MEASURING THE QUALITY OF AN E-GOVERNMENT FOLKSONOMY

HUDA. R. FARHAN AND MARK SANDERSON
Department of Information Studies University of Sheffield, UK
Sheffield, UK
H.farhan@shef.ac.uk *m.sanderson@shef.ac.uk*

Abstract

Evaluating information organization tools is crucial in determining the expected success or failure of the tool prior to implementation. This study used a previously developed evaluation criteria to assess the quality of a folksonomy as an information organization tool in an e-government context, as well as test the validity and reliability of the developed criteria. Two studies were conducted where three English oriented folksonomies were extracted from Delicious and analyzed. The evaluation criteria were found to be reliable, and folksonomy's creation was found to be not dependant on subject domain.

Keywords - Quality of folksonomy, E-government, Evaluation Criteria, research projects.

INTRODUCTION

We assume that in conjunction with conventional information organization tools such as classification schemes and controlled vocabularies, folksonomies can enhance information organization and retrieval within e-government portals. In order to better understand the potential usage of folksonomies in e-government context, an e-government related folksonomy needs to be identified and evaluated as an information organization tool. However, there is a lack of studies addressing the evaluation of a folksonomy's quality as an information organization tool. Therefore, this necessitates the exploration of folksonomy and the way it is developed and how it can be applied; an evaluation study must be conducted to assess the quality of folksonomy to support our call of folksonomy inclusion in official contexts such as e-government. The main aim of this study is to evaluate the quality of folksonomies as information organization and retrieval tool in e-government context using a previously developed evaluation criteria. In addition, the purpose of this study is to test the validity and reliability of the developed criteria and to explore the characteristics of folksonomy's tag formulation.

Consequently, two evaluation studies were conducted to test the validity and reliability of the evaluation criteria. In the first test, an e-government folksonomy was extracted from delicious and analyzed based on the criteria. From the first study, the criteria were modified and re-applied in a second study where two folksonomies: an e-government and a generic folksonomy were compared.

This paper is organized into five sections starting with general background of e-government concept followed by introducing the concept of folksonomy. After that, the first study was described followed by the data collection process, findings and discussion. The second study was introduced in section 4.1 Section 4.2 addressed the general observations, followed by the study conclusion and researchers implications in section 5.

1.1 BACKGROUND

E-government "refers to the process and structures pertinent to electronic delivering of government services to the public" Kasubiene (2007:68). E-government, the new way of governing gained interest by different countries. The number of UN members' countries that facilitate citizen's electronic participation in e-government rose from179 in 2005 to189 in 2008 (United Nation, 2008). e-government is service based context, where information are changing rapidly and the success of those services is based on the ability to reach all citizens who vary in their educational, social and technical background. Also, how services are used is a crucial issue in e-government which might be hindered by lack of information, or difficulty to locate information by citizens. Citizens need to find information and execute services without the need to navigate through unneeded information (Wagner et al., 2006). The need for information system that can cope with this nature and facilitate citizens' access of information regardless of their information skills or social background is urgent (Prokopiadou et al., 2004). E-government initiatives need to assure citizens' access to services and information.

1.2 NATURE OF INFORMATION IN E-GOVERNMENT

Understanding the nature of information in an e-government context might facilitate the identification of suitable classification methods that can be utilized to organize e-government information. Hence, improve the effectiveness of information retrieval in the e-government context. Information in an e-government context is large

and is produced from heterogeneous sources (Wagner et al., 2006; Lyer et al., 2006). West (2004) believed that information in e-government is non-hierarchical unlike traditional government. It is difficult to generalize such assumption. Being produced in different departments and from different managerial levels at government institutions, information in e-government context is expected to be hierarchical in nature.

As noticed by Klischewski and Jeenicke (2004) in e-government context, Information is used in three levels: 1- basic domain knowledge which includes information related to life events such as moving, birth and date registration.2- operational information about contacting the government organizations and opening hours. 3- Information about the source itself such as the physical and usage description as well as access restrictions. Information in e-government is changing continuously (Medjahed et al., 2003). Reacting to citizens' needs and changes in political law and regulations (Stojanovic et al., 2004), e-government initiatives may add new services, execute old services, join services or split services. Also, new laws and procedures may occur during the governmental and administrative transactions related to the communication between e-government's users. Information in e-government context is the result of dynamic processes, and under ongoing changes and modifications to reflect the changing nature of resources from which they evolve. These changes and modifications may result in new information.

