. Social tags under the title “Culture Matters: How Values Shape Human Progress” by Lawrence E. Harrison (<https://www.librarything.com/work/174367>).

only librarians, trained cataloguers only conduct the process of subject cataloguing using controlled terms²⁰.

Though, being originated from uncontrolled environment, social tags suffer from inherent issues, but still several information scientists identified their effectiveness and prescribed to incorporate them into libraries for document indexing²¹⁻²². Further, with the rapid growth of social science literature, it is important for librarians and cataloguers to provide quality subject access to those heterogeneous resources. Several researchers tried to measure both vocabularies from different perspective in different domains, but no study carried out in this direction. Therefore, the main goal of the study is to investigate the relation between both controlled and uncontrolled vocabularies and their effectiveness in three social science subjects.

2. LITERATURE REVIEW

Several information scientists have identified and discussed the importance and effectiveness of social tags in different domains in different ways. Some researchers like Heymann, *et al.*; Zubiaga, *et al.*; Bischoff, *et al.*; Noll and Meinel; and Syn and Spring determined social tags as subject metadata can be applied in libraries. Heymann, *et al.* compared Open Directory Project and Delicious bookmarking systems and found that experts take more time in metadata generation than bookmarking in Delicious system²³. Zubiaga, *et al.* also identified that user-generated annotation is better for categorising web pages²⁴. Bischoff, *et al.* found that social tags are better than expert-assigned metadata in the domain of music²¹. Noll and Meinel also compared social tags and author-generated keywords for web documents and found that assigned tags are closer than author-generated metadata for the same document²². Even, Syn and Spring identified that social tags are equal to author-generated keywords for describing academic papers²⁵. Again, some researchers like Yu and Chen; and Hoe Lian Goh, *et al.* also identified that social tags can improve the quality of metadata and facilitate in resource discovery. Yu and Chen tried to enhance library subject headings using folksonomy

tags and found that subject headings can be enriched by folksonomy tags²⁶. Hoe Lian Goh, *et al.* tried to find the effectiveness of 150 popular del.icio.us tags on resource discovery and they found that out of 150 tags, some tags are effective and can be used for resource discovery²⁷.

Further, some researchers like Lu, *et al.*; Carman; Thomas, Caudle and Schmitz; and Yi and Chan compared both LC subject headings with library thing tags and LC subject headings with social tags to identify the similarities and dissimilarities between them. Lu, *et al.* compared LC subject headings with Library Thing tags to measure the relation between them. They collected 176105 unique tags and 7628 LCSH descriptors for the comparison and found that a small portion (2.2 %) of social tags matched with LCSH descriptors which shows that metadata assigned by both experts and users are different. Moreover, the study recommends using both social tags and LC subject descriptors, libraries can enhance the subject access of library collections²⁸. Carman compared both LC subject headings and Library Thing tags in ten fiction and fantasy books and found good matching between both vocabularies²⁹. Thomas, Caudle and Schmitz also compared both metadata in order to identify whether both vocabularies complement each other and if so, to which extent they complement each other. They found that social tags provide additional subject access by providing terms other than experts³⁰. Yi and Chan compared folksonomy tags and LC subject headings in order to identify similarities between them. They found that two-thirds of all the tags matched with LCSH terms³¹.

Several studies identified the effectiveness of social tags and its relation with LC subject headings in different domains, but the present study is different from them because no such study even measures and compares the application and effectiveness of social tags in three subjects Economics, History and Sociology under social science domain. With that context, our study can shed light on the possibility of resource description using uncontrolled terms (social tags) in social science.

3. OBJECTIVES

We formulated some basic objectives based on which we carried out our investigation as follows:

- to identify whether uncontrolled terms are similar to controlled terms in Social Science.
- to identify whether uncontrolled terms can enhance the subject access to Social Science documents.
- to know whether uncontrolled terms complement controlled terms in the description and retrieval of Social Science books.

4. METHODOLOGY

The main purpose of this study is to identify whether uncontrolled terms (social tags) can be used with controlled terms (LCSH descriptors) in resource description in social science domain. To carry out the entire evaluation, the study has chosen three social science subjects like Economics, History and Sociology and two popular online databases like LibraryThing (<https://www.librarything.com/>) and the Library of Congress Online Catalogue (<https://catalog.loc.gov>). We preferred LibraryThing database (a social cataloguing website) for collecting social tags and LC Online Catalogue for collecting LCSH descriptors.

