

## **Social Tags of Selected Literature of William Shakespeare: A Comparative Study between Library Things Tags and OCLC Fast Subject Heading Lists**

**Dr. Ziaur Rahman<sup>1</sup> and Sanjukta Dutta<sup>2</sup>**

<sup>1</sup>Assistant Professor, Department of Library and Information Science, RabindraBharati University, 56A, BT Road, Kolkata, -700050. West Bengal, India

E-mail- [ziaur.rahman@rbu.ac.in](mailto:ziaur.rahman@rbu.ac.in)

ORCID ID  <https://orcid.org/0000-0002-3393-8690>

<sup>2</sup>Librarian, Morning Bells Academy, 301/311, East Ghospara Road, Shyamnagar, Kolkata, 743 127 West Bengal, India, Email: [sanjuktadutta333@gmail.com](mailto:sanjuktadutta333@gmail.com)

### **Abstract**

Information has become the most powerful tools which keeps us updated and is necessary for modern existence. To cope up with the need of the information seekers, the process of searching, creation and location as well as dissemination of information is also evolving. On one hand flexible searching is preferred by the users and on the other hand controlled vocabulary is universally accepted standard. Tags constitute the core of the text with those terms which are usually not obtainable from controlled vocabulary. This study examines the similarities and dissimilarities between Library Thing tags and OCLC FAST subject heading list. For collecting the data, some of the popular literatures of William Shakespeare were selected. The analysis shows the numbers of tags in Library Things are more in comparison to the OCLC Fast subject headings and 92 percent of the tags were found relevant. The study also shows that 'William Shakespeare' is the focal point for all the data gathered, but on the other hand 'Shakespeare' is not a preferred term in OCLC FAST subject heading.

**Keywords:** Tagging, Folksonomies, William Shakespeare, Social Tags, OCLC FAST, LibraryThing, Controlled Vocabulary, Information Retrieval, User Generated tags

## Introduction

In today's emerging knowledge society, information is being recognized as a primary source of wealth production. To cope up the evolving needs of the users the information retrieval process is also growing and evolving. As the whole generation now depends on online platforms, there social tags come to a play. Now a day's social tagging on online platforms has become a trend. On the other hand Subject Heading list which completely depends on Controlled Vocabulary is universally accepted. A comparative study has been done between Library Thing tags and OCLC FAST Subject Heading list. This study examines the similarities and dissimilarities between social tags and expert generated subject heading list. To carry forward this work tags from Library Thing website and subject heading list from OCLC FAST has been retrieved.

The traditional method of locating, identifying and retrieving resources, be it in non-digital environment or in digital environment was always on the top priority list of information professionals. The traditional methods followed by the libraries for the arrangement, classifying and for the easy retrieval of the resources was with the help of 'controlled vocabularies'. A vocabulary control device depends on a master list of terms that can be assigned to documents. Such a master list of terms is called "List of Subject Headings". A list of subject heading list contains the subject access terms (preferred terms) to be used in the cataloguing or indexing operation at hand. (Controlled Vocabulary, 2022)

FAST (Faceted Application of Subject Terminology) is an enumerative, faceted subject heading schema derived from the Library of Congress Subject Headings (LCSH). The purpose of adapting the LCSH with a simplified syntax to create FAST is to retain the very rich vocabulary of LCSH while making the schema easier to understand, control, apply, and use.

In contrast with the traditional method i.e., 'Controlled Vocabulary' comes the concept of 'Folksonomy'. **Folksonomies** also known as 'social tagging' are user-defined metadata collections. A folksonomy begins with tagging (Wall, 2005). A folksonomy is a decentralized, social approach to creating metadata for digital resources. It is spontaneous and Internet based information retrieval methodology consisting of collaboratively generated, open-ended labels or tags that categorise contents such as web resources, online photographs, and web links (Vaidya & Harinarayana, 2016). LibraryThing (LT) is a web-based open source software social cataloguing tool that is now used for organization of knowledge by individuals as well as many libraries (Van Tine & Walser, 2008; Bartley, n.d.). It is a social cataloguing web based application for listing cataloguing and sharing the books metadata and it's also a way to view other users' collections as well. LibraryThing can be used by everyone involved with books- individuals, libraries, authors and even publishers.