Looking at the nature of information in e-government context as being large in size, heterogonous, hierarchical, dynamic, and targeting users with noticeable diversity, raise the question of how to organize this type of information to facilitate citizens' access to government's information and services.

1.3 CONVENTIONAL CLASSIFICATION SCHEMES AND E-GOVERNMENT

Utilizing the suitable classification tool will lead to improvement of information retrieval process on websites (Nasir Uddin et al., 2006). Therefore, e-government websites like other websites utilized conventional controlled vocabulary and classification schemes to organize electronic information, and improve information retrieval effectiveness. In addition, different classification schemes were developed especially for e-government web sites representing concepts and services of e-government that cannot be found in other classification tools in such depth of representation such as TAGS Thesaurus of Australian Government subjects, Victoria online thesaurus- Australia, and IPSV- Integrated public sector vocabulary- UK. The Literature didn't show any evaluation process conducted on those projects; as a result it is difficult to determine the success or failure of deploying conventional classification schemes in the organization and retrieval of information on e-government web sites.

2.1 FOLKSONOMIES, CAN THEY PROVIDE THE SOLUTION FOR ORGANIZING INFORMATION IN E-GOVERNMENT?

One of semantic web applications that gained the most emphasis in the literature was Ontology "a hierarchical description of a set of concepts (is-a hierarchy), a set of properties and their relations, and a set of inference rules" (Lu et al. 2002:2). There were number of attempts to deploy ontology in organizing information in e-government context. Unfortunately, none of the ontology's projects in e-government was evaluated. Therefore, didn't report any success or failure, instead showed the different aspects of ontology. Despite the advantages of Ontology, it is laboured intensive and expensive to develop and maintain. To overcome this limitation ontology can be supported by other classification tools that are not financial and labour demanding. In addition to other advantages, folksonomy a Web 2.0 application fall under this description.

With web 2.0 a great shift has occurred in information retrieval, and users started to have more input in the information organization and retrieval process. In addition, it gave users the opportunity to organize their information resources- which can be a variety of blogs, articles, or bookmarks- by adding annotations and descriptive key words. . The set of key words or tags together form a folksonomy a "conceptual structure created by the people" (Hotho et al., 2006) that might be available not only for the person who created them, but also available for all users who are using the social tagging system.

In a service based context like e-government context -the focus of this study- where users' needs considered as an important factor in determining the efficiency of the services provided (Evans and Yen, 2006), it is crucial to rely heavily on professionals' organization of information. Therefore, involving end users in documents' organization through folksonomies provide more access points to those documents and adhere to users' prospectus (Trant, n.d.). According to Paralic and Sabol (2003:6) "retrieval accuracy depends on the quality of documents annotation". Hence, evaluating the quality of folksonomy is necessary before any attempt to implement it in any context. Therefore, this study aims to evaluate the quality of folksonomy as an information organization tool. The focus of this study is not on the tagging behaviour of users rather on the formulation of tags, types of dominant tags and the way in which tags related to documents in an e-government context. Furthermore, previously developed Evaluation criteria will be applied on folksonomies derived from de.ilico.us aiming at testing the validity and reliability of the developed evaluation criteria.

2.2 HOW CAN WE EVALUATE THE QUALITY OF FOLKSONOMIES

If we view folksonomies as classification tools for enhancing information retrieval, the quality of folksonomies will be identified as:

The ability to describe the content of electronic documents leading to the identification of access points recognized by the majority of users and facilitates information organization and retrieval.

Much of the research that dealt with folksonomies didn't address the quality of folksonomies as a classification tool. Thus no evaluation frame works could be identified and exploit. Therefore, an evaluation criteria was developed table (1) consisting of 6 variables each of which address an essential requirement of any prospectus tool of organizing and retrieving information. Findability will not be tested in this study owing to the nature of data extracted from delicious.