We have chosen thousand book titles (1000) from each three subject like Economics, History and Sociology and in total, three thousand books (3000) from Social Science. We developed few parameters before selecting books from the LibraryThing database, those books were selected that have at least three tags (≥ 3) assigned by social taggers in the tag display section and that have at least catalogued by ten members (≥ 10). Even, selecting books from LOC database, we preferred those books that have at least one subject descriptor (≥ 1) present under the MARC field 650 (Subject Added Entry - Topical Term) in the LOC database. Though, LC MARC database contains several fields under 6XX to represent subject-related information, but we preferred only those book titles that have field 650 (Subject Added Entry – Topical Term) and field 651 (Subject Added Entry – Geographic Name).

We only selected those tags in LibraryThing database that contain a tag frequency at least two (≥ 2) or more than that under the assigned tags section. Tag frequency reveals the usage of tags, which means the number of times the tag is being used for description. Before collecting descriptors from LOC database, we ensured that under field 650, the first indicator (level of subject) will be zero (0) to two (2), and the second

indicator will be zero (0), whereas, in case of field 651, the first indicator will be # (undefined) and the second indicator will be zero (0) to represent the Library of Congress Subject Headings. In traditional cataloguing systems, the experts create subject headings of books in general to specific form which looks like a string of words concatenated by a hyphen (-), e.g., *United States--History--Revolution, 1775-1783*, but the same is represented in MARC format through different subfields like |a United States |x History |y Revolution, 1775-1783. Each MARC field contain a few subfields that represent different subject facets, and those subject facets entirely form the subject heading under a given book. We only have chosen five subfields like \$a- Topical term or geographic name entry element; \$x- General subdivision; \$y- Chronological subdivision; \$z- Geographic subdivision; \$v- Form subdivision under field 650 and 651. The descriptors that appeared on those subfields were collected individually. While collecting those descriptors, it was found that some subject descriptors appeared in more than one subfield. In the below example, a book entitled ‘Why Nations Fail: The Origins of Power, Prosperity and Poverty’ (LCCN-2011023538), where it is found that the descriptor ‘developing countries’ appears on both subfield \$a and \$z simultaneously under MARC data.

650_0 |a Poverty |z Developing countries.

651_0 |a Developing countries |x Economic policy.

In that case, we collected that descriptor once per book. Even after collecting descriptors from different subfields individually, we accumulated all those descriptors and selected unique descriptor book-wise²⁸. Further, descriptors that appear with scope notes, like ‘*Elite (Psychology)*’ have been taken without scope notes in order to make parity with social tags because social taggers do not use scope notes for tags.

However, based on those selected parameters, we collected a good number of social tags and LCSH descriptors from both databases under three subjects separately (mentioned in Table 1). After removing duplicates, we prepared two datasets of unique social tags and unique LCSH descriptors based in three subjects separately based on their frequencies for the evaluation.

5. DATA ANALYSIS AND INTERPRETATION

5.1 Terminological Matching in Social Science

We compared both social tags and LCSH descriptor vocabulary to identify similarities and dissimilarities between them. We found that (mentioned in Table 2)

Table 1. Total terms and unique terms in three subjects

	Economics (E)		History (H)		Sociology (S)	
	TT	UT	TT	UT	TT	UT
ST	20699	2983 (14.41% of TT)	41313	6123 (14.82% of TT)	30292	4655 (15.37% of TT)
LCSH	4144	804 (19.40% of TT)	5387	1186 (22.02% of TT)	4508	1132 (22.11% of TT)

[TT = Total terms and UT = Unique terms]

in one side, overlapping terms comprise a small portion of social tags where Sociology (S= 14.80 %) has more matching than Economics (E= 12.77 %) and History (H= 8.06 %), respectively. Another side, overlapping terms comprise a major portion of LCSH descriptors under three subjects where Sociology (S= 60.87 %) has higher matching than Economics (E= 47.39 % and History (H= 41.65 %). That means there is more than 40 % chance in three subjects, where LCSH descriptors can be used as tags by taggers.