One of the most famous writers of all time in the English playwright, poet and actor named 'William Shakespeare', also popularly known as "Bard of Avon" (Bard is a poet and Avon is the town he lived in), in his lifetime, had written about a hundred and fifty-four sonnets, a few versus, two long narrative poems,

and thirty-eight plays. It is observed that Shakespeare produced most of his works between the years 1509-1613. Some of his best literary works are considered to be early plays of comedies and histories genre. Until about the year 1600, Shakespeare mainly wrote tragedies like Othello, Macbeth and king Lear which are considered to be some of his finest works. In the last few years of his life, Shakespeare mainly wrote tragic comedies and collaborated with other English playwrights. The literature of William Shakespeare has been selected to pursue this research work. William Shakespeare who is one of the great English playwrights has written many literatures. The study demonstrates the Folksonomy or social tags used by the readers as well as the subject heading list created by the experts to recollect and read works authored by William Shakespeare on the popular social tagging site Library Thing and OCLC FAST Subject Heading list.

### **Literature Review**

Edward T. O'Neill, et al. (2003), in their article entitled "FAST: Faceted Application of Subject Terminology." Published under "Subject Retrieval in a Networked Environment" stated that The Library of Congress Subject Headings schema (LCSH) is the mostly availed and widely accepted subject vocabulary for general application. But due its complicated syntax and rules, its application is limited. To preserve the rich vocabulary of LCSH while building the schema easier to manage, implement, understand and utilize, FAST was adapted. As an upward compatibility was sustained by the schema so, any authentic set of Library of Congress Subject Headings can be transformed to FAST heading. Louise F. Spiteri (2006) in her article entitled "The Use of Folksonomies in Public Library Catalogues" published under "The Serials Librarian" examines how folksonomies ought to be utilized in public library catalogues to allow end-users to arrange personal information spaces, create online communities of interest and supplement existing controlled vocabularies. Peishan Bartley (2009) in the article entitled "Book Tagging on LibraryThing: How, Why, and What Are in the Tags?" published under "Proceedings of the American Society for Information Science and Technology" tried to examined book tagging by exploring LibraryThing (LT) members' purposes for tagging; inquire how tags are used; and juxtaposing member assigned tags with terms in corresponding MARC records. Paul Heymann and Hector Garcia-Molina (2009) in the journal entitled "Contrasting Controlled Vocabulary and Tagging: Do Experts Choose the Right Names to Label the Wrong Things?" Examined and explored social cataloguing sites –tagging system wherein users tags book- offers very rare chance to distinct tags to other information organization system. A comparison has been made between

controlled vocabulary and tags. The finding shows that though many of the keywords designated by tags and LCSH are of a same kind but the utilization of those keywords by the interpreters are dissimilar. Westcott Jezmynne, et. (2009) in their article entitled “LibraryThing for Libraries at Claremont” published under “Library Hi Tech” tried to share the experience of implementing Library Thing for Libraries (LTFL) at the Claremont University Consortium’s libraries. The finding result of this study shows that LTFL proved to be an advantageous overlay of user assigned tags and recommended readings based on tags over the current bibliographic information provided in the traditional library catalog. Caimei Lu, Xiaohua Hu and Jung-ran Park (2010) in their article entitled “User Tags versus Expert-Assigned Subject Terms: A Comparison of LibraryThing Tags and Library of Congress Subject Headings.” Published under “J. Information Science” stated that as a latest approach for creating metadata, social tagging has seized the attention of library and information science researchers. This study makes a comparison between the social tags generated by users from LT website and the subject terms allocated by experts according to the Library of Congress Subject Headings (LCSH) as well as explores the differences and similarities between these two and also discovers the feasibility and barriers of imposing social tagging in library systems. Manish Gupta, Rui Li, Zhijun Yin and Jiawei Han (2011) in the paper entitled “An Overview of Social Tagging and Applications” published under “Social Network Data Analytics” encapsulated various methods employed to study numerous feature of tagging. Sheeji Kathuria (2011) in her article entitled “Content Analysis of Social Tags on Intersectionality for Works on Asian Women: An Exploratory Study of LibraryThing” inspected about the usability of social tags by the Library Thing users. She suggests it as a popular web 2.0 social networking site for cataloguing books and to elucidate works on Asian women in representing themes within the context of intersectionality. April Grey and Christine R. Hurko (2012) in their article entitled “So You Think You’re an Expert: Keyword Searching vs. Controlled Subject Headings” published under “Codex: The Journal of the Louisiana Chapter of the ACRL” examined the variations between keyword searching and controlled subject headings and their issues in conducting research. The students’ search habits in a controlled environment have been highlighted, as they often use Google to look for their queries. A.I.M. Jakaria Rahman (2012) in the article entitled “Social Tagging versus Expert Created Subject Headings” drew a comparison between social tagging and expert created subject heading. In his work, he investigated social tagging practice in science book context. The usefulness of social tag was ascertained as a subsidiary of controlled vocabulary to intensify the use of