Variables	Attributes
Orthography	The compatibility of tags constituting the folksonomy with the standard English (UK/USA), and an acknowledge populist source
Motivation	Taggers motivation as expressed by tags selection
Relevancy	Representation of the aboutness of documents in the folksonomy
Subject Coverage	Exhaustiveness and specificity of the folksonomy in covering the subject domain
Consensus	Frequency of taggers agreement upon individual/ discrete tags constituting the folksonomy
Consistency	The usage of a certain tag's form by different users
Findability	Folksonomy's ability to identify relevant documents

Table: 1 Folksonomy's quality evaluation criteria

3.1 METHODOLOGY AND DATA COLLECTION

Del.icio.us is a social web site was introduced by Joshua Schachter in 2003 (Yanbe et al., 2007). It was seen as the first site to utilize the concept of social tagging. (Cattuto et al. 2007). According to Macgregor and McCulloch (2006:295), del.icio.us is "arguably the most developed and possibly the most collaborative". For the above mentioned reasons and for the popularity of del.icio.us (Basile et al., 2007;) it was chosen for this study.

Deductive quantitative methods were used to measure to what extent the folksonomy under study met the developed evaluation criteria as well as testing the validity of the criteria. The data analyzed in this study were crawled in the first week of September 2007. A web crawler was designed to crawl del.icio.us through its RSS feed capturing recently posted bookmarks, along with details of the users who tagged these bookmarks, tags attached to the bookmarked pages, and the time of the post. In total, the data set was composed of 1,642,459 posts, 1100 distinct homogeneous URLs, and 54,209 distinct tags. After that, Government related URLs (.gov) were automatically extracted from the data set. Then a manual check was performed to assure the correctness of the data. Some URLs might have (.gov) as an extension but were not government URLs. The data can be summarized as the following: 15740 Gov posts, 11 Distinct URL with Gov extension, 1663 Distinct tags attached to URL with "gov" extension noting that the availability of websites as on the 17th of October 2007

3.2 FINDINGS AND DISCUSSION

To validate the developed criteria, the 1532 tags associated with the nine English oriented e-governments related URLs that have been posted 15209 were evaluated. It was expected that the folksonomy would have a portion of noise (Mathes, 2004), which has been treated differently in the literature; such as restricting the evaluated folksonomy with the use of common tags (Noll and Meinel, 2007). Therefore, before the manual analysis process, tags were treated according to the following guidelines:

Compound tags were decomposed manually by replacing the symbols used to join terms by two spaces, and then a manual check was performed to remove function words, or decompose non-separate compound tags formed by joining two or more terms without any space or symbols into single terms (e.g. primarysource, americanmemory).

Following Kipp and Campbell(2006) treatment of tags “only alphabetically identical tags were merged; spelling variations, plurals and tags distinguished only by additional punctuations were all counted as unique tags”

Punctuation such as (&,*_, #, ¿¿) and functional words were considered as unknown tags if used separately.

To consider a compound tag as a known tag, all terms forming the compound tag should be identified by the tool consulted excluding function words and punctuations.

Personal motivated tags will not go through the rest of the evaluation criteria assuming that personal motivated tags will have meaning only to the user who created them. Hence, will not be useful to other users in identifying resources.

Orthography
Tags were checked against the visual thesaurus for orthographic correctness. Tags which couldn’t be found in the visual thesaurus were checked against Wikipedia. Tags that were not found in both consulted tools were marked as unknown tags and removed from the folksonomy. Tags found in either of the consulted tools were labelled as known tags and included in the folksonomy. In general 1077 tags were identified as known tags; 951 found in the visual thesaurus and 126 in Wikipedia.

Only a small portion of the unknown tags were identified within the folksonomy under evaluation (375 out of 1532 tags). The rest were Standard English vocabularies based on the consulted tools. This finding goes in line with Spiteri (2007) who indicated the standard orthography of folksonomy. The unknown tags can be eliminated either by the folksonomy itself via users’ consensus, that will be discussed later in this paper, or via adding spelling checking facilities to the system to assist users when tagging documents. In this case the usefulness of such assisting tools needs to be explored. Furthermore, it was noticed that tags found in Wikipedia were mostly abbreviations and acronyms.