5.2 Rank Correlation of Matching Terms in Social Science

The Spearman’s Rank correlation of matching terms (mentioned in Fig 2) reveals that a strong rank correlation exists between terms under three subjects, where it is found that Sociology (r = 0.89) has a comparatively higher rank correlation than History (r = 0.88) and Economics (r = 0.75). That means matching terms have a strong association under Sociology than History and Economics. The following equation was used for the Speaman’s rank correlation with correction factors.

$$r = 1 - \frac{6 \sum d^2}{n(n^2 - 1)} + CF_1 + CF_2 \tag{1}$$

Correction Factor for Social tags (CF_1)

$$CF_1 = \frac{m(m^2 - 1)}{12} \tag{2}$$

Correction Factor for LCSH descriptors term (CF_2)

$$CF_2 = \frac{m(m^2 - 1)}{12} \tag{3}$$

Table 2. Subject-wise terminological matching scenario

	Economics (E)			History (H)			Sociology (S)		
	UT	MT	%	UT	MT	%	UT	MT	%
ST	2983	381	12.77	6123	494	8.07	4655	689	14.8
LCSH	804		47.39	1186		41.65	1132		60.87

[UT = Unique terms & MT = Matching terms]

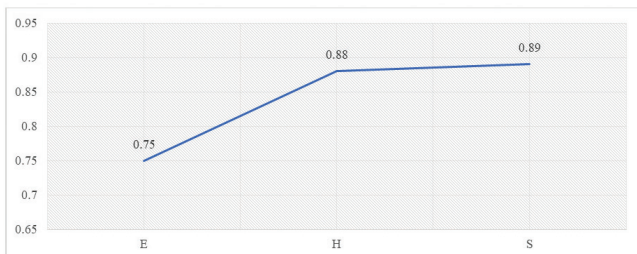


Figure 2. Spearman’s rank correlation in three subjects.

5.3 Jaccard Similarity Coefficient in Five Clusters of Terms in Social Science

Further, we tried to measure the similarity of highly used tags and descriptors between both vocabularies in three subjects using Jaccard Similarity Coefficient (J). We segregated highly used five hundred terms into five cluster of terms like top 100 terms, top 200 terms, top 300 terms, top 400 terms and top 500 terms. We found that (through Fig 3) Jaccard coefficient (J) is comparatively less under three subjects, where Economics (J= 0.15) and Sociology (J= 0.15) have higher average Jaccard coefficient (J) than History (J= 0.11). That means little similarity exists among top frequent terms between both datasets under three subjects.

The following equation was used for Jaccard Coefficient (J).

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|} \tag{4}$$

(Where A = social tags and B = LCSH descriptors)

5.4 Categories of Social Tags and LCSH Descriptors in Social Science

We compared both vocabularies to identify the categories like single-word, double-word and multi-word terms and preferences like topical, non-topical and personal terms assigned by social taggers and experts. We found that (mentioned in Table 3) social taggers mostly prefer single-word terms, whereas, experts mostly prefer double-word and multi-word terms. We also found that (mentioned in Table 4) experts use more topical terms (S= 73.14 %, E= 72.89 % and H= 61.05 %) than social taggers (S= 47.41 %, H= 42.12 % and E= 39.22 %), whereas, social taggers use more terms (E= 54.88 %, H= 54.21 % and S= 48.55 %) and some

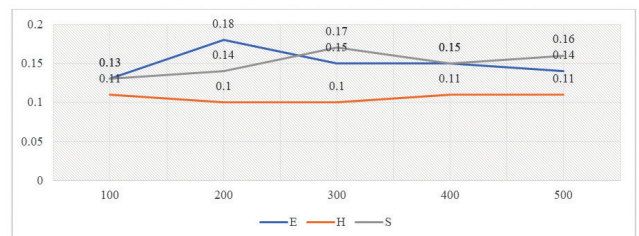


Figure 3. Jaccard similarity coefficient in five clusters of terms.

personal terms for their interest. Further, through term frequency analysis, we found that (mentioned in Table 5) social taggers mostly used topical terms whereas, experts used specific topical terms³².

5.5 Individual Book-wise Comparison of Social Tags with LCSH Descriptors in Social Science

We also compared social tags with LCSH descriptors at individual book level for three thousand books taken under the study. We found different percentage of matching starting from 0 % to 100 % between both vocabularies. We found that (through Fig. 4) the subject Economics (E= 92) has the lowest number of books than Sociology (S= 119) and History (H= 126), which have 0 % matching between both vocabularies. Again, the subject Economics contains major books (E= 247) which have 100 % matching between both vocabularies than Sociology (S= 202) and History (H= 78) respectively. Moreover, we found that all three subjects (E= 582, S= 555 and H= 512) contain more than fifty percent books that have 50 to 100 percent matching.