library resources. The study exhibited that neither the controlled vocabulary nor the social tagging practice is adequate.

### **Objective of the Study**

1. To find the differences as well as the similarities between the user assigned tags of Library Thing as well as subject headings of OCLC FAST (controlled vocabulary) are explored.
2. The usefulness of both tags and subject headings is also being explored.
3. To explain what extent this user generated tags can be used to enhance subject access.

### **Research Methodology**

A detailed study has been done upon Library Thing tags, Folksonomy, Tags, Subject Heading lists, and Controlled vocabulary. Content analysis method was used and found to be apt for conducting this study. The selected literatures of William Shakespeare have been chosen for pursuing this research work. Some of the famous literatures of Shakespeare like Hamlet, Macbeth, Othello have been selected on the basis of its uses, ranks and ratings that were found from both the websites of this study. The tags were retrieved from Library Thing website and subject headings were collected from OCLC FAST subject headings list.

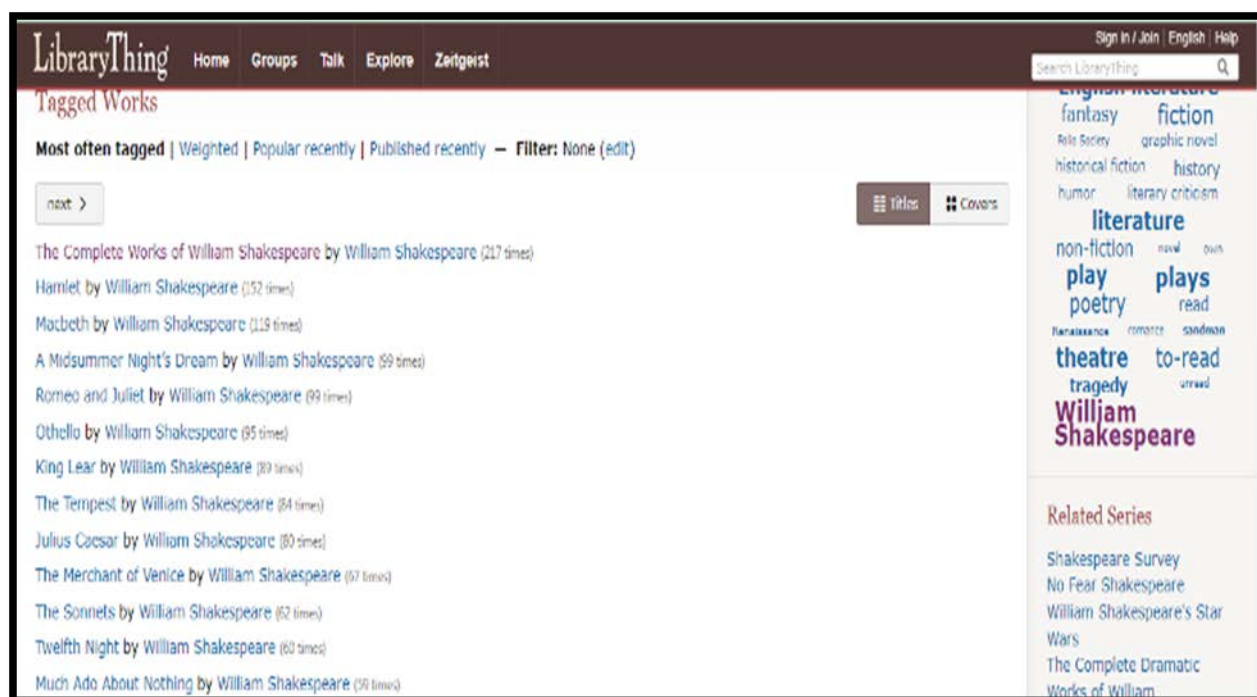
### **Data Collection**

For this study, secondary data such as tags were collected. A sample was needed that would be large enough to provide all the necessary data but also small enough so that the language used in the tags and subject headings could be easier to be studied in details. To accomplish these goals, a comparison has been made between Library Thing tags and OCLC FAST subject headings. To systematize these collected data, web content analysis method was used. For the collection of tags and subject headings, the keyword “William Shakespeare” was chosen.

The criteria followed for the selection of data are mentioned below:

- Books had to be written by William Shakespeare
- Metadata of those books had to be available on both the websites i.e., LibraryThing and OCLC FAST subject headings

**Fig5.1-** showing the number of uses of the literatures in Library Thing website



**Fig5.2-** showing the total number of uses of literatures in OCLC FAST

Heading	Facet	Uses
