

Technical Folksonomy of Tagging Information

Sonali Dapsi

Librarian, Raja Peary Mohan College, India.

Sudip Ranjan Hatua

Asst Prof. Rabindra Bharati University, Kolkata, India.

DOI: <http://dx.doi.org/10.21013/jems.v3.n3.p6>

How to cite this paper:

Dapsi, S., & Hatua, S. (2016). Technical Folksonomy of Tagging Information. *IRA International Journal of Education and Multidisciplinary Studies (ISSN 2455–2526)*, 3(3). doi:<http://dx.doi.org/10.21013/jems.v3.n3.p6>

© Institute of Research Advances



This work is licensed under a [Creative Commons Attribution-Non Commercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/) subject to proper citation to the publication source of the work.

Disclaimer: The scholarly papers as reviewed and published by the Institute of Research Advances (IRA) are the views and opinions of their respective authors and are not the views or opinions of the IRA. The IRA disclaims of any harm or loss caused due to the published content to any party.

ABSTRACT

The Study of the various articles published in Library and Information Science Journals in the resent times shows that the keywords provided by the authors along with their articles are mostly uncontrolled. They are basically phrases. In spite of the knowledge of controlled vocabulary and various subject heading scheme they mostly are using natural word and sentences to represent the thought content of their research outcomes. This is generating new trends of representing subjects known as technical folksonomy.

Keywords: Control vocabulary, Folksonomic Vocabulary, Technical Folksonomy.

1. Introduction:

The concept of Folksonomy has come into consideration after the usage of Computer where the new words are being uploaded .The term was coined in 2003 by an information architect, Thomas Vander Wal [10]. It is a neologism consisting of a combination of the words *folk* and *taxonomy*. Taxonomy is derived from the Greek words, *taxis* and *nomos*. *Taxis* mean Classification and *nomos* means Management. In spite of the knowledge of controlled vocabulary and various subject heading scheme they mostly are using natural word and sentences to represent the thought content of their research outcomes. This is generating new trends of representing subjects known as technical folksonomy.

2. Literature Review:-

Catherine Lyons viewed that combining cataloguing and other standard metadata practices with user-developed tags and Folksonomies is a good way to improve subject access to resources [9]. Slavko Knett told that the business world has of course discovered the considerable commercial potentials of folksonomy [13].

Ikki Ohmukai, Masahiro Hamasaki, and Hideaki Takeda said that social bookmark system using several metadata and personal network constructs a community-based ontology [2]. ZIXIN WU said that tagging communities are featured Web 2.0 phenomenon, where users describe a Web resource by using keywords (called tags) [11].

Terrell Russell said that Contextual Authority Tagging is the use of Folksonomies to discover and define Cognitive authority through reputation within communities of users [3]. Francisco Echarte, Jose Javier Astrain, Alberto Córdoba, Jesus Villadangos told that Ontology's and tagging systems are two different ways to organize the knowledge present in Web[6]. Alan Said, Robert Wetzker Winfried Umbrath and Leonhard Hennig investigated the problem and recommended during the first months of the collaborative tagging community Cite ULike[1]. Torben Knerr said that collaborative tagging represents the process by which many users describe resources (e.g. web pages or photos) with free-from keywords (tags). Web technologies to develop ontology for Folksonomies, making interoperability and automated processing feasible [10]. Emanuele Quintarelli said that Folksonomies attempt to provide a solution by introducing an innovative distributed approach based on social classification [5].

Massimiliano Dal Mas says that folksonomy gives an overview of current trends in manual indexing on the Web. Digital resources with tags (keywords) share their annotations with other users through tagging system [8]. Fabian Abel analyzed

the impact of tags on information retrieval [2]. Jesse Vig said that present tagging applications design the system [9]. Fabian Abels , Matteo Baldoni said that with the advent of Web 2.0 tagging became a popular feature in social media systems. People tag diverse kinds of content, e.g. products at Amazon [1]. Min Gyo Chung said that collaborative tagging activities that proposed scheme maintain video bookmarks, which contain some temporal or positional information about videos [4].

3. Problem Identification:-

1>Trends of usage phase, natural Language ,folksonomy are using in the authors in Library and Information Science journals are increasing since last two decades;

2>There is hardly found any attempt to build up a model system for folksonomical terms.

4. Hypothesis :-

1> Trends of usage of phrase, natural Language ,folksonomy by the authors in Library and Information Science journals in India are increasing since last two decades;

2> There is hardly found any attempt to build up a model system for folksonomical terms.

