An Exploratory Study on Semantic Arrangement of VDL-Based Iconic Knowledge Tags†

Xiaoyue Ma* and Jean-Pierre Cahier**

*School of Economics and Management, Xidian University, 266 Xinglong Section of Xifeng Road, Xi'an, Shaanxi, 710126, China, <xiaoyue.ma.utt@gmail.com>
**ICD/Tech-CICO Lab, Université de Technologie de Troyes, BP 2060, 10010 Troyes, France, <cahier@utt.fr>

Xiaoyue Ma is a lecturer in the department of information management of Xidian University since September 2013. She got her Ph.D. degree at the University of Technology of Troyes (France) in network, knowledge and organization. During the study, she developed an icon system for knowledge tagging. Her research interests currently focus on visual knowledge management, knowledge organization and sharing, in the field of which she has published about fifteen academic papers in international journals and conferences.

Primarily an engineer at L'Ecole Centrale de Lyon, France, J.-P. Cahier is since 2005 a searcher in knowledge engineering and cooperative work, in the Tech-CICO Lab (Troyes, France). During his Ph.D. work, he participated in early research on the social semantic web, by which the community can build a dynamic and collective meaning. He built the "Agorae" software tool, to build "hypertopic" knowledge maps. Today he focuses on visual and semiotic new approaches of collective knowledge management.

Ma, Xiaoyue and Cahier, Jean-Pierre. **An Exploratory Study on Semantic Arrangement of VDL-Based Iconic Knowledge Tags.** *Knowledge Organization. 41(1)*, 14-29. 49 references.

Abstract: VDL (Visual Distinctive Language)-based iconic knowledge tags are graphically structured icons for knowledge representation. VDL was developed and assessed to enhance the connection of iconic tags and the connection of tagged knowledge. The purpose of this paper is to present further investigation on an arrangement method for these special tags as well as the characteristics of better tag presentation in knowledge organization systems (KOS). An online experiment was conducted to compare tagging results of four types of iconic tag presentations: two types of iconic tags (VDL-based iconic tags and iconic tags without explicit structure) under two arrangement methods respectively (random arrangement and semantic arrangement). Tagging quality and tagging speed were measured to identify how users locate and locate again appropriate iconic tags for knowledge tagging. A supplementary test on tag structure identification was also carried out for each tag presentation. Semantic arrangement of VDL-based icons helped users to tag given articles with more appropriate tags in less time. Users identified better tag structure in this type of tag presentation. This in-depth work of VDL-based iconic tags is among the first to investigate how to visually structure knowledge tags, a problem neglected by previous studies on icon knowledge representation.

Received 26 May 2013; Revised 11 September 2013; Accepted 23 October 2013

Keywords: tags, tagging, group tagging, VDL-based icons, semantic arrangement

† This work is partly financed by the Fundamental Research Funds for the Central Universities in the project "Evaluation analysis of knowledge organization system," no. K5051306024.

1.0 Introduction

Knowledge Organization Systems (KOS's) (Hodge 2000) is a general term referring to the tools that present the organized interpretation of knowledge structures. It is intended to encompass all types of schemes for organizing knowledge. KOS's includes classification schemes that organize materials at a general level (such as books on a shelf), subject headings that provide more detailed access, and authority files that control variant versions of key information (such as geographic names and personal names). They also include less-traditional schemes, such as semantic networks and ontologies. A structured KOS serves as a bridge between the user's information need and the material in the collection. With it, a user should be able to identify boundary objects of interest (Bowker and Star 1999) without prior knowledge of its existence. Whether through browsing or direct searching, whether through themes on a web page or a site search engine, the structured KOS guides the user through a discovery process.

Knowledge tags (henceforth "tags") are employed to organize, share, and search information in KOS's. These short textual labels can be regarded as the keywords to imply the categorization of knowledge. For example, when an item of knowledge is marked by the tag "bus," it is considered to be sorted into the category "bus," while upper categories such as "transport," or sub-categories such as "minibus," might also be available. Tags of KOS's and their structure work as dynamic knowledge organization access (Kipp and Campbell 2010). Users are able to annotate sharing knowledge in KOS by predefined and recommended tags. New tags proposed by experts and users could be in turn added into certain category for potential searching use.

A tag cloud selects and presents a limited number of tags in a KOS to make a simple presentation of knowledge. It is visual interaction between users and knowledge resources by tagging. Besides the visual features of tag clouds such as size, color or font weight (Bielenberg and Zacher 2005; Shaw 2008; Bateman et al. 2008; Rivadeneira et al. 2007), a lot of previous studies tried also to find out which type of tag arrangement would improve the interaction quality of textual tag clouds (Kerr 2006; Chen et al. 2009; Knautz et al. 2010). Compared to several arrangement approaches, the most acceptable view on this issue was to semantically structure tag clouds (Schrammel et al. 2009). A whole tag cloud could be regarded as the combination of several clusters with the tags in each cluster representing topic-related terms.

In our research we are no longer interested in the tag arrangement of textual tag clouds. However, we need to make use of these empirical results for our new form of tags—VDL-based iconic tags (VDL stands for Visual Distinctive Language) (Figure 1). In former work (Ma and Cahier 2012), VDL-based iconic tags were created and

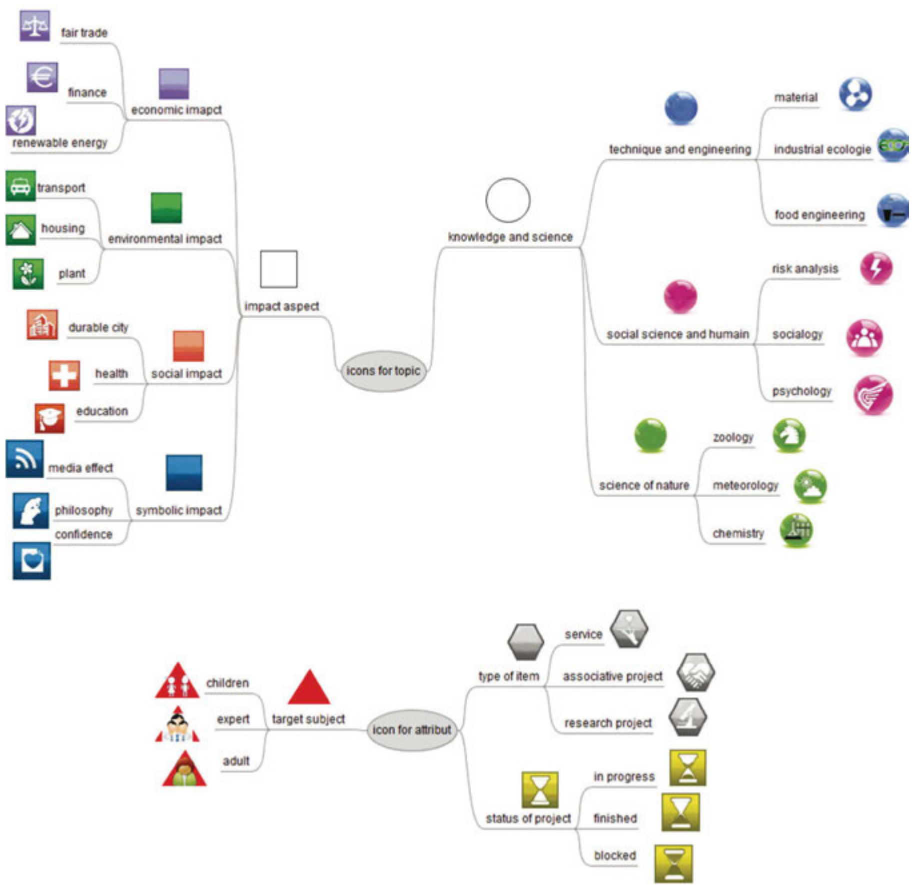


Figure. 1. Examples of VDL-based iconic knowledge tags in the field of sustainability (upper for topics and lower for attributes)

validated to improve the limits of textual tags, like incomprehension of tag meaning and neglect of controlled vocabulary structure of tags (Kipp and Joo 2010). They also contributed to improving visual thesauri (Shiri and Revie 2005), which bears more on social knowledge tags. However, tags for the former test were all randomly arranged; in other words, tags semantically related were not strictly clustered. To complete the conceptual proposition of VDL-based iconic tags, we continue to investigate how to arrange them for knowledge tagging in KOS's. Meanwhile observation on semantically structured textual tag clouds will be also verified whether applicable or not to VDL clouds. The results will be meaningful to the theoretical foundation of iconic tag clouds that can be implemented in KOS's and other tag-concerned systems. It could as well be meaningful to large-scale icon systems where icons are the main knowledge entities instead of a functional part.

