

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

Library Philosophy and Practice (e-journal)

Libraries at University of Nebraska-Lincoln

12-12-2021

Integration of Word Cloud and Tag Cloud with ILS OPAC for Enhancing the Folksonomy Based Services

Sukumar Mandal

Department of Library and Information Science, The University of Burdwan,
sukumar.mandal5@gmail.com

Sailendra Malik

Department of Library and Information Science, The University of Burdwan,
sailendra.malik113@gmail.com

Follow this and additional works at: <https://digitalcommons.unl.edu/libphilprac>



Part of the [Library and Information Science Commons](#)

Mandal, Sukumar and Malik, Sailendra, "Integration of Word Cloud and Tag Cloud with ILS OPAC for Enhancing the Folksonomy Based Services" (2021). *Library Philosophy and Practice (e-journal)*. 6733. <https://digitalcommons.unl.edu/libphilprac/6733>

Integration of Word Cloud and Tag Cloud with ILS OPAC for Enhancing the Folksonomy Based Services

Dr Sukumar Mandal

Assistant Professor

Department of Library and Information Science, The University of Burdwan

Email: sukumar.mandal5@gmail.com

&

Sailendra Malik

M.Phil Scholar

Department of Library and Information Science, The University of Burdwan

Email : sailendra.malik113@gmail.com

Abstract

Purpose

The significant purpose of conducting this research is to find out more about folksonomy-based services with the help of word cloud and tag cloud. Broader terms and narrower terms are searching on bibliographic records and linking web repositories. It shows the integration mechanism of the word cloud and tag cloud with ILS OPAC using Koha and HTML scripts. Designing and developing a process for easy retrieving of library resources based on folksonomy enabled services. These new innovative services are very much helpful for library users.

Methodology

This folksonomy-based integrated framework has been developed and designed based on global repositories. Word cloud and tag cloud are possible with the help of HTML scripts and Koha ILS OPAC. Developed a framework for incorporating the HTML script in Koha OPAC main user block based on tools and news options in staff client interfaces. The HTML script has been designed from the word cloud concept in the online repository. This integrated web 2.0 framework is very fruitful to library professionals because it depends on a LAMP architecture. The whole system and services have been developed on Ubuntu operating system.

Findings

Folksonomy-based services can be achieved for the users after proper configuration and adding of these concepts. It is possible to provide the web 2.0 services with the help of word cloud and tag cloud from Koha ILS OPAC. Related terms and links can be accessed using this integrated framework. So, folksonomy-based services have been provided by using these techniques.

Originality

The originality of this study is keyword visualization based on folksonomy services. This integrated framework belongs to web 2.0 concepts. It is possible to generate the word cloud and tag cloud in ILS OPAC in the form of visualization. This is very useful and conducive to library users. So all the libraries are very much attracted to these modern services and strategies. This integrated framework is based on web 2.0, and folksonomy-enabled services in tag and word cloud. Overall, it is possible to integrate and generate these services with OPAC to increase modern information retrieval systems and services.

Keywords: Word cloud, Tag cloud, Floksonomies, ILS-OPAC, Repository Linking, and Koha

Introduction

The automated and digital library system depends on providing current services based on user-centric and usability. These services can be achieved using modern information technology and social media. The concept of Folksonomies is increasing in the web 2.0 environment. Word cloud and tag cloud are essential components in folksonomies for managing electronic resources. Social bookmarking and social tagging are also crucial elements in the said information. These concepts have been designed and developed with the help of open-source strategies and techniques. Now, the modern integrated library system has provided the word cloud and tag cloud facilities. This can be achieved from the Koha ILS OPAC in two ways, such as integration and manually in-built. Word cloud integration is possible in ILS OPAC using HTML code to create OPAC main user block under system preferences. The Koha ILS OPAC provides another concept after adding the tag on title and subjects based on keywords. These keywords are highlighted when searching any document in ILS OPAC. However, Koha software has been selected based on global recommendations and modern web 2.0 features. Most of the parameters have been supported as per the recommendation proposed by the IFLA working group and digital library federation. Programming language has been developed and designed based on HTML and an online word cloud repository. It needs to carefully read the instructions and process to easy integration of word cloud and tag cloud in ILS OPAC for retrieving the broader and narrower information on any keywords.