5. Objective:-

The objectives of this study are:-

- a. To identify keywords available in the articles of Library and Information Science journals published in India during 1998 to 2012 and use them for the folksonomical study.
- b. To find out the trends of usage and observe their pattern e.g. control vocabulary, phrases, folksonomy etc.
- c. To design a model system on folksonomic vocabulary.

6. Scope of the study:-

The scopes of this present work are listed as follows-

- The subject scope restricted to Library and Information Science
- The geographical coverage restricted to Library & Information Science journals published in India;
- The time coverage is 1998-2012;
- Journals published by professional Associations and Institutions and LIS departments in India;
- Language covers the English.

7. Methodology:-

For collecting necessary data for the study, survey method has been practiced. The research design of the present study is based on descriptive design and probability sampling method. The sampling procedure selected for the study is based on stratified sampling.

Data has been categorized as follows-

- i) Journals published by Library and Information Science Departments of various Universities.
- ii) Journals published by professional body and Association.
- iii) Journals published by Institution or organization.

Among LIS Departments sample selection has been made by following criteria –

- i) University selected from four zones East, West, North, South
- ii) Departments which are 25 years or more have considered from each zone.

For professional body only national body and Association have been considered who are publishing LIS Journals in English Language.

For Institution and organization have been considered who are publishing LIS Journals in English Language.

8. Data Collection:

Study has been made of each and individual journal from above set of samples. The Total numbers of Universities calculated all over India is 514 and Library and Information Science Department available within these Universities are 178. Among 178 Universities we have found only 57 departments are 25 years or more old. The 57 departments we have been consulted through direct communication and e-mail. Unfortunately most of them replied as they don't published journal regularly. Only 7 LIS departments found who publishing journals regularly in English Language are. Apart from the 7 department, 3 journals published by Professional Association in national level institutions have taken into consideration for this study.

No of Total University	514
No of LIS Departments	178
No of Departmental consulted (Sample sample size)	57
No of Department who published journal of LIS	7
No of National Association	3
No of Institution	3

Table 1: *Study in Research area (Sample size)*

. Finally we have listed 7 Departments which have published Library and Information Science Journals and provides keywords most of their articles during 1998-2012 time span. The list of those departments has given in the table 2.

SI No	University Name	Place	Journals Name
1	University of Delhi	Delhi	Journal of Library & Information Science (JLIS)
2	University of Kashmir	Kashmir	Trends in Information Management, (TRIM)
3	University of Calcutta	Kolkata	Calcutta University Journal of Information Studies(CUJLIS)
4	Rabindra Bharati University	Kolkata	RBU Journal of Library and Information Science(RJLIS)
5	University of North Bengal	North Bengal	Advances in Library and Information Science(NALIS)
6	Jadavpur University	Kolkata	Librarian: A Journal of Library and Information Science(LJLIS)
7	Vidyasagar university	Midnapur	VU journal of Library and Information science(VJLIS)

8	NISCAIR	Delhi	Annals of Library and Information Studies (ALIS)
9	Sarada Ranganathan Endowment for Library Science	Karnataka	SRELS Journal of Information Management(SJIM)
10	Defence Scientific Information & Documentation Centre	Delhi	DESIDOC Journal of Library & Information Technology(DJLIT)
11	Indian Association of Special Library and Information Centre	Kolkata	IASLIC Bulletin(IB)
12	Indian Library Association	Delhi	Journal of Indian Library Association(ILSB)
13	University Library Teacher's Association	Andhra Pradesh (Hydrabad)	Pearl : A Journal of Library and Information Science(PJLIS)

Table 2: Collecting data (sample size)

From the table 2 we have found that total 13 Population collected from the different articles during 1998-2012. We have been found the uses of data are sample size. By studying all the journals listed in table 2. We have found following data given in table 3

Date Collection Statistics:-

Number of Journal studies	Number of Volume	Number of articles	Total terms found	Unique No of terms	Total Uncontrolled vocabularies	Unique no of uncontrolled vocabulary
13	322	2638	3604	2159	2924	1958

Table 3: Data collection Statistics

The Used terms available with the article are 3604 among which only 680 terms are found in various popular standard vocabulary tools used laid the LIS professionals. The remaining 2924 terms are not found any of the available standard vocabulary tools. These may be called technical folksonomy. These terms are given in table 3.