<a href="#">Shakespeare, William, 1564-1616</a>	person	139874
<a href="#">Hamlet (Shakespeare, William)</a>	title	9900
<a href="#">Macbeth (Shakespeare, William)</a>	title	5331
<a href="#">Romeo and Juliet (Shakespeare, William)</a>	title	4245
<a href="#">King Lear (Shakespeare, William)</a>	title	4059
<a href="#">Othello (Shakespeare, William)</a>	title	3677
<a href="#">Tempest (Shakespeare, William)</a>	title	3143
<a href="#">Merchant of Venice (Shakespeare, William)</a>	title	2861
<a href="#">Midsummer night's dream (Shakespeare, William)</a>	title	2692
<a href="#">Sonnets (Shakespeare, William)</a>	title	2596

**website (OCLC FAST (Faceted Application of Subject Terminology), 2022)**

The above mentioned criteria were fulfilled by the following literatures:-

- Hamlet
- Macbeth
- Romeo & Juliet
- King Lear
- Othello
- The Tempest
- The Merchant of Venice
- A Midsummer Night's Dream

As it is well known that there are so many popular literatures of William Shakespeare. So, it was quite difficult to choose only some of the literatures out of so many. The above mentioned literatures of William Shakespeare was selected on the basis of its 'number of uses', 'reviews', and 'ratings', that was found from both the websites of LibraryThing and OCLC FAST Subject Headings. (Fig.)

A search was carried out for this study. Data were collected from the search box of LibraryThing website and OCLC FAST subject headings, by retrieving the tags with the keyword "William Shakespeare". These data were later used for further data analysis.

### **Data Analysis and Interpretation**

The data collected for analysis was carried out and compared keeping in mind the followings:-

- In-depth analysis of LibraryThing tags
- Comparison between LibraryThing tags and OCLC FAST subject headings

Different categories have been formed for these collected tags as well as these tags have been coded for the data analysis part. Hence, the tags collected from LibraryThing website and OCLC FAST has been broadly categorized under the following groups:-

- **Exact synonyms:** These are the synonyms where there is an exact identical match between LibraryThing tags and OCLC FAST subject headings