In the next section, we will review the state-of-art of the semantic tag relations and semantically structured textual tag clouds. A presentation of previous work on VDL-based iconic tags in section two will specify the context of research and clear the motivation of this deeper study. Then section three will explain what semantic arrangement of VDL-based iconic tag clouds means and our hypothesis. A tagging test will be presented to confirm our hypothesis and discuss the characteristics of a better tags presentation in KOS.

2.0 Background

Before discussing the tag arrangement of VDL-based iconic tags, we need first to look back on the studies about tag arrangement of textual tags: what is defined as the semantic relations among tags and why semantically structured tag clouds have more advantages. In the latter part of the background, more details will also be presented about VDL-based iconic tags and the empirical demonstration of former experiments. All of the information is expected to give complete motivation on the research of semantic tag arrangement of VDL-based iconic tags.

2.1 Semantic relations within tags in KOS

The representation of tag structure (a group of tags in KOS) is as important as that of each single tag. On one hand, an explicit tag structure facilitates finding and finding again later an appropriate tag in a large group of tags. While searching tags for specific knowledge tagging purposes, relations among them allow users to find several alternatives referring to the closed topics. This leads to deeper comparison and selection among them in order to make better tag choices. On the other hand, tag structure offers a possible link between documents tagged by these tags. When tags reach a semantic consensus, knowledge tagged by them may be intuitively considered associated by common topics or attributes. This connection of documents is useful especially when dispersed documents are represented without clear categorization. Clear tag structure enhances the implicit network of tagged knowledge in KOS which provides easier organization and seeking.

In spite of the vocabulary problem existing (Sen et al. 2006; Downey et al. 2008; Macgregor and McCulloch 2006; Ames and Naaman 2007), there has been accumulating evidence suggesting that emergent structures do exist in social tagging systems (Golder and Huberman 2006; Cattuto et al. 2007). Most importantly, these emergent structures do seem to have the potential to help users to explore information by providing meaningful organization and indexing of information resources. Despite the diverse backgrounds and information goals of multiple users, co-occurring tags exhibited hierarchical structures that mirrored shared structures that were "anarchically negotiated" by the users.

To explore the hierarchical relations between tags, an intuitive way is to cluster the tags into hierarchical clusters. Wu et al. (2006b) used a factorized model, namely Latent Semantic Analysis, to group tags into non-hierarchical topics for better recommendation. Brooks and Montanez (2006) argued that performing Hierarchical Agglomerative Clustering (HAC) on tags can improve the collaborative tagging system. Later, HAC was also used for improving personalized recommendation (Shepitsen et al. 2008). Heymann and Garcia-Molina (2006) clustered tags into a tree by a similarity-based greedy tree-growing method. They evaluated the obtained trees empirically, and reported that the method is simple yet powerful for organizing tags with hierarchies. Based on Heymann and Garcia-Molina's work, Schwarzkopf et al. (2007) proposed an approach to modelling for users with the hierarchy of tags. Begelman et al. (2006) used top-down hierarchical clustering, instead of bottom-up HAC, to organize tags, and argued that tag hierarchies improve user experiences in their system. Most of the hierarchical clustering algorithms rely on the symmetric similarity among tags, while the discovered relations are hard to evaluate quantitatively, because one cannot distinguish similar from not-similar with a clear boundary.

People have also worked on bridging social tagging systems and ontologies in the semantic way (Fu et al. 2010). Mika (2005) proposed an extended scheme of social tagging that includes actors, concepts and objects, and used tag co-occurrences to construct ontology from social tags. Wu et al. (2006a) used hierarchical clustering to build ontology from tags that also use similar-to relationships. Later, ontology schemes that fits social tagging system were proposed, such as (Van Damme et al. 2007) and (Echarte et al. 2007), which mainly focused on the relation among tags, objects and users, rather than among tags themselves. Passant (2007) mapped tags to domain ontologies manually to improve information retrieval in social media. To construct a tag ontology automatically, Angeletou et al. (2007) used ontologies built by domain experts to find relations between tags, but observed very low coverage. Specia and Motta (2007) proposed an integrated framework for organizing tags by existing ontologies, but no experiment was performed. Kim et al. (2008) summarized the state-of-the-art methods to model tags with semantic annotations.

The idea of a social-semantic web (Bénel et al. 2009) has emerged over recent years adopting the notion of collaborative knowledge management (Ma and Cahier 2011). Contrary to the semantic web (Berners-Lee, 2000), the social-semantic web is not interested in formal semantics but in semantics depending on human subject and semiotic substrate. The knowledge model Hypertopic (www.hypertopic.org; see Zhou et al. 2006) was developed in the frame of the social-semantic web. It proposes to describe an item through its topics, attributes and resources. For each item, pertinent topics are listed to mention which type of subject is involved. These topics are supposed to be associated with certain viewpoints considered alongside potential users. In other words, the implied viewpoints represent the information goals of various people. Attributes and their corresponding values provide also complementary and objective information that cannot be modified according to different users' viewpoints. They are organized in pairs with the name and its values as a facet (Mas and Marleau 2009). Talking of resources, they characterize other vivid demonstrations of items, such as illustrating photos, URLs of websites or supporting document links.

On one hand, Hypertopic proposes a knowledge categorization method, especially emphasizing the concept of viewpoint, which is significant in collaborative knowledge classification (Ma and Cahier 2011). As illustrated in Figure 2, one item may be associated with more than one topic depending on subjects' viewpoints. Meanwhile the relations between two items can also be changed depending on them. For example, museum 1 and 2 are two items referring to the same topic category "educational place" from the viewpoint "function and value." However they will be categorized into two different topic categories when talking about the viewpoint "style of appearance"—museum 1 in Baroque while 2 in Gothic. Possible sub-topics such as "Baroque in 15th century," or "Baroque in 16th century" are supposed to continue specifying the period in which the style emerged. This type of categorization emphasizing the concept of viewpoint provides more flexible organization of items (knowledge) in a KOS. Categories of items are not solid but dynamic relying on users' opinions. It also allows collaborative participation of categorization from various users to search and retrieve an item under the viewpoints they prefer, even create a totally new viewpoint without changing current knowledge structures.

On the other hand, Hypertopic provides a meaningful structure to manage tags that stem from topics (viewpoints) and attributes. Both topics and attribute values recommend textual tags to specify knowledge categorization. Topic can be regarded as the "special" attribute consider-

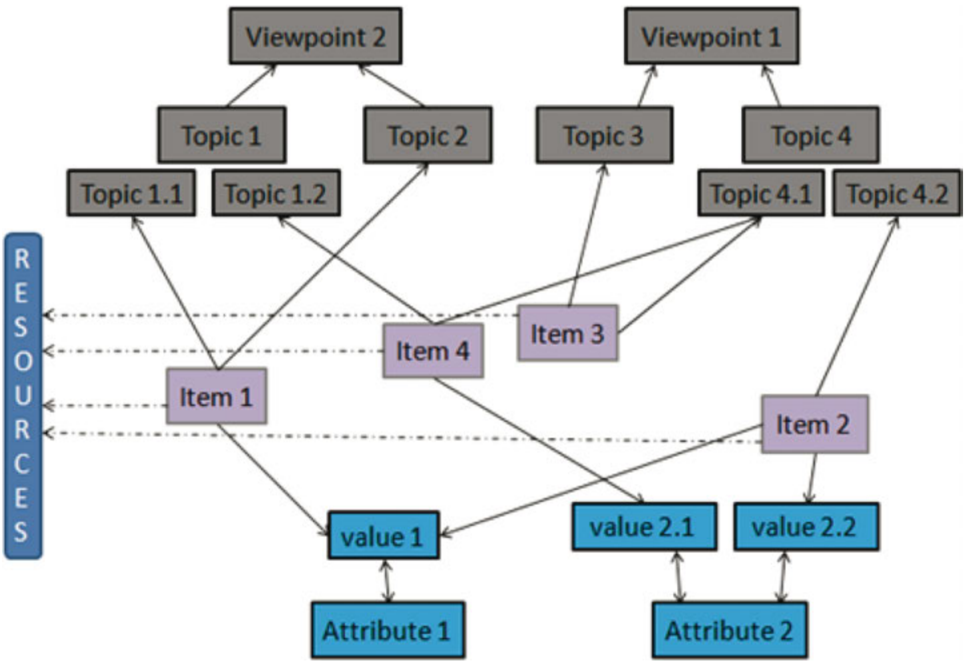


Figure. 2 Knowledge organization based on Hypertopic model: topics (viewpoints), attributes and resources (Ma and Cahier 2012)