Motivation
5% of the folksonomy was identified as having personal motivated tags, distributed among the different personal motivation attributes as the following: 8 tags under Self Presentation, 18 tags under Task Organizing and 53 tags under Opinion Expression. The majority of personal motivated tags fall under opinion expression attribute. Also, the small number of personal motivated tags (5%) emphasises the social aspect of folksonomy. Thereby, considering folksonomy as mainly a tool to organize information resources for personal incentives (Lee, 2006; Golder and Huberman, 2005) need to be reconsidered. It was not clear if this finding was peculiar to e-government domain or common when dealing with a special domain separately.

Consensus
High consensus on a small number of tags was noticed. Only 17 tags were used more than 100 times. It was clear that users’ consensus was greater when tags were known tags. Consensus on unknown tags was low. The majority of unknown tags were used only once (317 tags out of 376 tags).

Relevancy
Before exploring tag relevancy, a decomposition process was conducted, the purpose of which was to enable duplicate removal and then automatic term recognition. After this process, 1,104 terms remained. There was a noticeable variation in the number of relevant tags among URLs. The highest percentage of relevant tags was 41.9% and the lowest was 4.4%. Relevant tags represented a small portion of the folksonomy 149 out of 1104 (13.5%).

The low relevancy of the folksonomy might not represent the actual folksonomy’s relevancy to the described URLs. Tags pre-analysis treatments such as decomposition might reduce the degree of relevancy of folksonomy. Compounds tags gained relevancy to URLs only as part of the compound tag and not as a single term for example the compound tag astronomysites, astronomy occurred twice at the website while sites didn’t occur at all. Another probable explanation for the folksonomy’s low relevancy might be the use of automatic tools for relevancy judgement. Hence, tags representing the content of the URL but not mentioned at the website can’t be automatically recognized, for example tags like customer, help, and protection attached to consumer direct web page a site dedicated to help in consumer issues. Furthermore, single word tags that were formed by abbreviations and acronyms that might be relevant to the URLs but didn’t appear in the web page in the abbreviated form. Therefore can’t be recognized automatically, for example the tag picture occurred 4 times at the web site while pic didn’t occur at all. Also, plural and singular variations might reduce the degree of relevancy of folksonomy, for example the tag picture occur 4 times at the web site while pictures didn’t appear at all, In this case, term collapsing might be a solution to measure the actual folksonomy’s relevancy. Based on this analysis, it was decided to assess relevancy manually in future studies.

Consistency

To measure consistency of tag usage, we took the known tags used between 10 to 1680 times, the results were 126 tags. It was noticed that 97 out of 126 tags were consistently used. The types of inconsistent use varied between singular and plural, separated and non-separated compound tags, and short and long form tags. Looking at tags that were used in both singular and plural form, it was noticed that 23 tags were used in plural form with high consensus, and 21 tags used in singular form. Also, it was notice that although single word tags (663) numbered more than compound tags (413), the consensus was higher on compound tags. This finding revealed a similarity between users' terms preferences and information professionals' terms preferences that was expressed by Salton and McGill (1983:55) as "In many manual indexing situations where trained experts are involved the use of pre-coordinated compound terms is preferred".

Subject Coverage

The relevant tags were compared against the Integrated Public Sector Vocabulary (IPSV) including the scope notes: 109 out of 149 relevant tags were identified. Some tags were available in more than one class of the controlled vocabulary used, and we have to choose to which class a particular tag belong. It was found to be difficult to determine the degree of specificity of each tag owing to the fact that we were dealing with tags only without looking at the document. Thereby, the degree of subject specificity wasn't measured. In general most of the evaluated tags were used as a non-preferable terms in the ISPV. This finding was expected owing to the difference between users' language and professional language in describing documents. The high percentage of overlap between ISPV and our folksonomy (73%) was unexpected knowing that the folksonomy under evaluation was attached to URL that differ in scope and dedicated to citizens from different countries where ISPV specialized in British Public Sector information and services. Nevertheless, the percentage of subject coverage within relevant tags was considerably high.