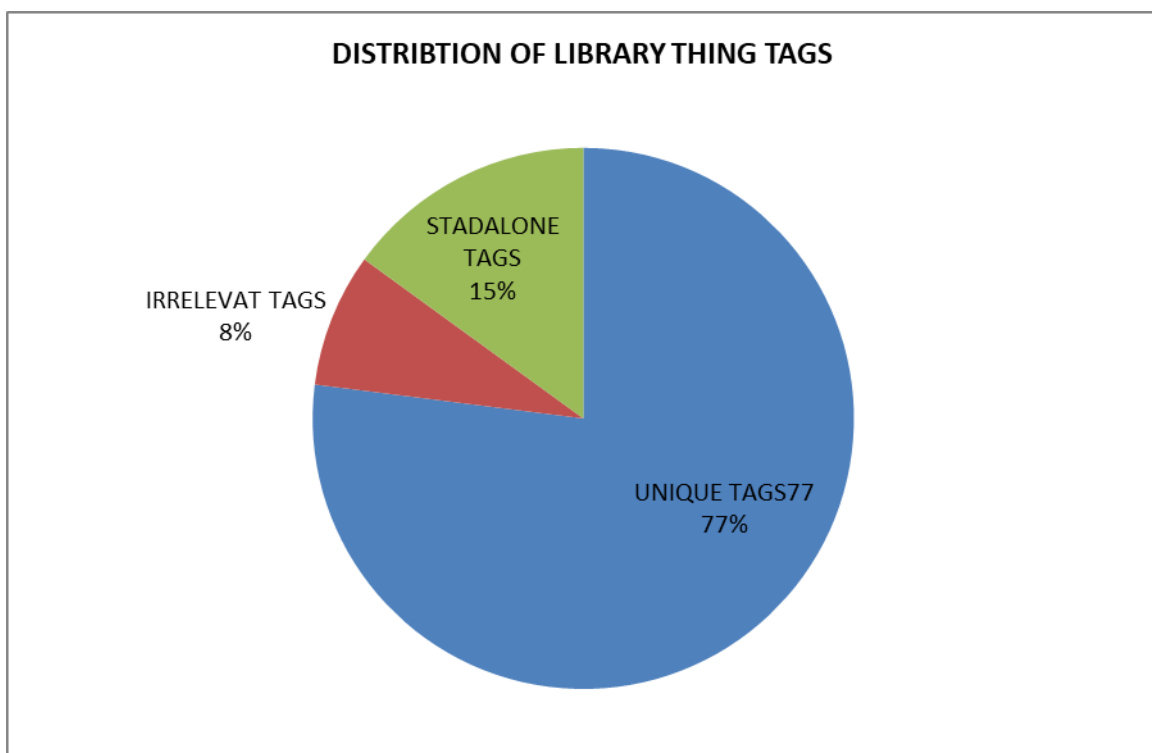
- **Near synonyms:** These are the synonyms where there is a minor variations between the matched terms of LibraryThing tags and OCLC FAST subject headings
- **Broader tags:** These are the tags which indicate broadly the contents in a book
- **Narrower tags:** These are the tags which are much more specific to the contents of the book
- **Stand alone tags:** These are the tags which was found in OCLC FAST subject headings and also in LibraryThing but with slight variations within the titles
- **Unique tags:** These are the tags which are not available in OCLC FAST subject headings
- **Irrelevant tags:** These are the tags present in libraryThing, which do not have any relevance from the access perspective, but were assigned to the resources for personal collection management.
- 

#### **In Depth Analysis of Library Thing Tags:-**

In total 747 tags were collected from the Library Thing website. Out of which 77 percent of the tags were found to be **unique tags**, 8 percent were **irrelevant tags** and 15 percent of the tags were **standalone tags**.



*Fig- showing distribution of LibraryThing Tags*



### Comparison between Library Things Tags and OCLC Fast Subject Headings

By considering the above mentioned terms like Exact synonyms, Near synonyms, Broader tags and Narrower tags, Unique tags, Stand-alone tags and etc., a correlative study has been carried out.

*Table: Representation of total number of tags*

<b><u>TOTAL REPRESENTATION OF TAGS</u></b>	
<b>SUBJECT HEADINGS &amp; TAGS</b>	<b>TOTAL</b>
Total number of LibraryThing tags	747
Total number of OCLC FAST	349
Exact Synonyms	36

Near Synonyms	87
Broader & Narrower tags	18
Unique tags	575
Stand-alone tags	112
Irrelevant tags	59

The above mentioned table exhibits that there were about 747 total number of LibraryThing tags, out of which 687 tags were relevant to the contents of the book, i.e., 92 per cent were found to be relevant tags. It is conjointly noticed that within the relevant tags 36 tags were found to be exact synonyms, 87 were near synonyms, 18 broader & narrower tags, 575 unique tags, 112 stand-alone tags and 59 irrelevant tags.

By using the below mentioned formula, distribution of tags according to their percentage of all book was found.