ing "topic" as the attribute name. These tags allow Hypertopic-based knowledge tagging in which users are able to tag knowledge through its topics and attributes. For each topic (viewpoint) or attribute value, more than one possible textual expression may exist, in kinds of synonym or languages. We consider them as a unit for a topic tag or an attribute tag, by which the tag structure is clearer and easier to be managed. All the topic tags are cataloged under the tree structure considering the common viewpoint as the "parent" node. Each topic tag may be followed by subtopic tags. The user is allowed to add new categories of topic tags by creating a parent node named "my viewpoint." This convenience encourages collaborative knowledge management (Ma and Cahier 2011) and collective tagging. However, if textual tags generated by Hypertopic are presented together without implying topic category or attribute name, the structure will be less explicit especially when users are not familiar with tag meaning. This problem is increasingly evident when tag numbers grow. In addition, sometimes one topic tag may be related with several topic categories. For example, renewable energy can be sorted in topic "energy" and topic "economy." In this case one textual form expression cannot reflect all possible relevant categories. A more explicit representation is required.

2.2 Semantically structured tag clouds

Although semantic relations do exist within tags, tag arrangement based on semantic clustering was not largely accepted at the beginning. Previous studies considered different types of arrangement to improve better interaction of tag clouds. Halvey and Keane (2007) investigated the effects of different tags clouds and listed arrangements comparing the performance for searching specific items. The setup included random and alphabetically ordered lists and tag clouds. Semantic ordering was not part of tested setups. They found that respondents were able to more easily and quickly find tags in alphabetical orders (both in lists and clouds). Rivadeneira et al. (2007) compared the recognition of single tags in alphabetical, sequential–frequency (most important tag at the left-upper side), spatially packed (arranged with Feinberg's algorithm) and list-frequency layouts (most important tag at the beginning of a vertical list of tags). Results did not show any significant disparity in recognition of tags. However, respondents could better recognize the overall categories presented when confronted with the vertical list of tags ordered by frequency. Hearst and Rosner (2008) discuss the organization of tag clouds. One important disadvantage of tag cloud layouts they mention is that items with similar meaning may lie far apart, and so meaningful associations may be missed.

The following studies started to focus on semantic relations within tags and tried to represent it in textual tag clouds. Hasan-Montero and Herrero-Solana (2006) claimed that the alphabetical arrangements neither facilitate visual scanning nor infer semantic relations between tags. They discovered that the users have difficulty comparing tags with small size and derived semantic relations. There might be wrong relation-interpretation with items placed near to each other. They proposed an algorithm using tag similarity to group and arrange tag clouds. Therefore, they developed a k-means algorithm to group semantic similar tags into different clusters and calculate tag similarity by means of relative co-occurrence between tags. Similar work can be found in (Provost 2008). Likewise, Fujimura et al. (2008) use the cosine similarity of tag feature vectors (terms and their weight generated from a set of tagged documents) to measure tag similarity. Based on this similarity they calculate a tag layout, where distance between tags represents semantic relatedness. Another very similar approach is proposed by (Berlocher 2008).

An empirical evaluation of semantically structured tag clouds (Schrammel et al. 2009) has demonstrated that topical layouts (semantically-structured tag clouds) can improve search performance for specific search tasks compared to random arrangements, but they still perform worse than alphabetic layouts. The semantic arrangement must be good enough otherwise users will not be able to distinguish it from random layouts. Semantic layouts therefore should only be used when the quality of the arrangement can be assured. Test participants also commented that it was difficult to identify clusters and relations beyond single lines.

2.3 Modelling well-structured iconic tags using Visual Distinctive Language

VDL-based iconic tags are well-structured icons working for better representation of tag structures and single tags. Because the semiotic representation of icons has been largely studied already, we are more interested in visualization of tag structures than in choosing symbols for each iconic tag. However the conclusion on imaged information (Paivio 1971) is as well accepted. To visualize the tag structure in KOS's, it has to first confirm the way of organizing tags, and then iconize them as well as their structure. Tags in KOS's can be regarded as the keywords to specify possible knowledge categorization, which means structuring tags is, in fact, recommending a method to organize information and knowledge.

The idea is to benefit from the categorization of textual tags made by Hypertopic (from topics and attribute values) and iconize it for better visualization of separate tags and their structures (see Figure 3) (Ma and Cahier

2012). Here we think of the simplest case, one-tag-one-icon tags, where one recommended tag corresponds to each topic and attribute value. What's more, one iconic symbol represents the current textual tag although various symbols can explain the same tag meaning. However, this approach will be extended equally for the many-to-many case where no constraint of textual tags and icons is applied. For example, the tag "nature" could also coexist with "mode of life," "environment," and other synonymous (or closed expressions) in different languages. This tag "nature" will be represented by iconic symbols of trees, flowers and other possible signs. No matter whether one-to-one or many-to-many cases, the tag structure is always consistent with the knowledge organization according to Hypertopic. As long as tags in given KOS obey this structure, we can iconize them in the same way.

The symbolic characters of icons convey explicitly the represented objects, while graphical characters help visualize relations within them. In particular, a special group of icons called "pre-icons" function to signify the categories of tags in a KOS: the same viewpoint, the same branch of topic or the same attribute name. Pre-icons act as the common base of iconic tags. Tags in each category will be specified by combining symbols with this corresponding iconic base. Nevertheless a pre-icon for attribute name is useless in some cases. For example, when iconizing the attribute values of "language," it is clear enough to represent them independently with national flags.

All the pre-icons in this model can be explained as a "graphical organizer" named Visual Distinctive Language (VDL) (see Figure 4), which aims to visually characterize the categorization made by Hypertopic protocol. Here the "language" is a wide notion (instead of a spoken word) that allows communicating with each other in a relatively effortless way (Nakamura and Zeng-Treitler, 2012). We call it Visual Distinctive Language because it provides visual consensus (pre-icons) on information structure (distinguishing one category from another). Users who accept it could communicate under this visual convention, like knowledge sharing, one of the communication means in KOS.

Among six visual variables illustrated by Bertin's graphical semiotic theory (Bertin 1983), three are in less accordance with the purpose of tag structure representation: size, orientation and value. It is difficult to distinguish two iconic tags in different sizes, different orientations or different values depending on the conditions of a computer screen. Considering aesthetic reasons, icons are preferably designed in unified size for software applications. Limited choices of orientation and value also make it less possible to design large scale tag presentations.

By contrast, three visual variables—shape, colour and texture—are chosen to create the pre-icons of VDL. For topics tags, all tags under common viewpoint are first designed by uniform shape (pre-icon), and then those sorted into different topic categories will be added with another

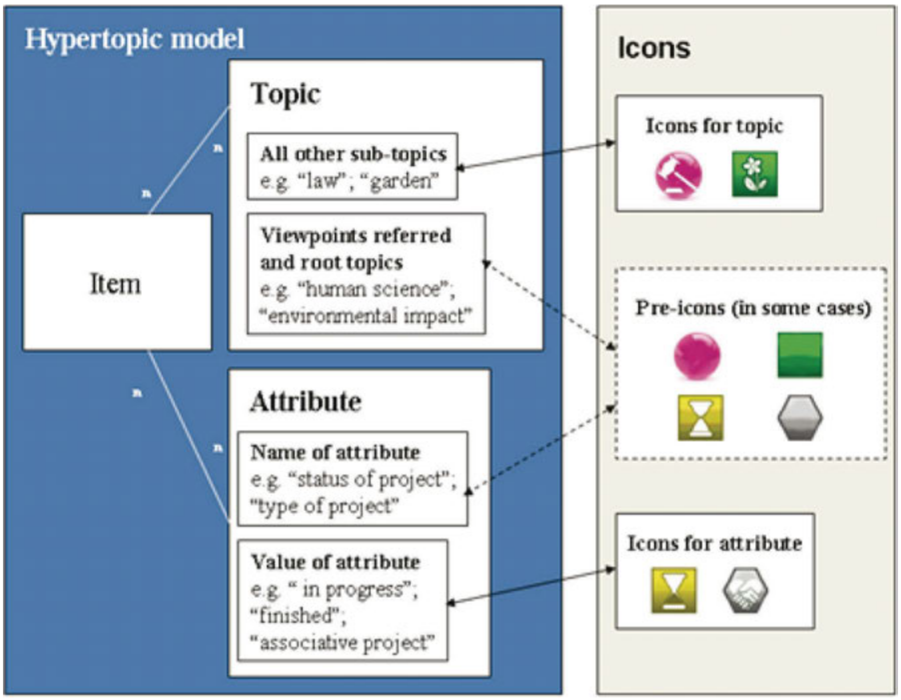


Figure 3. Iconized topics and attributes—two elements of Hypertopic—to form well-structured iconic tags (Ma and Cahier 2012)

visual variable colour to form updated pre-icons. Since topics tags are catalogued in tree structure, new visual variables would still have been added to create pre-icons for lower branches. However, on one hand the number of visual variables is limited; on the other hand excessive visual variables reduce the readability of iconic tag structures. To provide clearer and simpler VDL, iconic topic tags from the second level will always keep the same pre-icon without being distinguished by a new visual variable.