3.3 CONCLUSION

In general, the folksonomy under evaluation was formed according to the Standard English orthography, 70 % of tags were known tags. The noise tags were less than what was expected. Another unexpected finding was the small portion of personal motivated tags in the folksonomy. Only 5% of the folksonomy represented personal motivation. Thereby, the public motivation was dominant in the folksonomy formation. Also, 67% of the personal motivated tags fall under the opinion expression category that can be seen as an indirect form of recommendation.

Only 11.7% of tags were used from 10 times to 1680 times, and the high consensus was on known compound tags. The high consensus on a small number of tags emphasized the social authority on tag inclusion and elimination of noise. Although users' consensus considerably might help in eliminating noise from the folksonomy, removing terms without consensus reduces documents' access points. Users' consensus was on a small number of tags that were plural, non-separated compound, and long form tags. Terms' consistency was noticeably high, and 77% of tags consistent in usage. Users tend to agree on the usage of known tags.

The folksonomy under evaluation showed a low degree of relevancy to the URLs described. Only 13.5% of known tags were relevant based on an automatic relevancy judgement process, which encountered various difficulties: in future, human relevancy evaluation is recommended. As for the Subject coverage, 73% of relevant tags were identified in the ISPV showing a high overlap between the two schemes. Also, it was noticed that most of the folksonomy's tags were non-preferable terms, which emphasized the existence of the gap between professionals' language of describing documents and users' language of describing documents.

To evaluate the subject coverage exhaustivity and specificity, both the controlled vocabulary and the folksonomy should be dedicated to the same e-government initiative which couldn't be reached in our study. Therefore, subject coverage exhaustiveness and specificity will not be evaluated on the rest of the data. Also, the visual thesaurus will not be used to assess the orthography. Instead, WordNet will be used owing to its size, popularity and accessibility.

We concluded that the developed evaluation criteria were valid with few difficulties encountered when evaluating a folksonomy. Also, we noticed that folksonomy didn't suffer from a high portion of noise, and personal motivated tags were not dominant. Therefore, the success of folksonomy usage in e-government context is expected. This assumption is based on initial findings. Hence, the analysis of a larger e-government folksonomy will further support or contradict this assumption.

4.1 TESTING THE RELIABILITY OF THE EVALUATION CRITERIA

Building on the analysis described above, a follow up experiment was conducted to measure the quality of folksonomies in different domains. An e-government folksonomy and a generic folksonomy were chosen for this comparison. Folksonomies evaluated in this stage were crawled in the mid of December 2007 to the mid of January 2008. The data collection process resembled that described in section 3.1. Unlike the data set used for the first study, the second study data set was large in size which necessitates sampling. Furthermore, because the sampling was conducted on the posts in general instead of the distinct URLs, it was noticed that the number

of tags per URL was reduced unlike the case with the first study. The data set comprised of 145129 E-gov posts 20076621 Generic posts.

Knowing that del.icio.us is a bookmark system that is not restricted to a particular region, language, or subject domain. It was expected to have data of different languages. Dealing with this type of data will be difficult owing to the time constrains and technical issues. For the purpose of this study, only English oriented URLs were included.

It was noticed that e-government web pages lack consistency in the use of URL address extension. For example, www.parlament.ch The Federal Assembly - The Swiss Parliament; www.moi.gov.kw Kuwait ministry of interior; www.mze.cz/ Ministry of Agriculture - Czech Republic

Therefore, the inclusion of e-government URL was restricted to the URL representing web pages produced by official government institutions in different managerial levels, and the URL has "gov" extension. Therefore, there might be a bias in the data set towards certain number of e-government initiatives. Owing to the difference in size between the two folksonomies, a stratified sampling method was followed. The folksonomies were divided into two strata according to the URLs subject domains (e-government – generic). Based on an automated sampling formula, random samples were selected from each stratum forming a data set of 614 E-Gov-post with 189 distinct URLs and 634 generic post with 563 distinct URL.

After removing the non-English URLs the total number of distinct tags for e-government folksonomy and the generic folksonomy were 346,303 respectively. Looking at the number of tags in each folksonomy, it was noticed that the e-government folksonomy size was larger than the generic folksonomy. The average number of distinct tags per URL to both e-government and generic folksonomy were 1.7 and 0.66 respectively, Furthermore, a slight difference in the tagging process between e-government folksonomy and generic related folksonomy was noticed. The average number of posts per URL was 3.2% for the e-government folksonomy and 1.1% for the generic. These observations showed the intensity of e-government users' participation in the folksonomy development and the provision of larger numbers of new distinct tags compared to the generic folksonomy's users.