Category of tag chosen for the study

X

1

0

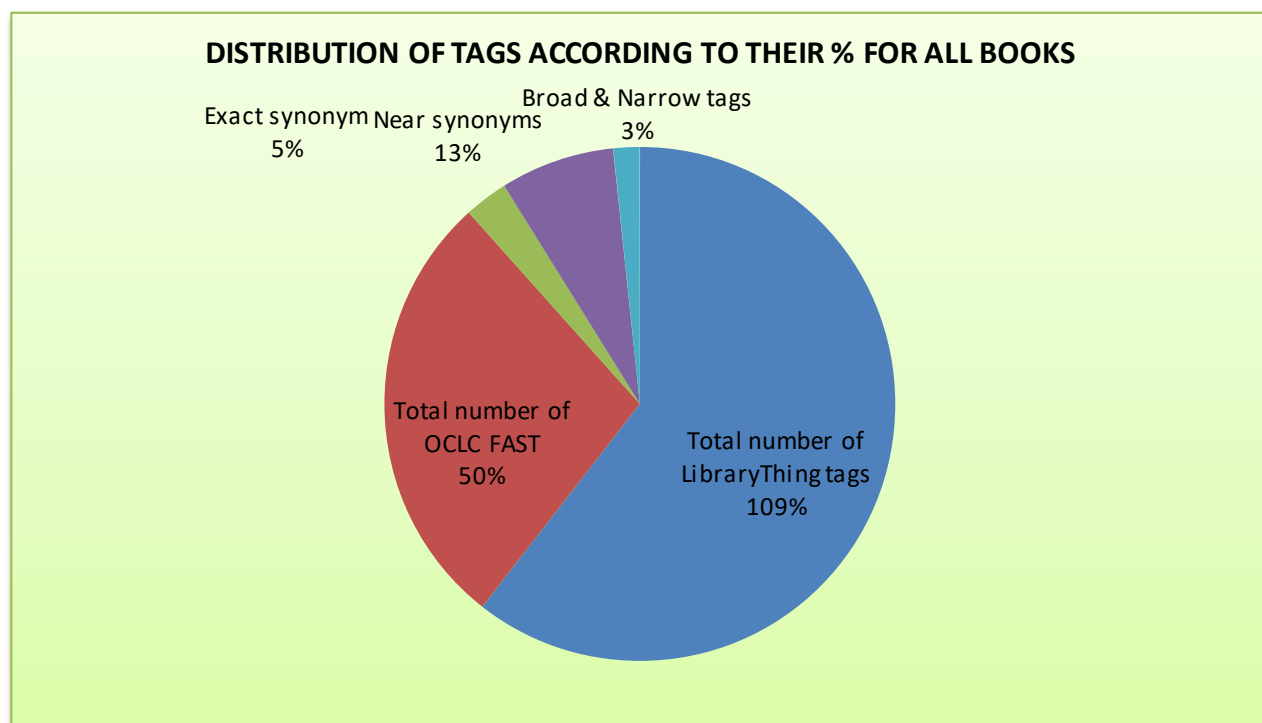
0

Tags found relevant to the content of the book

For this purpose, the earlier mentioned terms like Total number of LibraryThing tags, Total number of OCLC FAST subject headings, Exact synonyms, Near synonyms, Broader tags and Narrower tags were used. Each term was divided by the tag found relevant to the content of the book i.e.687, and the outcome was further multiplied by 100.