5.6 Social Tags Compared with LCSH Subdivisions in Social Science

We also tried to identify whether social taggers assign the similar types of terms as experts prefer to assign in MARC subfields. We found that (through Table 6) experts highly prefer subfield \$a (S= 958 terms, E= 726 terms and H= 699 terms) and less prefer subfield \$v terms (H= 23 terms, E= 22 terms and S= 17 terms), whereas, social taggers mostly prefer \$a terms as tags under Economics (85.40 %) and subfield \$z terms for History (74.1 %) and Sociology (71.64 %) and less prefer \$v terms as tags in Economics (27.27 %) and subfield \$y for History (5.5%) and Sociology (9.8 %). That means, social taggers prefer subfield \$a terms as tags but they don't prefer subfield \$y terms for tags^{28,32}.

5.7 Social Tags & LCSH Descriptors Compared with Each Book Title in Social Science

We compared social tags and LCSH descriptors with title words of three thousand books in order to identify whether experts and social taggers prefer title-based terms or not for document description. The quantity

Table 3. Categories of social tags and LCSH descriptors based on word length

Subjects	Single-word terms		Double-word terms		Multi-word terms	
	ST	LCSH	ST	LCSH	ST	LCSH
E	1546 (51.83%)	277 (34.45%)	1235 (41.4%)	391 (48.63%)	202 (6.77%)	136 (16.92%)
H	2596 (42.4%)	527 (44.44%)	2696 (44.03%)	362 (30.52%)	831 (13.57%)	297 (25.04%)
S	2258 (48.51%)	436 (38.52%)	1976 (42.49%)	466 (41.17%)	421 (9%)	230 (20.32%)

[E = Economics, H = History, S = Sociology]

Table 4. Categories of social tags and LCSH descriptors based on topical, non-topical and personal terms

Subjects	Topical terms		Non-topical terms		Personal terms	
	ST	LCSH	ST	LCSH	ST	LCSH
E	1170 (39.22%)	586 (72.89%)	1637 (54.88%)	218 (27.11%)	176 (5.9%)	0
H	2579 (42.12%)	724 (61.05%)	3321 (54.24%)	462 (38.95%)	223 (3.64%)	0
S	2207 (47.41%)	828 (73.14%)	2260 (48.55%)	304 (26.86%)	188 (4.04%)	0

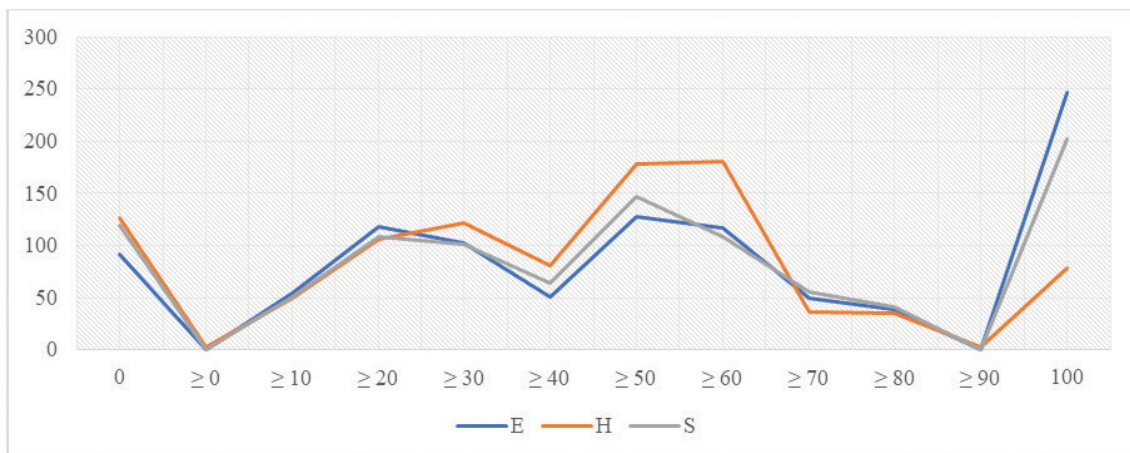


Figure 4. Individual book-wise matching of social tags with LCSH descriptors for three subjects.