The graphical rule is similarly applied to attribute tags. Attribute name is directly iconized into coloured shapes (pre-icons) and then attribute value is detailed by joining a symbol onto it (except special cases as mentioned in the preceding paragraph, such as "language"). The chosen colours in Figure 6 show an example of the idea of visualizing tag categorization by graphical variables, yet without a strict colour choice test. However, colour and shape are supposed to interact in VDL, neither is the dominant variable. Considering colour-blind cases, the version in black and white is created as well. Variable colour will be replaced by texture (see Figure 6) while preserving all the other rules from the coloured version. The final version of both iconic topic tags and iconic attribute tags bears no visual difference unless specifically marked for their originality. However, pre-icons allow indicating those from the same category of viewpoint, topic or attribute name.

To evaluate how VDL-based iconic tags improve tag presentation in KOS's, we have done the first "tagging on paper" experiment in 2011 (Ma and Cahier 2012). Considered that tagging effectiveness is a complex subject associated with numerous user-related cognitive factors, this experiment focuses on whether VDL-based iconic tags help finding more usable tags to annotate knowledge in KOS by visual representation of tags and tag structures. Figure 5 shows three tested tags: textual tags, iconic tags without explicit structure and VDL-based iconic tags.

Across several tests in the experiment, early results demonstrated that VDL-based iconic tags have more advantages compared with iconic tags without structure and textual tags. Participants announced they easily located and located again later a tag from a tag presentation essentially through graphical tag structure and partly through iconic symbols. The knowledge resources tagged by VDL-based iconic tags were also supposed to be strongly connected in a KOS. The former test allowed us to confirm the first hypothesis on VDL-based iconic tags: visual codes of VDL improve knowledge tagging in a KOS. However there was no discussion of tag arrangement methods (all the tags in the experiment were arranged randomly).

Consequently, in this paper, we propose to produce a more in-depth study of VDL-based iconic tags. More precisely, we propose to:

– verify whether a conclusion on semantically structured textual tag clouds can be applied to VDL-based iconic tag presentations; and,
– develop a supplementary experiment to get more complete view on how to construct better VDL-based iconic tag presentations, which will be meaningful for creating iconic tag clouds in KOS's.

3.0 Semantic arrangement method for VDL-based iconic tags

Several arrangement methods are available for textual tag clouds, such as alphabetic arrangement, random arrangement, folksonomy-based arrangement, or semantic (linguistic-based) arrangement. While for iconic tags, only random arrangement and semantic arrangement are considered according to tag format. Since former studies demonstrated that semantically-clustered textual tag clouds yielded better tag presentation and interface, we are considering similarly semantically clustering VDL-based iconic tags. First we define what semantic arrangement refers to for VDL-based iconic tags.

Tag presentation in KOS's is dynamic wherein users choose recommended tags for thier own tagging and searching goals and in turn update useful tags for later use. Thus the tag arrangement should be convenient both for locating an existing tag and for adding new tags. VDL-based iconic tags improve the limits of textual tags in knowledge tagging. The symbolic characters of icons convey explicitly the represented objects and the graphical characters enhance connection among tags and documents tagged. In particular, a special group of icons called "pre-icons" function to signify the categories of tags: the same viewpoint, the same branch of topic or the same attribute name (tag structure proposed by Hypertopic). Pre-icons in VDL act as the common base of iconic tags. The tags in each category will be specified by combining symbols with this corresponding iconic base. Here we still think of the simplest case, the one-to-one tag-icon case as mentioned in section 2.3. However, this approach will also be extended for the knowledge tags to which no vocabulary (symbol) constraint of textual tags (icons) is applied. In that many-to-many case, more than one textual tag (icon) will be proposed for knowledge in each category.

The semantic relations within VDL-based tags are integrated from both graphical relations and semiotic relations of icons taking advantage of Visual Distinctive Language. Thus the semantic arrangement means iconic tags with the same pre-icons will be clustered. To arrange the tags in one category (one viewpoint, one branch of topic or one attribute name) requires only to put the tags with the same graphical characters together (same colour, same shape). Particular tags from different topic branches of

Figure 6. Semantic arrangement of VDL-based iconic tags (taking an iconic tag cloud for example)

the same viewpoint are displayed closer together (see Figure 6). It is hypothesized that this type of tag presentation will present clearer boundaries of tag clusters than those randomly arranged. Users might find and add tags easily even they do not understand completely icon representations. The semiotic interpretation of tag meaning will be concerned less since not only icon symbols but also pre-icons confirm the categorization of tags.

Because semantically structured textual tags have been studied before, we investigate only semantically structured VDL-based iconic tags and semantically structured iconic tags without explicit structure, taking comparison with those randomly arranged. Here iconic tags without explicit structure perform as a control group to see whether tag format or tag arrangement is more important for tag presentation. It is assumed that semantically structured VDL-based iconic tags will facilitate locating and relocating tags for knowledge tagging.

4.0 Experiment

A computerized experiment was conducted to investigate the tag arrangement for VDL-based iconic tags. There were four types of iconic tag presentations in this experiment (four groups A, B, C, D shown in Figure 7). Comparison took place in three sessions: group A and group B (to see whether semantic arrangement improves tag presentation for no visual structure iconic tags compared to random tags); group C and group D (to see whether semantic arrangement improves tag presentation for VDL-based iconic tags compared to random tags); group B and group C (to see whether semantic arrangement or VDL-based icons are more critical to improving tag presentation). Particularly, comparison between group A and group C has been made in a former study (random arrangement of two types of iconic tag clouds). Each group of participants was asked to tag 24 given documents (like a simulated KOS) by using the tags from tag presentations. We assume that users in group of VDL-based iconic tags and semantically arranged tags will find more appropriate tags (greater accuracy) in less time (speedier) compared to other patterns. In addition, we also traced participants' behaviors: the time spent to tag an item and its changing tendency, the frequency of asking for the instruction and the proportion among tags considered to choose and those finally being chosen.

4.1 Participants

Forty-eight French speaking students, 26 male and 22 female with computer science as their master major in the University of Technology of Troyes participated in this experiment. They were divided into four groups corresponding to four types of tested tag systems: group A for iconic tags without explicit structure and randomly arranged (12 persons); group B for iconic tags without explicit structure and semantically arranged (12 persons); group C for VDL-based iconic tags and randomly arranged (12 persons); group D for VDL-based iconic tags and semantically arranged (12 persons).

4.2 Material

The material for this online experiment included four types of tag presentations (see Figure 7), and 24 knowledge articles (see Figure 8). Tags in each presentation are knowledge tags referring to seven topical categories (from two viewpoints) and three attribute names on the topic of sustainable development. Tag presentation type one (type three) differs from type two (type four) on the tag arrangement while type one (type two) and type three (type four) differ on the tag format. We chose the same icon symbols for all four presentations to avoid the impact on semiotic interpretation (icon choosing). What we wanted to test was the influence produced by visual structure and arrangement of iconic tags. The twenty-four web articles were the same as those used in the first experiment. They were short texts with a large range of interest in the field of sustainability and each is represented by title, image and description.