9.1 Data Analysis:-

We have analysed the data as per following sequence-

- i) Each and every 2638 articles from 322 volumes have been thoroughly studied and listed 3604 terms/phrases etc. Provided by the author/ editor along with their respective article.
- ii) We have listed five established Controlled Vocabulary tools given in the table (4) and compare each and every term and phrases with these tools and find their availability or appearance on those tools.
- iii) Separated Controlled Vocabulary terms and uncontrolled Vocabulary.
- iv) Analyse the growth pattern of those listed uncontrolled vocabulary.
- v) Analyze the frequency of uncontrolled vocabulary.

Sl No	Title of the Books	Edition
1	Library of Congress Subject Heading	25 th ed,2002
2	Sears List of Subject Heading	10 th ed,1972
3	Dewey Decimal Classification	23ed, 2011
4	Colon Classification	6 th ed, 27 th reprint 2006
5	Thesaurus	http://www.thesaurus.com/ (retrieve from 01.3.2012)

Table 4: List of Comparative standard tools

.Year wise Growth of all control terms in all journals

Journal Name	98	99	00	01	02	03	04	05	06	07	08	09	10	11	12	Total
IB	0	0	0	1	5	7	4	6	18	26	19	26	17	20	32	181
ALIS											1			1	33	35
DJLIS														4	81	85
TIM								2	7	3	7	4	11	14	21	69
SJIM	11	20	13	16	26	14	11	12	09	04	10	17	15	16	05	199
PJLIS															17	17
LJLIS			1		1	3	4	4			3	1				17
RJLIS											1		3			4
CUJLIS					3	3						6	1		2	15
NALIS														2		2
JLIS													11			11
ILAB			10	9	5				3			1	7	8	2	45
	11	20	24	26	40	27	19	24	35	36	39	56	65	65	193	680

Table 5 Year wise Growth of all control terms in all journals

For consulting table 5, we can see from the above table that the Control term/ Keywords Phrases used in various Journals released during a span of fifteen years starting from the year 1998 to 2012. The used of control terms in the starting years 1998 and 1999 was nil, which gradually increase from the year 2000 and reached its maximum level in the year 2012 which was around 680 numbers.

Year wise growth of Uncontrolled term in all journals:

Journal Name	98	99	00	01	02	03	04	05	06	07	08	09	10	11	12	Total
IB				4	7	27		23	55	56	77	71	47	84	55	489
ALIS															45	45

DJLI S															1 7 8	178
TIM								1 0	2 3	2 8	2 7	4 3	4 4	8 2	3 4	291
SJIM	5 0	7 9	5 2	9 7	9 1	7 4	6 9	9 9	1 0	1 0	1 4	7 3	1 5	1 4	1 5	148 6
PJLI S															5 3	53
LJLI S			2	3	5	5	3	3	1		1		1	7		106
RJLI S										2	3		5			10
CUJ LIS				1	1 0	9	5	5	4	1 0	1 1	1 6	4	1	8	84
NAL IS													4			4
JLIS													2 2			22
ILA B			2 2	4 4	1 4					5		6	1 0	2 5	3 0	156
	5 0	7 9	7 6	1 4	1 2	1 1	1 0	1 6	1 8	2 1	2 5	2 0	2 9	3 4	5 5	292 4

Table6: Year wise growth of Uncontrolled term in all journals

For consulting table 6, we can see from the above table that the Uncontrolled term/ Keywords Phrases used in various Journals released during a span of fifteen years starting from the year 1998 to 2012. The used of Uncontrolled terms in the starting years 1998 and 1999 was nil, which gradually increase from the year 2000 and reached its maximum level in the year 2012 which was around 2924 numbers.

9.2 Considering Folksonomy:

Information communication		
Technology	36	UGC-INFONATE 9
Authorship Pattern	20	Information seeking 8
Scientometric	19	Behavior 8
Bibliometric Studies	16	Job Satisfaction 8
Citation Analysis	16	Metadata 8
Librarian	16	Analytico Synthetic 7
Library Service	12	Classification 7
World Wide Web	11	Author Productivity 7
Information Literacy	11	Internet Use 7
DSpace	10	Case Study 6
Publication productivity	10	Informatics 6
Publication productivity	10	Reading Habit 6
User Studies	10	Virtual Library 5
LIS Professional	10	Webometrics 5
Institutional Repository	10	Information Technology 5
		Electronic Books 5
		Electronic Learning 5