Review of Literature

This research explored a new tactic to classifying arbitrary literature in the context of the increasing quantity of digital reader data, such as wikis, blogs, and newsgroups (Cosh, Burns & Daniel, 2008). Hopefully, visual tag clouds visibly convey tag text and hierarchy to portray annotated content for future retrieval. Merely some few publications, but at the other extreme, explicitly evaluated and highlighted the aspects of various graphic approaches using design tag clouds (Ma & Ma, 2020). The study's intention sought and seeing which visual extraction platforms were always the smoothest to get and the most incredible accuracy. Nonetheless, for seven of the ten assignments, respondents of the ontology approach would suggest an effective aggregate. Still, even so, users preferred textbox-based search between cloud-based or annotation discovery (Bar Ilan et al, 2012). This study compared and assessed that each state's new interface, abilities, and behaviors stack up against each other. deciding to see where the web apps are technologies enable and empower visitors with elements to allocate, discover, search, and then use tags within their involvement with social tagging websites was the clear objective of the initial study analysis (Shiri, 2009). Dewey Decimal Classification is indeed an accepted theoretical organizing system to boost content identification and reinforcement learning, which may be the goal of this research (Golub, Lykke & Tudhope, 2014). Several computerized archives don't make it possible for similar subject indexing and segmentation. The research study examines strategies to enhance social tagging integrity for good information exploration and retrieving (Matthews et al, 2010). Researchers analyze keyword excavation for word embedding formation in sparsely tagged contexts with only a few personal annotations. Authors develop the label, a folksonomy-based phrase method that elevates phrases used by the public to the list box. This would be reflected in the concordance seen between automated tagging of the experimental objects and the cloud concepts obtained from the items' data. These findings also suggest the approach's stability instead of specific other cloud generating strategies that are insensitive to small datasets (Carmel et al, 2012).

It's just a smartphone application that uses visual representations to show information and images about a destination. This strategy of obtaining facility data deviates from a more popular online map method. PediaCloud invents word clouds with Wikipedia interconnections. Get a controversial word cloud that quantifies the content related to just that term to help focus your metadata inquiry. And you can are using textboxes to explore a wide range of topics. An evaluation of the program's

alternative uses and system integration selections is introduced here. London, a large city with lots of location data, was used to test PediaCloud in 2014. On PediaCloud, you will find local information by chance rather than searching for it (Tessem et al. 2015). With EdWordle, anyone can manipulate word clouds systematically. EdWordle is a particular phrase program that allows anyone to relocate and adjust specific words without maintaining one's environmental components unchanged. So, scientists integrate a robust frame approximation with standalone system Wordle practises to reload the server and build things compositions. Word clouds can now take on new forms thanks to EdWordle's consistently good functionality. Readers could now create narrative clouds in which the word consolidation is deliberately controlled. These results indicate that we can enhance consumer efficiency, user acceptance, and appearance by contrasting your strategy with existing guidelines (Y. Wang *et al.*, 2018).

Visualizations are a prominent way to illustrate word or text data and are increasing in popularity as the massive data century comes. Nowadays, several digital text documents makers enable users to duplicate a sentence or gather text data from the web domain. Furthermore, most modern word cloud producers exclusively support English symbols, restricting access to English-speaking visitors. Although they are not browsers, there is already word embedding compilers for computer languages (like Python and R). This activity aims at designing a GUI program to produce word cloud mappings with simplicity. The Python programming language is covered in depth. Ultimately, the program runs to be used (Jin, 2017). An approach of instantly viewing the news subjects covered by diverse sources online has been suggested in this study. Word clouds assist readers in assessing a news publication's partiality considerably quicker. As a consequence of all these evaluations and others who have been performed on the techniques being used to retrieve the powerful words contained by a news site (Porter stemmer, for example), authors can confidently state that they are helpful. Again, to analyze the approaches, the harvest time and the frequency of correctly recognized terms were also examined. Due to the appropriate mix with both efficiency and agility, tf-idf is a better rake and text ranking methodology (Chandrapaul et al, 2019).