4.3 Procedure

This experiment was composed of three parts: pre-questionnaire, tagging test and post-questionnaire. There

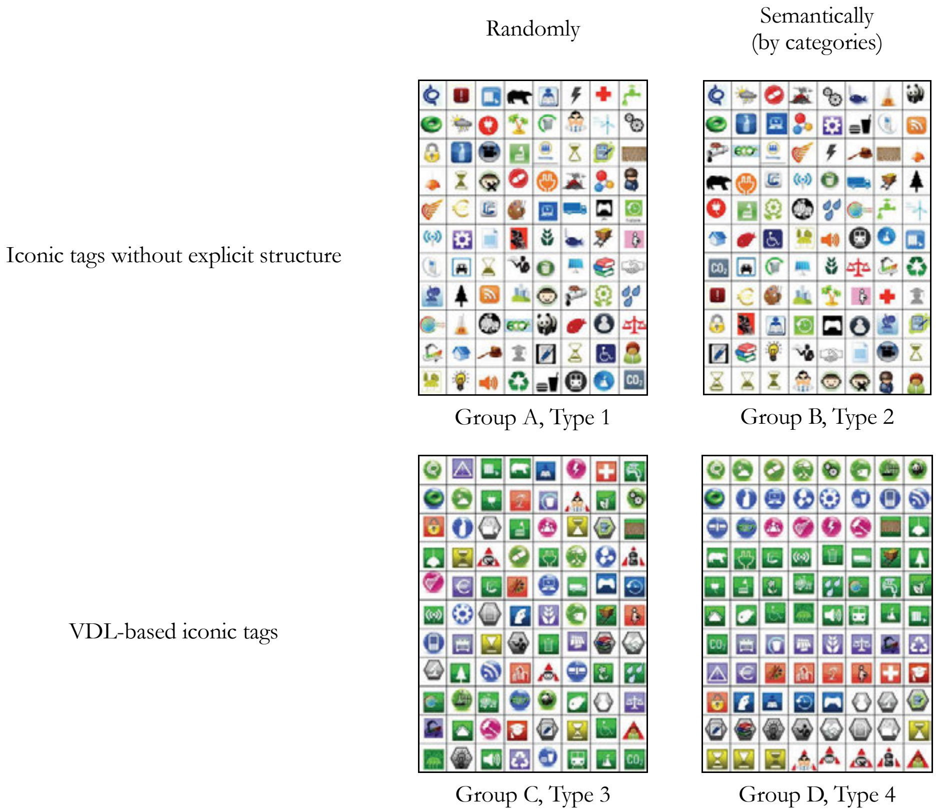


Figure 7. Four groups and their corresponding tag presentations in the online tagging test

Figure 8. Tagging test platform and articles to tag (example of item 1 for group B)

was no unified time constraint over the whole process but it proceeded without permit to suspend. All of the participants logged into the system with their e-mail addresses and assigned passwords. The system produced automatically for each of them a group code in order (A1, B1, C1, D1, A2, B2, C2, D2 ....). The letter of this code corresponded to the type of tag presentation they used. In order to understand the level of prior knowledge in the field of sustainable development, each participant first completed a pre-questionnaire of 10 questions: five concerned academic knowledge in the field while others were about personal understanding and awareness of sustainable development.

Once participants finished pre-questionnaires, they started tagging texts using given tags. A "Help" button was displayed in the upper right corner to give instruction if necessary. A double left click on an icon allowed submitting it into a tag-selection zone (choose an iconic tag) while a double left click on the icon in the selection zone was to return it to the former location of tag presentations; as well a simple right click on the icons made corresponding text of the icon visible. Participants could confirm tagging choices for an article and continue on to the next one by clicking the button "next item." Once a tagged text was confirmed, it could not be modified. Similarly, an untagged item could not be shifted up to the next one. When participants clicked the "finish the tagging" button on the final article, they arrived at the post-questionnaire to test tag structure identification using the four types of iconic tag presentations. They had the same operation of clicks as before to submit and cancel an icon. However, they could no longer get help from the textual meaning of iconic tags. The post-questionnaire was used in order to explore which type of iconic tag presentation explicates better semantic tag clusters.

Additionally, several new variables were also tested in this experiment. First, tagging duration was one of these variables. We were interested not only in average tagging duration for one item, but also in any changing tendency from the first article to the last one. Second, the proportion between chosen tags (tags selected) and final tags (tags confirmed for one item) was also meaningful. Here selected tags were placed in the tag selection zone while confirmed tags were those tags finally appearing in the tag selection zone when clicking "next item." This proportion could also be seen as the probability of confidence. The higher the average proportion was, the more participants were confident with their choice of tags. This percentage also implied the understanding level and learning result of iconic tags and their structure. Finally, asking for instruction revealed whether users had difficulty on operations in the test. This statistical record was considered as part of the prior knowledge.

4.4 Results

4.4.1 Prior knowledge test

Each question in the pre-questionnaire had one correct answer from three options (a, b or c). A participant who managed to find that answer won one point while a participant who could not find it did not earn any points. After the test, there was a list of points earned (10 in total) by each person. Participants whose point total was above or within the range from 6 to 2 were not considered in the final analysis. That is to say, they were excluded from the average level of the prior domain, which influenced the outcome of the experiment. Individual difference also was implied by the frequency of clicks on the "Help" button. Participants who asked more frequently for "Help" could show a worse understanding of the test. Levene's homocedasticity test[2] revealed no significant heterogeneity between the variances on the score in the pre-questionnaire (P=0.572) and instruction reading (P=0.812). The mean scores on the pre-questionnaire for the four groups were 8.5 for group A, 8 for group B, 8.4 for group C and 9 for group D. An ANOVA conducted on the subjects' performances in the pre-questionnaire revealed no significant difference (F<1). As far as the instruction reading was concerned, the mean times were 2 for group A, 1.7 for group B, 1.7 for group C and 2.2 for group D. The performances of the subjects revealed also no significant difference (F<1). The two results suggested that there was no significant individual difference on the prior knowledge test which could influence the later tagging test.

4.4.2 Tagging process

Here we must first explain the method of evaluating the tagging process that was applied in the former experiment. Two factors were considered in the evaluation: tagging quality (more appropriate tags found) and tagging speed (less time spent to tag). The method for analyzing the quality of tagging remained the same as in the previous experiment (Ma and Cahier 2102) using an expert matrix and Rx^2 criterion, which will be explained below. Eighty-seven tags each had a unique tag number from 1 to 87. Five experts on sustainability were invited to tag the texts with these 87 tags. For each text, they were required to rank all of the chosen tags with a number from 0 to 5 to represent the degree of correlation. Five indicated that the tag was certainly relevant to the item while 0 meant not relevant. The average of the five experts comprised a matrix, called the expert matrix showing the correlations between tags and items (see Table 1).

Similarly, the tagging result of all the participants filled 48 participant matrixes. The unique difference from the expert matrix was that the participant matrixes were filled

	Item 1	*Item 2*	*Item j*	…	*Item 24*
Tag 1	0	0	1	…	0
Tag 2	3	4	2	…	1
Tag i	2	2	1	…	0
…	…	…	…	…	…
Tag 87	5	4	5	…	3

	Item 1	*Item 2*	*Item j*	…	*Item 24*
Tag 1	0	0	1	…	0
Tag 2	1	1	0	…	1
Tag i	0	1	1	…	0
…	…	…	…	…	…
Tag 87	1	0	1	…	1

Table 1. Expert matrix (on the left) and participant matrix (on the right) to evaluate tagging quality (Ma and Cahier 2012)

merely with either 1 or 0: 1 refers to the tags used while 0 to the tags not marked in boxes.

To analyze the tagging result of participant x, the formula below was applied.

$$Rx = \sum_{i=1}^{87} \sum_{j=1}^{24} TE_{ij} \bullet TPx_{ij} \qquad (1)$$

TE_{ij} : number in row i column j of the expert matrix
TPx_{ij} : number in row i column j of the participant matrix (participant x)

Rx is a variable implying the tagging quality which refers to the degree of appropriate tags that have been chosen. It reveals high quality of tag cloud interactions such as locating and locating again useful tags considered relevant by experts. All the RXs in one group were considered as a one-dimensional table to perform an ANOVA analysis among groups.