Based on lesson learned from the previous evaluation process, some modifications were applied to the evaluation criteria. First, relevancy will be measured based on human judgement instead of calculating tags occurrence within documents. Second, a subject related reference tool must be consulted when measuring folksonomy's subject coverage.

Orthography
The orthography of both folksonomies was evaluated against WordNet 3.0 and Wikipedia. In the e-government folksonomy 281 tags were identified as known tags, the rest were labelled as personal or unknown tags. 257 tags were identified in WordNet and 24 tags were identified in Wikipedia showing the high degree of overlap 91.5% between the e-government folksonomy and WordNet. Tags that were not identified in WordNet and found in Wikipedia were mostly abbreviations, technology terms and compound terms formed with one or more abbreviations as well as new emerging users' terms (populist) (e.g. cyberbullying, web 2.0, gov, Fafsa). Only 2% of the e-government folksonomy was identified as unknown tags forming a small size of noise that can be recognized in the e-government folksonomy.

As for the generic folksonomy, 317 distinct tags were labelled as known tags, the rest of the folksonomy labelled as Unknown and personal tags. 258 were identified in WordNet and 59 were identified in Wikipedia. The overlap between the generic folksonomy and WordNet was about 81.4%. Like the case of e-government folksonomy, tags identified in Wikipedia were mostly abbreviations and technology terms. The size of unknown tags was considerably small 3%.

Motivation
In both e-government and generic folksonomies, the size of personal motivated tags was too small (5%) to analyse

Consensus
Only 6 distinct known tags had users' consensus more than 10 times with 33 times usage as the highest consensus, where 206 distinct known tags were used only once in e-government folksonomy. As for the generic folksonomy also only 6 distinct known tags were used more than 10 times with 16 times usage as the highest consensus. The majority of the distinct known tags (246 tags) were used only once.

Consistency
It was noticed that in both folksonomies there was a high consistency on tags' usage. only two tags were used in different forms; government used 33 times and gov used only 3 times, the second tag used in different forms was

library used 10 times while library: used once. Owing to the small number of tags with high consensus in both folksonomies it was difficult to identify users' preferences of morphological variations.

Relevancy

The degree of relevancy between folksonomy and URLs was judged by the researcher using Subjective relevance assessment (Saracevic, 1996), based on a three level scale; highly relevant, partially relevant, irrelevant. Owing to restrictions of some sites, downloading the entire URLs sample wasn't accomplished, decreasing the size of the sample to 132 URLs for e-government folksonomy, and 366 URLs for generic folksonomy. The remaining government folksonomy consisted of 438 known tags distributed based on relevancy as the following: 319 highly relevant tags, 66 partially relevant tags and 53 irrelevant tags. The generic folksonomy consisted of 395 known tags distributed based on relevancy as the following: 266 high relevant, 81 partially relevant and 48 irrelevant. From the figures above we noticed that the percentage of highly relevant tags were somewhat higher on e-government folksonomy 72.8 % than the generic folksonomy 67%. Which supported our assumption that folksonomy describe the content of the document to which it is attached and that end users apply tags that can be considered as relevant to the documents' content. Furthermore, it was noticed that relevancy between the evaluated folksonomies and URLs increased after changing the relevancy measurement method from simply calculating tags' occurrence within a document to human judgment. Hence, emphasizing the effect of the decomposition process on folksonomy's relevancy to related URLs that was noticed in the first evaluation study.

4.2 GENERAL OBSERVATIONS

Golder and Huberman (2005) identified categories of folksonomy tags as the following:

Identifying What (or Who) it is About.

Identifying What it Is.

Identifying Who Owns It.

Refining Categories.

Identifying Qualities or Characteristics.

Self Reference.

Task Organizing.