Word clouds it has received significant interest for condensing textual information. They overlay the most frequent words of a page. Furthermore, a handful of word embedding graphics enables matching numerous sources. The multiple methods offer identical content outlines, although missing structure, readability, and white space. Researchers suggest MultiCloud, a technique for visualizing miscellaneous articles within single phrase clouds, to address this problem. These consider users to explore visualisations by making adjustments and visual images. Users can customize factors to optimize utilization and then get a graphical depiction that facilitates easy word impact on different documents. Graphic researchers and five humanities professors analyzed MultiCloud (John et al, 2018). Textual data is fragmented primarily and massive, proving challenge is to understand caused by a lack of recognized approaches. Consequently, this study was conducted to see if the significant data word embedding approach, one among text algorithms, could be exploited (Lee, 2020). From distributed programming towards virtualization, distributed systems, and so now cloud technology, Technology has revolutionized data centers. The terminology "cloud" is an analogy for the Online platform as a which was before cloud computing provider. Software packages, apps, storage, information, and computational devices all seem to be accessible on the Internet for nothing (Sadiku, Musa & Momoh, 2014). Tag clouds seem popular recently for presenting a concise summary of a website's or document's topic. The tree cloud is a revolutionary representation that offers additional knowledge. Like a word cloud, it highlights one of the most frequent words in the content; however, the keywords are structured on a tree to emphasize their meaningful closeness. Animated tree clouds could be used as analytics tools and discover important things. They still provide techniques for analyzing the tree cloud's accuracy and essential construction activities (Gambette & Véronis, 2010).

Word cloud and tag cloud concepts are belongs to cloud computing in web 2.0 subject area. From the above discussions it is clear that the authors and researchers have discussed both theoretical and practical concepts on word cloud and tag cloud. Now important areas mentioned by the authors these are tag based image retrieval, folksonomy enabled term extraction, analysis of news, clouds of contents, tree cloud display, social tagging of DDC, word cloud wostware Python, MultiCloud techniques, unstructured text data, comparative study of tag clouds, knowledge organization system through word cloud, opportunities and challenges, features of social tagging, locative information, editing of word cloud using EdWordle. These concepts and techniques are discossed by the authors and researchers for developing the word cloud and tag cloud in cloud computing environment. But finds gaps in the areas of integrated library management system. These research paper shows the integration process of word cloud and tag cloud with ILS OPAC using Koha open source software and scripts have been designed and collected from word cloud online environment which clearly discussed in methodology section. This integrated framework provides the folksonomy enabled services towards fulfilling the intrinsic features of web 2.0.

Objectives

The important and essential objectives of this study are pointed as below:

- (i) To explore the integration process and mechanism for word cloud and tag cloud in ILS OPAC;
- (ii) To study the online web repositories on word cloud and tag cloud for easy formation of HTML code;
- (iii) To configure the global system preferences in Koha for retrieving these facilities among the users;
- (iv) To provide the multiple languages from the word cloud and tag cloud in Koha OPAC.

Integration Methodology

Word cloud belongs to folksonomy in modern information technology. It is also known as a tag cloud. It fully supports the word visualization. This is the novelty of the word cloud or tag cloud concept. However, it needs to integrate with ILS OPAC to easily access keywords and linking resources available in a web repository. Now, follow some important instructions and steps are (i) study the word cloud tools and techniques for easy integration with Koha OPAC; (ii) go to the word cloud website <https://www.wordclouds.com/> based on intrinsic features and facilities which consists of weight, word, color, URL, export CSV and import CSV, file, size, shape, word list, Gap size, theme, colors, font, and wizard—then required to configure and add the words from different web repositories like Koha, Google Books, Amazon Books, etc. It is possible to open the website from the word cloud with the help of the link option. Data entry is likely based on multilingual English, Bengali, Hindi, Nepali, Odia (Oriya), Malayalam, Arabic, Chinese, etc. So, regional languages are managed with the help of Unicode and UTF-8. After creating the word cloud, it needs to save in HTML format by applying the save image option, which consists of different file formats such as jpg, png, pdf, and HTML. Finally, click on the save option, and it will store under the download directory, e.g., file name as wordcloud.html. It is possible to export and import in CSV format for easy data conversion from one system to another. Then configure and add the HTML word cloud code in Koha OpacMainuserBlock under the global system preferences by applying the following URL as <http://localhost:8001> and after a click on the save option for easy integration of this concepts. Finally, the results of word cloud will appear in Koha OPAC through the URL of <http://localhost:8002>. This is very user-friendly and informative. All the word cloud can easily be