Table 6. Terms appeared on LCSH subfields

	Economics (E)					History (H)					Sociology (S)				
Total no. of terms appeared	1810	925	172	116	78	2216	1679	990	657	189	2850	1146	214	438	65
Unique terms	726	223	76	54	22	699	169	291	139	23	958	118	51	67	17
No. of terms used as tags	620	102	24	18	6	355	73	16	103	8	543	54	5	48	10
%	85.4	45.74	31.58	33.33	27.27	50.79	43.2	5.5	74.1	34.78	56.68	45.76	9.8	71.6	58.82

Table 7. Total terms & unique terms appearing on book titles subject wise

	Economics (E)		History (H)		Sociology (S)	
	TA	UT	TA	UT	TA	UT
Social tags	1264	444 (14.88 % of TUST)	1457	673 (10.99 % of TUST)	1284	564 (12.12 % of TUST)
LCSH descriptors	603	175 (21.77 % of TULCSH)	481	196 (16.53 % of TULCSH)	425	195 (17.23 % of TULCSH)

[TA = Terms appeared, UT= Unique Terms, TUST= Total Unique Social tags, TULCSH= Total Unique LCSH descriptors]

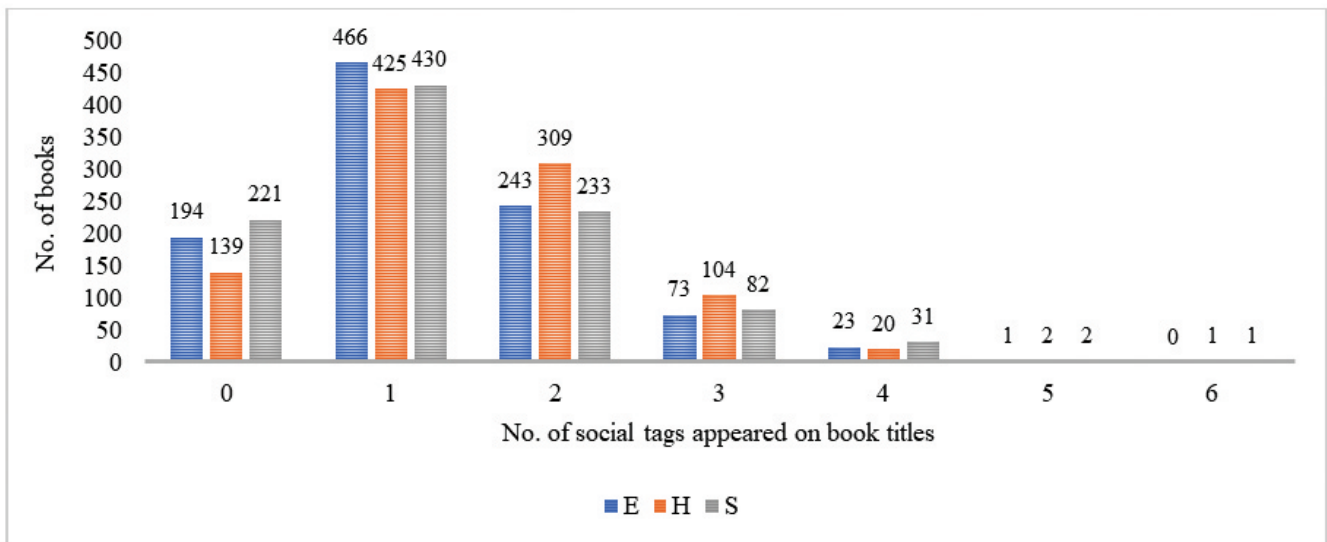


Figure 5. Scenario of social tags appearing on book titles.

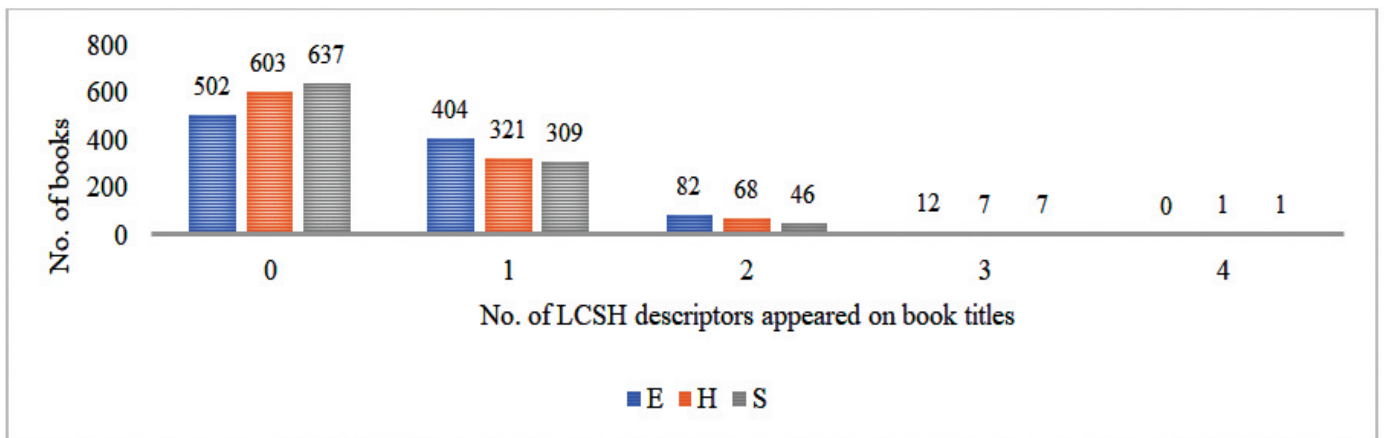


Figure 6. Scenario of LCSH descriptor appearing on book titles.