From our evaluation in the previous mentioned study, it was noticed that a folksonomy contained information about documents to which it was attached more than the function of the document or the content. Therefore other categories of tags were identified and would be added to Golder and Huberman's categories as the following:

Type of document: reference, video, games, tools, podcast

Nature of usage: free, open source

Technology used to create the document or required to use the document: flash, plug in

Interval of updating the document or availability of the information: daily, temporary

Also, users tagged some information which is not part of the document content such as links and advertisement which supports (Hayman, 2007, Lin et al. 2006) believes. Owing to the size of the evaluated folksonomies, this finding require further investigation to judge how much emphasis users put on other types of information such as links and advertisement when bookmarking a document.

4.3 CONCLUSION

While applying the evaluation criteria, we didn't encounter noticeable difficulties which emphasize the reliability of the developed evaluation criteria. Also, from the findings above and in comparison with the findings of the first evaluation study, similar results were acquired which lead us to argue that the tagging process is not dependant on the subject domain.

Both folksonomies were compatible to a high degree with Standard English orthography according to the consulted tools. The overlap between e-government folksonomy and WordNet was higher than the overlap between generic folksonomy and WordNet. Therefore, known tags in e-government folksonomy were higher than generic folksonomy. Also, High user consensus and tags consistency was noticed in both folksonomies.

The size of noise in e-government folksonomy and non-government folksonomy was small; hence significant difference couldn't be identified. The size of noise was small: 3% compared to what was expected and mentioned in literature (Marlow et al., 2006). The consulted controlled vocabulary tool couldn't recognize most of tags which could be considered as new emerging users' terms (populist). Thereby, controlled vocabularies weakness in matching the growth of users' terms and vocabularies was revealed. For example, terms like Microsoft, ebooks, and web 2.0 which are widely used couldn't be found in the consulted controlled vocabulary.

As for the degree of relevancy between the evaluated folksonomies and the described documents, both were similar with more than 80% of both folksonomies containing relevant tags demonstrating a folksonomy's ability to describe the abouteness of documents. Hence, folksonomy could be used to retrieve needed information.

5.1 IMPLICATIONS FOR RESEARCHERS

This study provides several implications for researchers. First, there is an evident that social tagging systems are used to store and retrieve e-government related documents. Therefore, further studies are needed to explore e-government users' tagging behaviour as a step towards implementing social tagging systems in an e-government context. Second, this study provides evidence that further studies need to be conducted to evaluate folksonomies derived from a particular e-government initiative. Some criterion couldn't be applied on the data derived from a public worldwide used tagging system such as delicious. Third, dealing with folksonomy requires manual treatment that is known to be time and labour demanding. However, relaying on automatic tags' analysis risks the fidelity of the evaluation results. Finally, folksonomies creation is furthermore complicated and lack standardization than mentioned in literature. However, standards and evaluation measures can be developed to evaluate folksonomies.