accessed from this integrated framework. Users are very much benefitted by using this technique both in the offline and online environment. Apart from this, Koha also provides the extra facilities named tag cloud. It is achieved by adding the keywords under the tag option for each bibliographic record in any item. So, folksonomy-based services can be provided from Koha OPAC with the help of the tag cloud tab.

Word Cloud and Tag Cloud in ILS OPAC

The Word cloud interface is shown in Figure-1 towards achieving the folksonomy enabled services from the Koha ILS OPAC to enhance the advanced level tasks and activities. It can perform in two ways, local and global databases, with the help of linking options based on the methodology section from the word cloud web repository. It is possible to access the other repositories regarding the keywords by click on the word, and it will retrieve the correct information from the local and global level repositories. Multilingual visibility is possible in the word cloud interface because it fully supports the UTF-8 standard. So, this interface is conducive to all the users for easy management of regional languages.



Figure-1: Folksonomy services in ILS OPAC

Conclusion

Folksonomy-enabled services are achieved from ILS OPAC in Koha. Word cloud and tag cloud search systems are possible in Koha ILS OPAC. This integrated framework is very conducive and attractive to library users. It is possible to generate the web 2.0 features with the help of word visualization in the library OPAC. All the library professionals and library users have been benefitted by using this integrated folksonomy web 2.0 framework. It can enhance the advanced

level digital services for the users. It is possible to access multiple languages by clicking on any word in the above Figure to increase library users' web-enabled services. Related resources can easily access from the other web repositories with the help of word visualization in the library OPAC. So OPAC and OPAC 2.0 facilities are also achieved and consumed from this integrated domain-specific system and services. The bibliographic and authority data fields search system is enabled in word and tag cloud services on particular components such as author, title, subject, keywords, uncontrolled terms, link, etc. These are very new and innovative digital services provided for library users.