of terms appeared in both vocabularies (mentioned in Table 7) under three subjects reveals that social taggers use more title-based terms than experts, comparatively high in History (H = 673) than Sociology (S = 564) and Economics (E = 444). Though social taggers prefer title-based terms, we found that social tag vocabulary contains 15 % terms as title-based terms and other 85 % terms as non-title-based terms. Again, we measured the range of tags and descriptors those appeared on book titles. We found that (through Fig 5 & Fig 6) major books (E = 466, S = 430 and H = 425) have one tag appeared on book titles, whereas, major books (S = 637, H = 603 and E = 502) have zero descriptors appeared on book titles (S = 637, H = 603 and E = 502)^{28,33-34}.

6. RESULT AND DISCUSSION

We investigated the effectiveness and applicability of uncontrolled terms in comparison with controlled terms in document description under three subject like Economics, History and Sociology in Social Science domain. We found that History subject taggers assign more tags than Sociology and Economics subject taggers which shows that all three subject taggers have good intentions towards social tagging but comparatively, History subject taggers are more active in tagging than other subject taggers. Though there is a good number of tags assigned but the similarity level between both vocabularies is very less. Again, the uncontrolled vocabulary is very large and contains more than 80% of terms other than controlled terms for three subjects. Again, though History subject taggers are more active for assigning tags but Sociology and Economics subject taggers perform better than History in vocabulary matching scenario. Further, through our study, we identified tagging preferences among social taggers and experts in three subjects like social taggers mostly prefer good number of single-word non-topical terms, small percentage of topical terms and personal terms, whereas, experts prefer mostly double-word and multi-word topical terms and little percentage of non-topical terms.

Frequency analysis also reveals that social taggers prefer general topical terms but experts prefer specific topical terms. Even, we also identified tagging preferences differ from subject to subject among social taggers like Economics subject taggers mostly prefer topical terms (\$a subfield terms) for tagging but they don't prefer form division terms (subfield \$v terms) as tags, whereas, History and Sociology subject taggers prefer geographic name terms (subfield \$z terms) for tagging but they don't prefer chronological terms (\$y) for tagging. Moreover, we found that social taggers don't prefer chronological terms and form division terms for tagging. Again, social taggers mostly prefer terms from document titles but experts prefer little terms from document titles under three subjects.

7. CONCLUSION

The present study tries to present a subject-wise

comparative scenario of social tagging application under social science domain. It is clearly identified that both controlled and uncontrolled vocabularies are mostly different and uncontrolled vocabulary is large and comprised of different terms other than expert terms and title terms which can be investigated in further studies. With the growing information needs of users and speedy growth of social science literature draw challenges among information professional about the subject access of those literature. In pursuance this, it is essential to revise the information organisation process, so an effective and user-centric information organisation can be developed that can augment better subject access as well as information retrieval experiences.

Further, the emergence and popularity of social tagging or social cataloguing indirectly force information scientists to rethink its incorporation in libraries. As a result, many information scientists recommend through their studies how tags can be incorporated into their library OPACs. Though uncontrolled terms majorly differ from controlled terms, those terms facilitate libraries to improve their resource description process by supplementing terms other than expert-generated controlled terms³⁵⁻³⁶. That means library catalogues would be an environment to introduce 'hybrid metadata ecology', which combines controlled vocabularies, classification and folksonomies³⁷. However, by incorporating both controlled and uncontrolled terms, libraries can extend their resource description process and enhance the subject access of social science documents in such a way that they can retrieve social science documents from various search dimensions. Through this way, uncontrolled terms could complement controlled terms and strengthen the capacities of libraries²³⁻³⁸.