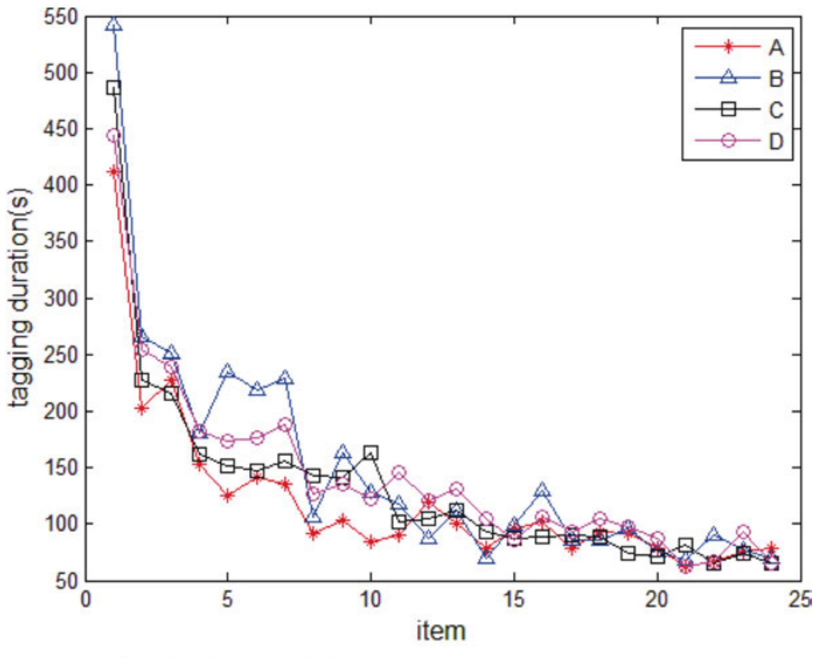
Tagging speed was originally reflected by the duration of tagging, from the selection of the first tag for the first text to the ending of the final tag for the final text. The final statistical results compared were Rx/tagging duration of each participant, representing tagging quality in per unit time. Levens's homocedasticity test indicated significant heterogeneity between the variances on the tagging process: Rx/tagging time, $P<0.05$. Consequently, these performances were analyzed using a nonparametric Kruskal-Wallis test. This latter test implied a significant effect of the semantically structured VDL-based icons on subjects' tagging performances, N=40, $P<0.05$. A more thorough analysis using a Mann-Whitney test indicated a significant difference between group D (M=342.1) and group C (M=238.2), Mann-Whitney U=32, P= 0.04. Similarly, the performances of group D were significantly better than group B (M=215.2), Mann-Whitney U = 5, P <0.05. As demonstrated before, group A (M= 154.4) was significantly poorer than group C, Mann-Whitney U = 15, P<0.05. In contrast, the performances obtained for group A and B did not differ significantly for the tagging process, Mann-Whitney U = 32, P=0.173.

4.4.3 Time changing tendency

Apart from average tagging time, dynamic change tendency is also useful for analysing user behaviour. It can be seen from Figure 9 that users in four groups revealed close changing tendencies. Tagging duration decreased from item 1 to item 24 in all the groups without significant difference emerging on the rate of change.

4.4.4 Post-questionnaire

The critical prediction of structural identification of tags was to compare the categories proposed by participants with predefined VDL categories (seven categories of topics and three categories of attribute names, the same as before). Participants who were in complete correspondence with one of these categories earned 2 points. Those whose category was partially correspondent were scored 1



duration for one article (four groups)

Figure 9. Average tagging duration for one article (four groups)

point. No points were awarded to participants who mixed more than one proposed category. From the name of the suggested category we also knew whether they only identified the visual structure of the tags by the graphical regularity of VDL or whether they understood the meaning of the tag and confirmed it by the graphical regularity. It was also assumed that the group working with VDL-based iconic tags and presented by category could offer more categories corresponding to the categorization of tags, but maybe there would be the risk that two provisions of VDL-based iconic tags demonstrated the same capacity.

After confirming that the homocedasticity of the variances was not statistically significant ($P<0.05$), Kruskal-Wallis test revealed a significant difference among four groups, $H=40$, $P<0.05$. More precisely, group D ($M=12.4$) performed significantly better than group B ($M=1.6$), Mann-Whitney $U=8$, $P=0.001$ and group C ($M=3.2$) Mann-Whitney $U=12$, $P=0.004$. As was observed in the former experiment, Rx of group C was significantly higher than that of group A ($M=0.6$), Mann-Whitney $U=26.5$, $P=0.037$. In contrast, group B did not obviously improve compared to group A, Mann-Whitney $U=44$, $P=0.465$.

4.4.5 Selection proportion

Levene's test implied significant differences between variances in four groups ($P=0.025$). The latter Kruskal-Wallis test revealed no significant difference on selection proportion among the four groups ($P=0.149$).

4.5 Discussion

The results are partially in accordance with out predictions. Semantically structured VDL-based iconic tag presentation showed better effectiveness in the tagging process (considering tagging quality and tagging speed) than the other three types.

4.5.1 Group C vs. group D to see whether semantic arrangement improves interaction of tag clouds for VDL-based iconic tags compared to random arrangement.

As demonstrated with textual tag clouds, semantically structured tag clusters led to a quicker and more accurate localizing of specific tags. Similarly, semantically structured VDL-based iconic tags also revealed better guidance in tag selection. Compared between tag presentations type 3 and type 4, semantically structured tags showed more clearly the layouts of tag clusters using visual signals, such as different colours or different shapes. Instead of spending time to identify VDL in group C, testers in group D got rapid graphical information about tag structure. Users' comments implied some of evidence. The participants in group D said that as soon as they saw tag presentations, they found clear icon categories represented in several graphical bases in common. In contrast, those in group C, although identifying the visual structure of tags, took much more time than the semantically structured group to catch this implicit information. The significantly better performance on structure identification in the post-questionnaire also validated this.

The advantage of semantically-arranged, VDL-based icons was demonstrated also in tagging topic-related articles. Users are likely to tag them with the same tags or at least with the tags in one category. For example, if they tagged a text on environment with a green tag, this tag or other green tags was supposed to be used again for another environmentally-concerned text. In the case of randomly arranged VDL-based iconic tag presentations, users knew that there were still other choices of green tags in the display. However, these green tags again took time and risked omitting some that were not used before. Otherwise semantically arranged VDL-based iconic tags might avert this problem. Green tags means all of the green tags were always listed together. Once one tag in a category was found, all other tags in that category appeared one by one. Using this not only saves time localizing a tag, but also increases the tagging quality because all the alternatives are listed together, with the same structure information implied by visual code, influencing users' selection accuracy and confidence.

Similar to the explication in the previous experiment, users became accustomed to selecting the tags from each visual category. Finding and choosing a tag from 88 options turns out to be a choice from seven small groups. In semantically arranged VDL-based iconic tag presentations, this method was better applied. Most of the testers in group D stated that they started the tagging process by consulting all the tag categories in every visual base, and then they preferred to locate at each visual category to select the useful tags. In group C, although they said they tried as well to choose tags from each visual category, it was not easy to find all icons in one category since they were scattered in the presentation. They always forgot which tag in this category had been browsed. When they decided to look back for a second ttime at a certain tag, they could not easily pick it out.

4.5.2 Group A vs Group B (to see whether semantic arrangement improves interaction of tag clouds for no visual structure iconic tags compared to random one).

However, semantically-structured, iconic tags without explicit structure did not reveal significantly better performance on the tagging process compared with randomly arranged groups, nor did they in the post-questionnaire. Testers in group A and B earned almost the same score in the identification of tag structure. The semantically structured arrangement did not bring a supplementary effect. As declared in semantically structured tag clouds (Schrammel et al. 2009), the semantic arrangement must be good enough otherwise users will not be able to distinguish it from random layouts and semantic layouts. Therefore it should only be used when the quality of the arrangement can be assured. Iconic tags without explicit structure did offer graphical interpretation of tags, yet they did not provide visual information on tag structure—semantic relations within them. Consequently, users used semantically arranged icons totally as they did randomly arranged icons, which was previously shown to be poorer than randomly arranged VDL-based icons (Ma and Cahier 2012) in the tagging process.

4.5.3 Group B vs Group C (to see whether semantic arrangement or VDL-based icons is more critical to improves interaction of tag clouds).