REFERENCES

1. BASILE, P., et al. (2007) Recommending Smart Tags in a Social Bookmarking System. Avaliable from: http://www.kde.cs.uni-kassel.de/ws/eswc2007/proc/RecommendingSmartTags.pdf [8th September 2007].
2. CATTUTO, C., et al.(2007) Semiotic dynamics and collaborative tagging.Avaliable from: http://www3.isrl.uiuc.edu/~junwang4/langev/localcopy/pdf/cattuto06semioticDynamics.pdf [17th July 2007].
3. EVANS, D. and YEN, D. C. (2006) E-Government: Evolving relationship of citizens and government, domestic, and international development .Government Information Quarterly, 23(2), 207-235.
4. GOLDER, S. and HUBERMAN, B. (2005)The Structure of Collaborative Tagging Systems. Avaliable from: http://www.hpl.hp.com/research/idl/papers/tags/tags.pdf [24 August 2007].
5. HAYMAN, S. (2007) Folksonomies and Tagging: New developments in social bookmarking. Avaliable from: http://www.educationau.edu.au/jahia/webdav/site/myjahiasite/shared/papers/arkhayman.pdf [24 August 2007].
6. HOTHO, A., et al. (2006) Information Retrieval in Folksonomies: Search and Ranking. In: The Semantic Web: Research and Applications. Berlin / Heidelberg, Springer
7. KAŠUBIENĖ, L. and VANAGAS, P. (2007) Assumptions of E-government Services Quality Evaluation. Engineering Economics (5(55)), 68-74.
8. KLISCHEWSKI, R. and JEENICKE, M.(2004)Semantic Web Technologies for Information Management within e-Government Services:Proceedings of the37th Annual Hawaii International Conference on System Sciences (HICSS'04) - Track 5 - Hawaii, USA,IEEE Computer Society
9. KIPP, M. and CAMPBELL, G.(2006)Patterns and Inconsistencies in Collaborative Tagging Systems: An Examination of Tagging Practices:Proceedings of theAmerican Society for Information Science and Technology Annual Conference Austin, Texas (US),American Society for Information Science and Technology
10. LEE, K. J.(2006)What goes around comes around: an analysis of del.icio.us as social space:Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work.Banff, Alberta, Canada,ACM,191-194
11. LIN, X., et al.(2006)Exploring Characteristics of Social Classification:Proceedings of the17th Workshop of the American Society for Information Science and Technology Special Interest Group in Classification Research Austin,Texas,
12. LU, S., et al. (2002) The Semantic Web: opportunities and challenges for next-generation Web applications. International Journal of Information Research,7(4).
13. LYER, L. S., et al. (2006) Knowledge management for government -to-government (G2G) process coordination Electronic Government, 3(1),18-35.
14. MACGREGOR, G. and MCCULLOCH, E. (2006) Collaborative Tagging as a Knowledge Organisation and Resource Discovery Tool. Library Review, 55(5).
15. MARLOW, C., et al.(2006)HT06, tagging paper, taxonomy, Flickr, academic article, to read:Proceedings of the Seventeenth conference on Hypertext and hypermedia.Odense, Denmark ACM,31-40
16. MATHES, A. (2004) Folksonomies - Cooperative Classification and Communication Through Shared Metadata. Avaliable from: http://www.adammathes.com/academic/computer-mediated-

communication/folksonomies.html [17th July 2007].
17. MEDJAHED, B., et al.(2003)Semantic web enabled e-government services:Proceedings of the Annual national conference on Digital government research.Boston, MA,Digital Government Society of North America, 1 - 4
18. NASIR-UDDIN, M., et al. (2006) Information description and discovery method using classification structures in web. Malaysian journal of library and information science, 11(2), 1-20.
19. NOLL, M. and MEINEL, C.(2007)Authors vs. readers: a comparative study of document metadata and content in the www:Proceedings of the2007 ACM symposium on Document engineering.Winnipeg, Manitoba, Canada ACM,177-186
20. PARALIC, J. and SABOL, T. (2003) Work with Knowledge for Support of e-Government. Avaliable from: http://people.tuke.sk/jan.paralic/papers/Znal03_invited.pdf [29th June 2008].
21. PROKOPIADOU, G., et al. (2004) Integrating knowledge management tools for government information Government Information Quarterly, 21(2), 170-198
22. SPITERI, L. F. (2007) Structure and form of folksonomy tags: The road to the public library catalogue Webology, 4(2).
23. STOJANOVIC, L., et al. (2004) On Managing Changes in the Ontology-Based E-government.In:On the Move to Meaningful Internet Systems 2004: CoopIS, DOA, and ODBASE. Berlin / Heidelberg, Springer
24. TRANT, J. (n.d.) Exploring the potintial for social tagging and folksonomy in art museums:proofofconcept.Avaliablefrom: http://www.archimuse.com/papers/steve-nrhm-0605preprint.pdf[24thAugust 2007].
25. United Nations.(2008)UN E-Government Survey 2008: From e-Government to Connected Governance. Department of Economic and Social Affairs. New York
26. WAGNER, C., et al. (2006) Building Semantic Webs for e-government with Wiki technology. Electronic Government, an International Journal, 3(1), 36-55.
27. WEST, D. M. (2004) E-Government and the Transformation of Service Delivery and Citizen Attitudes. Public Administration Review, 64(1), 15-27.
28. YANBE, Y., et al.(2007)Can social bookmarking enhance search in the web?:Proceedings of the7th ACM/IEEE-CS joint conference on Digital libraries Vancouver, BC, Canada ACM,107-116