REFERENCES

1. Fagan, J.C. Discovery tools and information literacy. *J. Web Libr.*, 2011, **5**(3), 171-178.
2. Hunter, R. N. Successes and failures of patrons searching the online catalog at a large academic library: A transaction log analysis. 1991, **30**(4), 395-402. <https://www.jstor.org/stable/25828813> (Accessed on 10 December 2023).
3. Siegfried, S.; Bates, M.J. & Wilde, D.N. A profile of end-user searching behavior by humanities scholars: The Getty Online Searching Project Report No. 2. *J. Am. Soc. Inf. Sci.*, 1993, **44**(5), 273-291.
4. Golub, K. Potential and challenges of subject access in libraries today on the example of Swedish libraries. *Int. Inf. Libr. Rev.*, 2016, **48**(3), 204-210.
5. Kumar, G. & Kumar, K. Theory of cataloguing. Vikas Publishing House, New Delhi, 1982.
6. Safari, M. Metadata and the Web. *Webology*, 2004, **1**(2). http://eprints.rclis.org/7318/1/Metadata_and_the_Web.pdf (Accessed on 5 May 2023).
7. 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.

8. Smith-Yoshimura, K. Social metadata for libraries, archives, and museums: Executive summary. OCLC, Dublin, Ohio, 2012. <https://www.oclc.org/content/dam/research/publications/library/2012/2012-02.pdf> (Accessed on 10 January 2023).
9. Sears, M.E. Sears list of subject headings 19th edition. HW Wilson, New York, 2007.
10. Sarkhel, J. Subject Cataloguing-Problems. IGNOU. <http://egyankosh.ac.in/handle/123456789/33119> (Accessed on 15 March 2023).
11. Ni, D. Subject Cataloging and Social Tagging in Library Systems. *J. Libr. Info. Sci.*, 2010, **36**(1), 4-15.
12. Smith, C. Controlled vocabularies: Past, present and future of subject access. *Catalog. Classi. Quar.*, 2021, **59**(2-3), 186-202. doi: 10.1080/01639374.2021.1881007.
13. Wu, D.; He, D.; Qui, J.; Lin, R. & Liu, Y. Comparing social tags with subject headings on annotating books: A study comparing the information science domain in English and Chinese. *J. Inf. Sci.*, 2013, **39**(2), 169-187.
14. Sauperl, A. & Saye, J.D. Subject determination during the cataloging process: An intensive study of five catalogers. *Adv. Classi. Res. Online*, 1998, **9**(1), 119-138.
15. Golub, K. Subject access in swedish discovery services. *Knowl. Org.*, 2018, **45**(4), 297-309. doi: 10.5771/0943-7444-2018-4-297.
16. Markey, K. The online library catalogue: Paradise lost and paradise regained? *D-Libr. Mag.*, 2007, **13**(1/2). doi: 10.1045/january2007-markey (Accessed on 10 December 2023).
17. Golub, K.; Tyrkko, J.; Hansson, J. & Ahlstrom, I. Subject indexing in humanities: A comparison between a local university repository and an international bibliographic service, *J. Doc.*, 2020, **76**(6), 1193-1214.
18. Furnas, G.W.; Landauer, T.K.; Gomez, L.M. & Dumais, S.T. The vocabulary problem in human-system communication. *Comm. ACM*, 1987, **30**(11), 964-971.
19. Hedden, H. How semantic tagging increases findability. *E-Content*, 2008, **31**(8), 38-43.
20. Steele, T. The new cooperative cataloging. *Libr. Hi Tech*, 2009, **27**(1), 68-77. doi: 10.1108/07378830910942928.
21. Bischoff, K. Can all tags be used for search? In 17th ACM Conference on Information and Knowledge Management, 26-30 October 2008, Napa Valley California, USA. 2008. pp. 193-202. doi: 10.1145/1458082.1458112 (Accessed on 20 December 2023).
22. Noll, M.G. & Meinel, C. Authors vs. readers: A comparative study of document metadata and content in the Www. In the 2007 ACM symposium on Document engineering, 28-31 August 2007, Winnipeg Manitoba, Canada. 2007. pp. 177-86. doi: 10.1145/1284420.1284465 (Accessed on 05 January 2024).
23. Heymann, P. & Garcia-Molina, H. Contrasting controlled vocabulary and tagging: Do experts choose the right names to label the wrong things? In Second ACM International Conference on Web Search and Data Mining, 9-12 February 2009, Barcelona, Spain. 2009. pp.1-4. <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=5e703880637f833bab0ff249371363052844c17d> (Accessed on 26 March 2023).
24. Zubiaga, A.; Martinez, R. & Fresno, V. Getting the most out of social annotations for web page classification. In 9th ACM Symposium on Document Engineering, 16-18 September 2009, Munich, Germany. 2009. pp. 74-83. doi: 10.1145/1600193.1600211 (Accessed on 15 November 2023).
25. Syn, S.Y. & Spring, M.B. Tags as keywords—comparison of the relative quality of tags and keywords. *Proc. Am. Soc. Inf. Sci. Technol.*, 2009, **46**(1), 1-19.
26. Yu, W. & Chen, J. Enriching the library subject headings with folksonomy. *Elect. Libr.*, 2020, **38**(2), 297-315. doi: 10.1108/EL-07-2019-0156.
27. 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), 568-83. <https://hdl.handle.net/10356/91019>.
28. Lu, C.; Park, J.R. & Hu, X. User tags versus expert-assigned subject terms: A comparison of library thing tags and Library of Congress Subject Headings. *J. Inf. Sci.*, 2010, **36**(6), 763-779. doi: 10.1177/0165551510386173.
29. Carman, N. Library thing tags and library of congress subject headings: A comparison of science fiction and fantasy works. University of Wellington. 2009. <http://researcharchive.vuw.ac.nz/handle/10063/1272>
30. 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.
31. Yi, K. & Chan, L.M. Linking folksonomy to library of congress subject headings: An exploratory study. *J. Doc.*, 2009, **65**(6), 872-900. doi: 10.1108/00220410910998906.
32. Rolla, P.J. User tags versus subject headings. *Libr. Resour. Tech. Ser.*, 2009, **53**(3), 174-184. doi: 10.5860/lrts.53n3.174.
33. Aanonson, J. A comparison of keyword subject searching on six British University Opacs. *Online Rev.*, 1987, **11**(5), 303-313. doi: 10.1108/eb024248.
34. 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.
35. 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.
36. Spiteri, L.F. The use of folksonomy in public library catalogues. *The Ser. Libra.*, 2006, **51**(2), 75-89. https://www.tandfonline.com/doi/abs/10.1300/J123v51n02_06 (Accessed on 10 March 2023).
37. Morville, P. *Ambient Findability: what we find changes who we become*, O reilly Media, 2005.
38. Samanta, K.S & Rath, D.S. User generated social tags versus librarian generated subject headings: A comparative study in the domain of history. *DESIDOC J. Libr. Inf. Technol.*, 2020, **40**(03), 176-184. doi: 10.14429/djlit.40.03.15413.