Table 4.15.1: Considering folksonomy

In view of the usage keywords used in the last fifteen years. We have found the usage of uncontrolled keywords have come in to consideration accounts. We have found that thirty terms are used 315 times. We can confuse that if a keywords is being used for more than ten tomes, then we can term them as “folksonomy “.but the gradation of commonly used keywords is not bound to more than 10 times. In recent future, these can be considered as “Folksonomy”. “Information Communication Technology” and “Scientrometrics”,” Author ship Pattern” Keywords are used 36 and 20 times. The keywords used less then5 to 10 times can be used as folksonomy of their usage are found to be rapidly increase.

There is no guideline or rules standard to identify a term as folksonomic term. We have no idea, after how many occurrences or after how many years a term becomes a folksonomic term. The general conceptions of a folksonomical term are that which generates automatically and accepted or used by the common people. For our consideration we have found four terms are used more than ten times. But is also not included any standard vocabulary tools. So we have considered these terms are folksonomic terms. Among of them “Information communication Technology” terms is used thirty six times. Rest of twenty eight terms are used in more than five times, so in future, we considered that these terms may be folksonomical terms.

Folksonomy is the uncontrolled usage terms/phrases. Folksonomies are thus created by the people for the people on the basis of the premise that people can create a categorization that will better reflect the people's conceptual model. Keeping in mind this concept, we may consider those terms given in example above the proposed model of technical folksonomy. eg.-

Bibliographic Control

UC	Bibliographic record control
BC	Library and Information Science
NC	Metadata Database
RC	Authority Control
R	Use for Library Science

9.3 Conclusion:-

Finally the study clearly shows that trend of usage of folksonomy are increasing day by day in LIS domain. This tiny research study is based on broad concept of folksonomy. We did not find any model folksonomic vocabulary system. Thus the hypothesis we set become positive.

With the invention of internet, the style representation and publication nascent micro thought have been changed remarkable in various subjects. Library and Information Science is no exception. Last few year years it has been found that to represent the thought content the authors of Library and Information Science are using more and more phrases , natural languages than the standard control vocabulary and thus developed a new style and terms popularly known as folksonomy to represent the thought content of subjects. So we should accept and

welcome Folksonomical tagging system for creating metadata system and to create a proposed model which may include any standard tools. And information retrieval process for multi- indexing and also adding future social OPAC (SOPAC) system for library and information centre.

Reference:-

1. Abel F, Leveraging search and content exploration by exploiting context in folksonomy systems, *New Review of Hypermedia and multimedia*, 16 (2) (2010)5-9.
2. An M G, Video summarisation based on collaborative temporal tags, *Online Information Review*. 23(2) (2011) 303-325.
3. Bateman, S Collaborative tagging : Folksonomy ,metadata, visualization, E-Learning , Thesis. University of Saskatchewan .
4. Brathen, J. (2009). An analysis of image folksonomy generation.
5. Chan, K. Y. (2009). Linking folksonomy to Library of Congress subject headings:an exploratory study. *Journal of Documentation* .
6. Francisco Echarte, J. J. (2007). *Ontology of Folksonomy: A New Modeling Method*.
7. Ikki Ohmukai, M. H. (2004). *A Proposal of Community-based Folksonomy with RDF Metadata*.
8. Karch, M. (2006) *Folksonomy based tag recommendation*. *Journal of library Metadata*.
9. Knerr, T. (2005). *Tagging Ontology – Towards a Common Ontology for Folksonomies*.
10. Martin Szomszor, H. A. (2008). *Semantic Modelling of User Interests Based on Cross-Folksonomy Analysis*. SIWS .
11. Mas, M. D. (2010). *Intelligent interface architectures for folksonomy driven structure network*.
12. Mathes, A. (2004). *Folksonomies - Cooperative Classification and Communication Through Shared Metadata*. . *Computer Mediated Communication* .
13. Noruzi, A. (2006). *Folksonomies: (Un)Controlled Vocabulary? Knowledge Organization* , 119-133p.
14. Russell, T. (2005). *Contextual Authority Tagging:Cognitive Authority through Folksonomy*. SLIS .
15. Vig, J. (2010). *Intelligent Tagging Interfaces: Beyond Folksonomy*. TIST .
16. Wu, Z. (2002). *Tagging:Marrying folksonomy and ontology*. Beijing Broadcasting Institute .