**Fig- Distribution of tags according to their percentage for all books**

The pie diagram shown in the figure above displays that the total number of

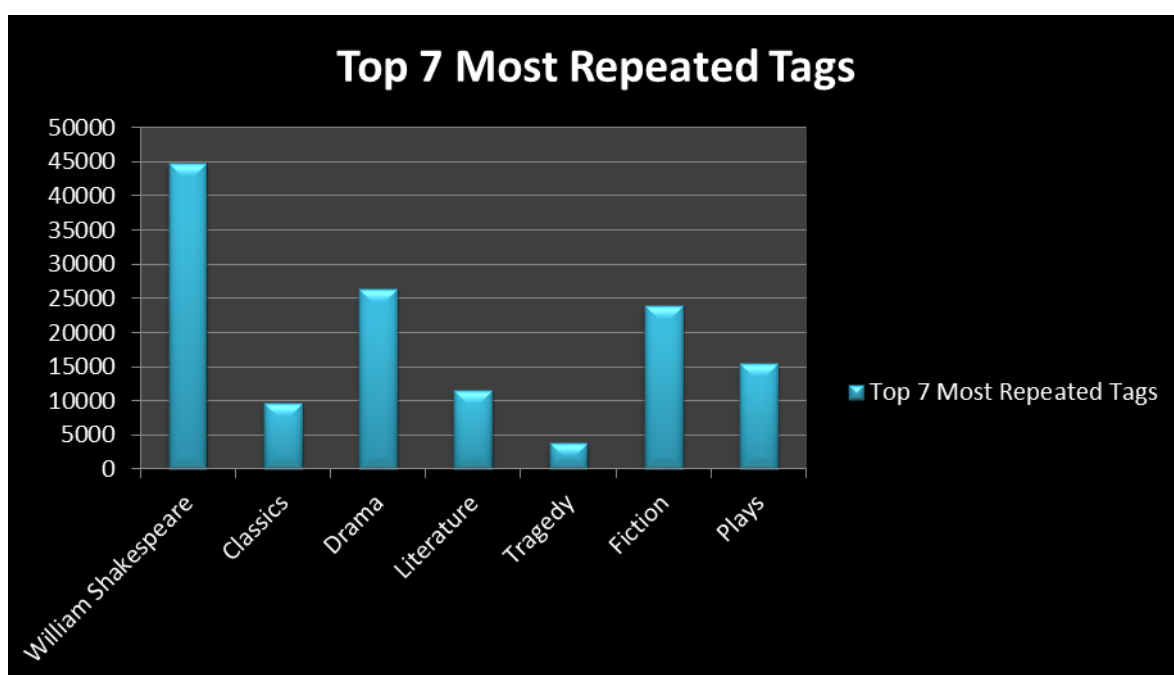


LibraryThing tag is 109 percent, Total number of OCLC FAST subject headings is 50 percent, exact synonym is 5 percent, near synonym is 13 percent and broader & narrower tag is 3 percent.

**Top Seven most repeated Tags in all the eight literature**

A detailed examination was conducted by considering the 7 most repeated tags that has been retrieved from the LibraryThing website. From the figure- given below, it can be seen that the tag William Shakespeare appeared 44803 times, Classics 9680 times, Drama 26420 times, Literature 11470 times, Tragedy 3852 times, Fiction 23867 times and Plays 15513 number of times. But the presence of such tags was not found in the OCLC FAST Subject Headings.

*Fig- showing the top 7 Most Repeated Tags*



## **Conclusion**

This study of comparative study between Library Thing tags and OCLC FAST subject heading list tries to highlight the similarities and differences between social tags generated by users and controlled vocabularies which is expressed in mathematical value and not by semantic meaning of the words chosen. As per the findings, the numbers of tags are more in comparison to the subject headings and 92 percent of the tags were found relevant to the study. While analyzing the 'seven most repeated tags', it is found that 'William Shakespeare' is the focal point for all the data gathered, but on the other hand 'Shakespeare' is not a preferred term in OCLC FAST subject heading. Tags constitute the core of the text with those terms which are usually not obtainable from controlled vocabulary. Tags are richer in text as well as flexible, dynamic, user centric and democratic in nature. On the other hand controlled vocabularies are generally used to maintain precision and to avoid conflict while locating and retrieving required resources. Controlled vocabularies are also used to maintain the standardization. Both tags and controlled vocabularies are complementary to each other. To sum up it can be said that neither tags nor controlled vocabulary is perfect. Although tags are flexible, user centric and dynamic, but they lack precision whereas though controlled vocabularies are précised and universal, they lack flexibility and are rigid for the use of users.

But in this changing society where information explosion is taking place in each and every blink of an eye, as well as the need of the users are also growing with passing days, the methods to satisfy the quest of the users are also developing and taking a new shape. Therefore tagging has become an integral part of the information society in today's date. Due to information explosion new disciplines are also emerging with only with quite a few resources to elucidate them. This is where tagging system will come to the play. So, the tagging system needs to be reinforced and used in a robust manner for providing excellent search and recommendation service in this modern century.

The present study leaves scope for further researches and exploration on various other facets of tags as well as of controlled vocabularies. This research has been conducted by taking into consideration the literatures of William Shakespeare. Similar kind of research can be conducted by considering the literatures written by other famous writers or famous personalities like Charles Dickens, Jane Austen, George Orwell, Rabindranath Tagore, R.K.Narayana, Arundhati Roy, Khuswant Singh, Sarojini Naidu, and Kazi Nazrul Islam, who is also known as the Rebel Poet.