CONTRIBUTORS

Dr Kalyan Sundar Samanta has been working as Librarian at Prabhu Jagatbandhu College, Andul, Howrah, West Bengal, India.

He has obtained BLISc., MLISc., and PhD from Vidyasagar University. His area of interest is Information retrieval, ICT applications in libraries, and Open-source software applications in libraries. The present study is part of his doctoral research. 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 has been working as Professor in the Department of Library and Information Science, Vidyasagar University, Midnapore, West Bengal, since December 2012. He has obtained PhD & MCom from Vidyasagar University and ADIS from Indian Statistical Institute. His area of interest is Computer application in LIS, Library classification and Application of artificial intelligence in LIS. The present study has been conducted under the guidance of Dr Durga Sankar Rath. His contribution to this study is the conceptualisation, improvement in contents and preparation of the final draft of the study.

Table 5. Top ten frequent social tags and LCSH descriptors in social science

Economics	History					Sociology						
	Social tags Freq. in soc. tag. voc.	LC SH term voc.	LC SH terms in soc. tag. voc.	Social tags Freq. in soc. tag. voc.	LC SH term voc.	LC SH terms in soc. tag. voc.	Social tags Freq. in soc. tag. voc.	LC SH term voc.	LC SH terms in soc. tag. voc.	Social tags Freq. in soc. tag. voc.	LC SH term voc.	
economics	899	173	296	history	9941	633	history	633	994	863	328	-
non-fiction	838	-	239	non-fiction	902	-	united states	319	-	787	173	504
to-read	564	-	173	to-read	693	-	biography	187	340	622	164	73
history	557	239	138	American history	514	-	politics and govt.	151	43	504	105	787
politics	495	-	135	unread	468	-	world war, 1939-1945	131	-	467	97	82
business	393	-	114	read	461	-	campaigns	101	17	441	97	13
finance	355	37	86	usa	459	-	great britain	98	152	432	56	52
capitalism	343	135	79	military history	426	-	20th century	81	401	342	52	127
economy	311	-	62	us history	413	-	civil war, 1861-1865	75	-	337	41	334
economic history	280	86	48	20th century	401	81	19th century	72	275	334	36	230