Seen from the assessment results, semantically-arranged tags improved the tagging process with the condition that the semantic structure was solid and clear enough for all users, as was demonstrated in group C and D. If not, it will act just like randomly arranged tags, like A and B. How to define a solid and clear semantic structure or said semantic layout among a group of tags is a crucial topic to discuss. On the one hand, if tags are in text or in icons without explicit structure, they have to be in such high accordance with daily comprehension that users easily recognize the tag cluster, using less ambiguous words. On the other hand, if tags can be sorted into several layers, they have to add complementary information for specifying their structure, such as VDL and pre-icons. Meanwhile, this information saves the users' time identifying semantic layers because of a more precise and intuitional tag structure. What's more, testers in group C did better than those in group B, which also leads to an interesting argument. It is assumed that in tag presentations tag format (representation of a single tag and its structure) is more essential than tag arrangement. Comparing group B with group C, one changes tag arrangement to semantically structured based on group A, while the other alters tag format by adding pre-icons to original icons in group A. However, the statistical results implied significant improvement between A and C (Ma and Cahier 2012) but not between A and B. In the absence of visual structure tags, even though tags are semantically arranged, they will not ameliorate the tagging process. As a result, reforming tag presentation requires first making better representations of tag and tag structure, and then implementing the arrangement. All of these observations on tag format and tag arrangement are meaningful to create visual tag clouds in a KOS.

4.5.4 Other events

Average tagging time of group C and D was longer than that of group A and B. It may be assumed that testers with VDL-based icons could find more appropriate iconic tags and they spent more time to select them considering pure tagging duration. In particular, testers using iconic tags without explicit structure merely selected limited icons because it was difficult to find more interesting tags among a huge number. Even though group D took a little longer tagging time, it still showed a significantly better tagging quality in per unit time, which signified that the pure tagging quality of group D was much more higher than other groups, including group C. Tagging duration decreased from article 1 to article 24 in all four groups, which signified progressive user learning on tags and getting skilled on tagging activity. It is assumed that the participant could learn gradually the sense of tags and their structure, and this could reduce tagging time. Meanwhile, the calculation of change tendency enabled us to argue that no matter which type of iconic tag was used, users showed similar changing regularity.

In particular, there was no significant difference on the proportion between final tags and selected tags (as defined above). This proportion did not make any significant difference among the four groups, which could be partly illustrated by the argument in the former experiment (Ma and Cahier 2012) that both types of iconic tags had equivalent capacity in tag interpreting and memorizing. From the present experiment, we can enhance the argument by another explanation that two types of arrangement, randomly and categorically, did not influence tag interpreting and memorizing. Users are supposed to have a close degree of confidence due to comprehension of tag representations. In other words, neither VDL nor semantically arranged structure will improve the comprehension and memorizing of tags except for the symbols of iconic tags.

5.0 Conclusion

The research findings in this paper have validated semantic arrangement of VDL-based icon tags providing better tag presentations for knowledge tagging. This advantage was mainly produced by visual representation of semantic tag structures by pre-icons. The observation is relatively consistent with that of semantically clustered textual tags. It is seen once again that the semantic arrangement must be good enough otherwise users will not be able to distinguish it from random layouts. What's more, results demonstrated that a tag format such as VDL-based is more critical compared to tag arrangement for knowledge tag presentations. This provides a possible interface for a KOS, such as tag clouds, to make a visual bridge between tags and knowledge. Well-structured tag clouds need to be built up by VDL-based iconic tags and arranged by semantic clusters based on empirical observations. Meanwhile, the explicit structure of tags will also help users in better understanding and identifying the organization of knowledge.

Notes

1. Levene's homocedasticity test is an inferential statistic used to assess the equality of variances for a variable calculated for two or more groups.
2. Variable predefined to analyze tagging effectiveness among four groups. Details can be seen in the previous paper (Ma and Cahier 2012).

References

Ames, Morgan and Naaman, Mor. 2007. Why we tag: The motivations for annotation in mobile and online media, In *Proceeding CHI '07 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York: ACM Press, pp. 971-80.

Angeletou, Sofia, Sabou, Marta, Specia, Lucia and Motta, Enrico. 2007. Bridging the gap between folksonomies and the semantic web: An experience report. In *The 4th European Semantic Web Conference 2007 (ESWC 2007), 3-7 Jun 2007, Innsbruck, Austria.* Available http://oro.open.ac.uk/23608/1/semnet2007.pdf

Bateman, Ivan, Lee, Kyung-il and Kim, Kono. 2008. TopicRank: Bringing insight to users. In *Proceeding SIGIR '08 Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval.* New York: ACM Press, pp 703-4.

Begelman, Grigory, Keller, Philipp and Smadja, Frank. 2006. Automated tag clustering: Improving search and exploration in the tag space. Paper presented at *Collaborative Web Tagging Workshop of 15th International World Wide Web, 23-26 May, Edinburgh, Scotland.* Available http://www.ra.ethz.ch/cdstore/www2006/www.rawsugar.com/www2006/20.pdf

Bénel, Aurélien, Zhou, Chao and Cahier, Jean-Pierre. 2009. Beyond web 2.0 ... and beyond the semantic web. In *Design of Cooperative Systems*, Chapter 1, Springer.

Berlocher, Ivan, Lee, Kyung-Il, and Kim, Kono. 2008. TopicRank: bringing insight to users. In *Proceedings of SIGIR 2008*, ACM Press, pp. 703-04.

Berners-Lee, Tim. 2000. Semantic web on XML. In *Proceedings of XML 2000*, Washington, DC.

Bertin, Jarques. 1983. Semiology of graphics: diagrams, networks, maps. W.J. Berg trans. University of Wisconsin Press, Translation.

Bielenberg, Kai and Zacher, Marc. 2005. *Groups in social software: Utilizing tagging to integrate individual contexts for social navigation*. Masters thesis, Bremen: University of Bremen, 2005.

Bowker, Geoffrey C. and Star, Susan L. 1999. *Sorting things out: classification and its consequences*. Cambridge, MA: MIT Press.

Brooks, Christopher H. and Montanez, Nancy. 2006. Improved annotation of the blogosphere via autotagging and hierarchical clustering. In *Proceedings of the 15th International Conference on World Wide Web (Edinburgh, Scotland, May 23 - 26, 2006). WWW '06*. New York: ACM Press, pp. 625-32.

Cattuto, Ciro, Loreto, Vittorio and Pietronero, Luciano. 2007. Semiotic dynamics and collaborative tagging. *Proceedings of the National Academy of Sciences of the United States of America 104*: 1461-4.

Chen, Ya-Xi, Santamaría, Rodrigo, Butz, Anderas and Therón, Roberto. 2009. Tagclusters: Semantic aggregation of collaborative tags beyond tagclouds. *Smart graphics lecture notes in computer science* 5531: 56-67.

Downey, Douglas, Dumais, Susan, Liebling, Dan. and Horvitz, Eric. 2008. Understanding the relationship between searchers' queries and information goals. In Shanahan, James G., Amer-Yahia, Sihem, Manolescu, Ioana, Zhang, Yi, Evans, David A., Kolcz, Aleksander, Choi, Key-Sun and Chowdhury, Abdur, eds., *Proceedings of the 17th ACM Conference on Information and Knowledge Management - CIKM 2008*. New York: ACM Press, pp. 449-58.

Echarte, Francisco, Astrain, José Javier, Córdoba, Alberto and Villadangos, Jesús. 2007. Ontology of folksonomy: A new modeling method. In Handschuh, S., Collier, N., Groza, T., Dieng R., Sintek M., and de Waard A., eds., *Proceedings of Semantic Authoring, Annotation and Knowledge Markup Workshop (SAAKM2007)*, 28-31 October, Whistler, British Columbia, Canada: CEUR. Available http://ceur-ws.org/Vol-289/p08.pdf.

Fu, Wai-Tat, Kannampallil, Thomas, Kang, Ruogu and He, Jibo. 2010. Semantic imitation in social tagging. *Journal of ACM transaction on computer-human interaction 17*: 1-37.

Fujimura, Ko, Fujimura, Shigero, Matsubayashi, Tatsushi, Yamada, Takeshi and Okuda, Hidenori. 2008. Topigraphy: Visualization for large scale tag clouds. In *Proceeding WWW '08 Proceedings of the 17th international conference on World Wide Web*. New York: ACM Press, pp. 1087-8.

Golder, Scott A. and Huberman, Bernardo A. 2006. Usage patterns of collaborative tagging systems. *Journal of information science 32*: 198-208.

Halvey, Martin J. and Keane, Mark T. 2007. An assessment of tag representation techniques. In Williamson, Carrie and Zurko, Mary Ellen, eds., *WWW 2007: proceedings of the 16th International Conference on World Wide Web, Banff, (Alberta, Canada), May 8-12, 2007*. New York: ACM Press, pp. 1313-4.