## References

1. O'Neill, Edward T., et al. "FAST: Faceted Application of Subject Terminology." *Subject Retrieval in a Networked Environment*, edited by I.C. McIlwaine, DE GRUYTER SAUR, 2003, pp. 140–46. DOI.org (Crossref), <https://doi.org/10.1515/9783110964912.140>.
2. Wall, Thomas Vander. "Vanderwal.net." *Folksonomy Definition and Wikipedia :: Off the Top*, 2 Nov. 2005, <https://www.vanderwal.net/random/entrysel.php?blog=1750>.
3. Spiteri, Louise F. "The Use of Folksonomies in Public Library Catalogues." *The Serials Librarian*, vol. 51, no. 2, 2006, pp. 75–89. [www.academia.edu](http://www.academia.edu), [https://doi.org/doi.org/10.1300/J123v51n02\\_06](https://doi.org/doi.org/10.1300/J123v51n02_06).
4. Heymann, Paul, and Hector Garcia-Molina. *Contrasting Controlled Vocabulary and Tagging: Do Experts Choose the Right Names to Label the Wrong Things?* Stanford InfoLab, 2009, pp. 1–4. [ilpubs.stanford.edu:8090](http://ilpubs.stanford.edu:8090), <http://ilpubs.stanford.edu:8090/955/>.
5. Westcott, Jezmynne, et al. "LibraryThing for Libraries at Claremont." *Library Hi Tech*, edited by Brad Eden, vol. 27, no. 1, Jan. 2009, pp. 78–81, <https://doi.org/10.1108/07378830910942937>.
6. Bartley, Peishan. "Book Tagging on LibraryThing: How, Why, and What Are in the Tags?" *Proceedings of the American Society for Information Science and Technology*, vol. 46, no. 1, 2009, pp. 1–22. Wiley Online Library, <https://doi.org/10.1002/meet.2009.1450460228>.
7. Lu, Caimei, et al. "User Tags versus Expert-Assigned Subject Terms: A Comparison of LibraryThing Tags and Library of Congress Subject Headings." *J. Information Science*, vol. 36, Dec. 2010, pp. 763–79. ResearchGate, <https://doi.org/10.1177/0165551510386173>.
8. Gupta, Manish, et al. "An Overview of Social Tagging and Applications." *Social Network Data Analytics*, edited by Charu C. Aggarwal, Springer US, 2011, pp. 447–97. Springer Link, [https://doi.org/10.1007/978-1-4419-8462-3\\_16](https://doi.org/10.1007/978-1-4419-8462-3_16).

9. Kathuria, Sheetija. "Content Analysis of Social Tags on Intersectionality for Works on Asian Women: An Exploratory Study of LibraryThing." *Undefined*, 2011. [www.semanticscholar.org](http://www.semanticscholar.org), <https://www.semanticscholar.org/paper/Content-Analysis-of-Social-Tags-on-for-Works-on-An-Kathuria/b5436e90010a500e27ca867a3bc8edc8129a3a13>.
10. Grey, April, and Christine R. Hurko. "So You Think You're an Expert: Keyword Searching vs. Controlled Subject Headings." *Codex: The Journal of the Louisiana Chapter of the ACRL*, vol. 1, no. 4, 4, Feb. 2012, pp. 15–26.
11. Rahman, A. I. M. Jakaria. *Social Tagging versus Expert Created Subject Headings*. 1 June 2012. *ResearchGate*, <https://doi.org/10.13140/RG.2.1.2426.0968>.
12. Vaidya, Praveenkumar, and N. S. Harinarayan. "The Role of Social Tags in Web Resource Discovery: An Evaluation of User-Generated Keywords." *Annals of Library and Information Studies*, vol. 63, Jan. 2016, pp. 289–97.
13. *Controlled Vocabulary*. (n.d.). Retrieved June 4, 2022, from <https://www.librarianshipstudies.com/2020/03/controlled-vocabulary.html>
14. *OCLC FAST (Faceted Application of Subject Terminology)*. (n.d.). Retrieved June 4, 2022, from <https://fast.oclc.org/>