References

- BarIlan, J., Zhitomirsky Geffet, M., Miller, Y. and Shoham, S. (2012). Tag-based retrieval of images through different interfaces: a user study. *Online Information Review*, 36 (5), 739-757. <https://doi.org/10.1108/14684521211276019> (accessed 16 August 2021)
- Carmel, D., Uziel, E., Guy, I., Mass, Y., & Roitman, H. (2012). Folksonomy-based term extraction for word cloud generation. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 3 (4), 1-20. <https://doi.org/10.1145/2337542.2337545> (accessed 18 August 2021)
- Chandrapaul, Soni R., Sharma S., Fagna H., Mittal S. (2019). News Analysis Using Word Cloud. In: Rawat B., Trivedi A., Manhas S., Karwal V. (eds) *Advances in Signal Processing and Communication. Lecture Notes in Electrical Engineering*, Vol. 526. Springer, Singapore. https://doi.org/10.1007/978-981-13-2553-3_6 (accessed 21 August 2021)
- Cosh, K.J., Burns, R. and Daniel, T. (2008). Content clouds: classifying content in Web 2.0. *Library Review*, 57 (9), 722-729. <https://doi.org/10.1108/00242530810911824> (accessed 23 August 2021)
- Gambette, P., & Véronis, J. (2010). Visualising a text with a tree cloud. In *Classification as a Tool for Research* (pp. 561-569). Springer, Berlin, Heidelberg. Doi: 10.1007/978-3-642-10745-0_61 (accessed 25 August 2021)
- Golub, K., Lykke, M. and Tudhope, D. (2014). Enhancing social tagging with automated keywords from the Dewey Decimal Classification. *Journal of Documentation*, 70 (5), 801-828. <https://doi.org/10.1108/JD-05-2013-0056> (accessed 27 August 2021)
- Jin, Y. (2017). Development of word cloud generator software based on python. *Procedia engineering*, 174, 788-792. <https://doi.org/10.1016/j.proeng.2017.01.223> (accessed 29 August 2021)
- John, M., Marbach, E., Lohmann, S., Heimerl, F., & Ertl, T. (2018). MultiCloud: Interactive word cloud visualization for the analysis of multiple texts. In *Proceedings of the 44th Graphics Interface Conference* (pp. 34-41). <https://doi.org/10.20380/GI2018.06> (accessed 15 September 2021)
- Lee, W. J. (2020). A Study on Word Cloud Techniques for Analysis of Unstructured Text Data. *The Journal of the Convergence on Culture Technology*, 6 (4), 715-720. <https://doi.org/10.17703/JCCT.2020.6.4.715> (accessed 18 September 2021)
- Ma, X. and Ma, H. (2020). Comparative study of graphic-based tag clouds: theory and experimental evaluation for information search. *Online Information Review*, 44 (5), 1135-1160. <https://doi.org/10.1108/OIR-12-2019-0372> (accessed 10 September 2021)
- Matthews, B., Jones, C., Puzoń, B., Moon, J., Tudhope, D., Golub, K. and Lykke Nielsen, M. (2010). An evaluation of enhancing social tagging with a knowledge organization system. *Aslib Proceedings*, 62 (4/5), 447-465. <https://doi.org/10.1108/00012531011074690> (accessed 16 September 2021)

- Sadiku, M. N., Musa, S. M., & Momoh, O. D. (2014). Cloud computing: opportunities and challenges. *IEEE potentials*, 33 (1), 34-36. DOI: [10.1109/MPOT.2013.2279684](https://doi.org/10.1109/MPOT.2013.2279684) (accessed 6 October 2021)
- Shiri, A. (2009). An examination of social tagging interface features and functionalities: An analytical comparison. *Online Information Review*, 33 (5), 901-919. <https://doi.org/10.1108/14684520911001909> (accessed 4 October 2021)
- Tessem, B., Bjørnstad, S., Chen, W. & Nyre, L. (2015). Word cloud visualisation of locative information. *Journal of Location Based Services*, 9 (4), 254-272, DOI: 10.1080/17489725.2015.1118566 (accessed 5 October 2021)
- Wang, Y., Chu, X., Bao, C., Zhu, L., Deussen, O., Chen, B., Sedlmair, M. (2018). EdWordle: Consistency-Preserving Word Cloud Editing. In *IEEE Transactions on Visualization and Computer Graphics*, 24 (1), 647-656, doi: 10.1109/TVCG.2017.2745859. (accessed 7 October 2021)

Additional Readings

- Bakhshayesh, S.M., Ahmadi, A. and Mohebi, A. (2021). A new subject-based retrieval and search result visualization approach for scientific digital libraries. *The Electronic Library*, ahead-of-print No. ahead-of-print. <https://doi.org/10.1108/EL-08-2020-0243> (accessed 24 September 2021)
- Ma, X., Xue, P., Zhang, S., Matta, N., Qin, C., Cahier, J.-P. and Wang, K. (2019). Information search by applying VDL-based iconic tags: an experimental study. *Journal of Documentation*, 75 (4), 807-822. <https://doi.org/10.1108/JD-08-2018-0127> (accessed 3 October 2021)
- Radoiu, D., Enachescu, C. and Adjei, O. (2006). A systematic approach to scientific visualization. *Engineering Computations*, 23 (8), 898-906. <https://doi.org/10.1108/02644400610707784> (accessed 5 October 2021)
- Qin, C., Liu, Y., Mou, J. and Chen, J. (2019). User adoption of a hybrid social tagging approach in an online knowledge community. *Aslib Journal of Information Management*, 71 (2), 155-175. <https://doi.org/10.1108/AJIM-09-2018-0212> (accessed 7 October 2021)