Hassan-Montero, Yusef and Herrero-Solana, Víctor. 2006. Improving tagclouds as visual information retrieval interfaces. In *I International Conference on Multidisciplinary Information Sciences and Technologies, InSciT2006. Mérida, Spain. October 25-28, 2006*. Available http://nosolousabilidad.com/hassan/improving_tagclouds.pdf

Hearst, Marti A. and Rosner, Daniela. 2008. Tag clouds: Data analysis tool or social signaller. In *HICSS '08 Proceedings of the International Conference in Waikoloa, Big Island, Hawaii, USA, 2008*. Washington DC: IEEE Computer Society, pp. 160-9.

Heymann, Paul and Garcia-Molina, Hector. 2006. *Collaborative creation of communal hierarchical taxonomies in social tagging systems. InfoLab Technical Report 10*. Available http://ilpubs.stanford.edu:8090/775/1/2006-10.pdf.

Hodge, Gail. 2000. *Systems of knowledge organization for digital libraries: Beyond traditional authority files*. Washington DC: The Council on Library and Information Resources. Available http://www.clir.org/pubs/reports/pub91/pub91.pdf

Kerr, Bernard. 2006. TagOrbitals: A tag index visualization. In *SIGGRAPH 2006 sketches*. New York: ACM, p.158.

Kim, Hak Lae, Scerri, Simon, Breslin, John G, Decker, Stefan and Kim, Hong Gee. 2008. The state of the art in tag ontologies: A semantic model for tagging and folksonomies. In Greenberg, Jane and Klas, Wolfgang, eds., *Metadata for semantic and social applications proceedings of the international conference on Dublin Core and Metadata Applications, Berlin, 22 - 26 September 2008, DC 2008: Berlin, Germany*. Göttingen: Universitatsverlag Göttingen, pp. 128-37.

Kipp, Margaret E.I. and Campbell, D. Grant. 2010. Searching with tags: Do tags help users find things? *Knowledge organization* 37: 239-55.

Kipp, Margaret E.I. and Joo, Soohyung. 2010. Application of structural equation modelling in exploring tag patterns: A pilot study. In Proceedings *Annual Meeting of the American Society for Information Science and Technology, Pittsburgh, Pennsylvania, USA*. Available http://www.asis.org/asist2010/proceedings/proceedings/ASIST_AM10/submissions/325_Final_Submission.pdf

Knautz, Kathrin, Soubusta, Simone and Stock, Wolfgang G. 2010. Tag clusters as information retrieval interfaces. In Sprague, Ralph H., ed., *Proceedings of the 43rd Annual Hawaii International Conference on System Sciences: 5-8 January, 2010, Koloa, Kauai, Hawaii*. Los Alamitos, Calif.: IEEE Computer Society Press, pp. 1-10.

Ma, Xiaoyue and Cahier, Jean-Pierre. 2011. Iconic categorization with knowledge-based icon systems can improve collaborative KM. In *Proceedings of CTS2011 in Philadelphia, USA, 2011, IEEE Conference Publications*, pp. 216-23.

Ma, Xiaoyue and Cahier, Jean-Pierre. 2012. Visual distinctive language: Using a hypertopic-based iconic tagging system for knowledge sharing. In Drira, Khalil and Reddy, Sumitra Mitra, eds., *Proceedings: 21st IEEE International WETICE Conference: WETICE 2012: 25-27 June 2012, Toulouse, France*. Los Alamitos, Calif.: IEEE Computer Society, pp. 456-61.

Macgregor, Geroge and McCulloch, Emma. 2006. Collaborative tagging as a knowledge organization and resource discovery tool. *Library review 55*: 291-300.

Mas, Sabine and Marleau, Yves. 2009. Proposition of a faceted classification model to support corporate information organization and digital records management. In *42nd Hawaii International Conference on System Sciences (HICSS). 5-9 January, 2009, Waikoloa, Big Island, Hawaii*. DOI= http://doi.ieeecomputersociety.org/10.1109/HICSS.2009.874.

Mika, Peter. 2005. Ontologies are us: A unified model of social networks and semantics. In Gil, Yolanda, ed., *The Semantic Web- ISWC 2005: 4th International Semantic Web Conference, ISWC 2005, Galway, Ireland, November 6-10, 2005: proceedings*. Berlin: Springer, pp. 522-36.

Nakamura, Carlos and Zeng-Treitler, Qing. 2012. A taxonomy of representation strategies in iconic communication. *International journal of human-computer studies* 70: 535-51.

Passant, Alexandre. 2007. Using ontologies to strengthen folksonomies and enrich information retrieval in weblogs. In *Weblogs and Social Media proceedings of the international conference in Boulder, Colorado, USA, 2007*. New York, ACM Press, pp. 128-37.

Paivio, Allan. 1971. *Imagery and verbal processes*, New York: Holt, Rinehart, and Winston.

Provost, James. 2008. *Improved document summarization and tag clouds via singular value decomposition*. Master thesis. Kingston, Canada: Wueen's University.

Rivadeneira, A.W., Gruen, Daniel M., Muller, Michae J. and Millen, David R. 2007. Getting our head in the clouds: Toward evaluation studies of tagclouds, In *CHI 2007 proceedings of the international conference in San Jose, California, USA, 2007*. New York: ACM Press, pp. 995-8.

Schrammel, Johann, Leitner, Michael and Tscheligi, Manfred. 2009. Semantically structured tag clouds: An empirical evaluation of clustered presentation approaches. In *CHI 2009 proceedings of the international conference in Boston, USA, 2009*. New York: ACM Press, pp. 2037-40.

Schwarzkopf, Eric, Heckmann, Dominik, Dengler, Dietmar and Kröner, Alexander. 2007. Mining the structure of tag spaces for user modeling. In *Data mining for user modeling ICUM'07 proceedings of the international conference in Corfu, Greece, 2007*, pp. 30-1.

Sen, Shilad, Lam, Shyong K., Rashid, Al Mamunur, Cosley, Dan, Frankowski, Dan, Osterhouse, Jeremy, Harper, Maxwell F. and Riedl, John. 2006. Tagging, communities, vocabulary, evaluation. In *Conference on Computer Supported Cooperative Work: 20th anniversary: November 4-8, 2006, the Fairmont Banff Springs Hotel, Banff, Alberta, Canada: conference proceedings*. New York: ACM Press, pp. 181-90.

Shaw, Blake. 2008. *Utilizing folksonomy: Similarity metadata from the del.icio.us system*. Available http://www.metablake.com/webfolk/web-project.pdf

Shepitsen, Andriy, Gemmell, Jonathan, Mobasher, Bamshad and Burke, Robin. 2008. Personalized recommendation in social tagging systems using hierarchical clustering. In *Recommender systems proceedings of the international conference in Lausanne, Switzerland, 2008*. New York: ACM Press, pp. 259-66.

Shiri, Ali Asghar and Revie, Crawford. 2005. Usability and user perceptions of a thesaurus-enhanced search interface. *Journal of documentation 61*: 640-56.

Specia, Lucia and Motta, Enrico. 2007. Integrating folksonomies with the semantic web. In Franconi, E., Kifer, M., May, W., eds., *ESWC '07 Proceedings of the 4th European conference on The Semantic Web: Research and Applications,* pp. 624-39.

Wu, Harris, Zubair, Mohammad and Maly, Kurt. 2006a. Harvesting social knowledge from folksonomies. In *Hypertext and hypermedia proceedings of the international conference in Odense, Denmark, 2006*. New York: ACM Press, pp 111-4.

Wu, Xian, Zhang, Lei and Yu, Yong. 2006b. Exploring social annotations for the semantic web. In *15th World Wide Web proceedings of the international conference in Edinburgh, Scotland, 2006*. New York: ACM Press, pp. 417-26.

Van Damme, Céline, Hepp, Martin, and Siorpaes, Katharina. 2007. Folksontology: An integrated approach for turning folksonomies into ontologies. In *Proceedings of the ESCW Workshop Bridging the gap between semantic web and web 2.0,* pp. 57-70.

Zhou, Chao, Lejeune, Christophe and Bénel, Aurélien. 2006. Towards a standard protocol for community driven organizations of knowledge. In *Proceedings of the 2006 conference on Leading the Web in Concurrent Engineering: Next Generation Concurrent Engineering.* Amsterdam: IOS Press, pp. 338-49.