Elements of social representation theory in collaborative tagging systems

Elementos da teoria da representação social em sistemas colaborativos de marcação

Patricia Zeni MARCHIORI[1]

Andre Luiz APPEL[2]

Eduardo Michellotti BETTONI[3]

Denise Fukumi TSUNODA[1]

Frank Coelho de ALCÂNTARA[4]

Abstract

This article discusses the information representation process based on the Moscovici's Social Representation Theory and domain analysis in Information Science. The aim was to identify mechanisms and constituent dimensions of social representation in collaborative tagging systems/social bookmarking systems. Scientific knowledge was defined as the object/phenomenon of representation in these systems; and the tag as the shareable structure of meaning that connects participants and resources. The empirical research involved descriptive statistical techniques applied to a *corpora* of tags available in CiteULike, which is a social tagging system developed for the academic community. The data analysis, performed in a sample of groups derived from the dataset, showed that the users' reuse of their own tags resembles the anchorage mechanism. The reuse of tags by other participants - in the same group - reveals some evidence of the objectification mechanism. Some speculation arose about the cognitive effort made by the individual, under group influence, with regard to the tagging activity, user's choice of resources, and sharing styles. Further studies on social bookmarking systems depend both on a "gain scale" of users and items tagged, requiring techniques and procedures redesigned by Information Science, Statistics, Network Analysis, Linguistics/Sociolinguistics and Social Psychology.

Keywords: Information representation. Information sharing styles. Social bookmarking systems. Social representation theory.

Resumo

O artigo discute um recorte na temática de representação da informação explorando a Teoria das Representações Sociais e a abordagem da análise de domínio da Ciência da Informação. Teve como objetivo geral identificar mecanismos e dimensões constituintes da representação social em grupos de participantes de sistemas de colaborativos de marcação (ou sistemas de marcação social). Definiu--se o conhecimento científico como objeto/fenômeno de representação em tais sistemas; delimitou-se a tag/marcação como unidade de registro e de significado; e se considerou os usuários, o conjunto total de tags, e os itens marcados como unidade de contexto. A pesquisa empírica envolveu técnicas de estatística descritiva aplicada à corpora de tags disponíveis em datasets do CiteULike - um sistema de marcação social orientado para a comunidade científico-acadêmica. A análise dos dados em uma amostra de grupos

derivada dos datasets permitiu verificar que o reúso das tags do próprio usuário se assemelha ao mecanismo da ancoragem; o reúso de tags por outros participantes do grupo revela indícios do mecanismo de objetivação. Especulou-se sobre as condições que afetam o esforço cognitivo na marcação de itens e em relação aos estilos de compartilhamento entre os elementos dos grupos. Estudos sobre sistemas de marcação dependem de um "ganho em escala" tanto de usuários como de itens marcados exigindo técnicas e procedimentos reelaborados entre a Ciência da Informação, a Estatística, a Teoria de Redes, a Linguística (e a Sociolinguística) e a Psicologia Social.

Palavras-chave: *Representação da informação. Estilos de compartilhamento. Sistemas de marcação social. Teoria das representações sociais.*

Introduction

Environments mediated by technology have led to the increasing autonomy of individuals in the process of information representation. Within this context, a set of applications - known as collaborative tagging systems or social bookmarking systems - aim to stimulate a "shared" effort to find and tag items in a joint collection of resources. The tags, which can be words, phrases, codes or other strings of characters, may either represent the features of the tagged resources (or resource-tagger relationships), as well as become representations or descriptions that can be used by search services, allowing people to find resources that are of interest to them at particular times (Furner, 2007).

The nature of this subject encourages studies on recommendation systems based on information resources selected by non-expert individuals, as well as research that explores modeling algorithms protocols related to disambiguation of words/terms used as tags (or the tagging activity itself). Word Sense Disambiguation (WSD) involves the association of a given word (in a text or discourse) with a definition or meaning/sense which is distinguishable from other meanings that could be potentially given to such a word. Disambiguation must involve the step of gathering all the different meanings for every word relevant to the text (or discourse under consideration) and assigning the appropriate word (which carries the intended meaning) to each occurrence and, therefore, excluding the non-significant ones (Ide & Véronis, 1998). Another approach defies the concept of "social" in these systems as it seeks to describe and explain the individuals' strategies when selecting the resources available on the Internet and their assumptions when choosing a term for this representation. Within this approach, it is assumed that conceptual statements - potentially epistemic and not always explicit - conceal a hidden set of dynamics that could be exposed and analyzed with the help of Psychology, Sociology, and Linguistics theories.

This study was based on Moscovici's Social Representation Theory (SRT) (1978, 2009). According to the SRT and Information Science, language and communication patterns are indicators of cooperation processes among individuals who share a given domain of knowledge, discipline or environment (Hjørland, 2002). By including SRT to the discussions related to social tagging systems, three premises have been established: 1) a given individual/user, who has his own reference framework, assumes a dynamic and dialectical relationship with the group she/he is involved with through the tagging activity; 2) by adopting a social software, an individual conveys elements and dimensions that shape the social representation of "knowledge as an object/phenomenon"; 3) the tag is a shareable structure of meaning among social software users.

The possibility of extracting implicit information from datasets (*corpora*) of tags, in which the actual "tagging effort" cannot be directly scrutinized, emerged with the purpose of identifying, in social bookmarking systems, how a community or scientific domain unveils the mechanisms and dimensions that constitute a (shared) social representation. The paper also intends to offer inputs to the field of multidisciplinary analysis, as well as in the visualization and analysis of large and complex network studies. Furthermore, the association of the socio-cognitive sciences and computational modeling, such as cognitive architecture and social simulation, can be explained from the SRT standpoint (Sun, 2006).

Searching for patterns in tags: Users, communities and sharing

According to Vander Wal (2005), the tagging activity can be seen as a "narrow folksonomy" that

reinforces one's Personal Information Management (PIM) allowing a certain individual to identify and classify an information resource using his/her own vocabulary. This kind of tagging action is dominant on Flickr, where a person or a few people apply a group of tags to retrieve one (or more) specific resource(s). On the other hand, even if a person involved in the tagging process uses terms derived from his/her own vocabulary in collaborative tagging (or "broad folksonomy"), there are more people tagging the same object/resource. A power curve (or a network effect) comes forth as a result to the number of persons involved in the tagging activity such as in Delicious. Therefore, collaborative tagging has required wider debates and a more significant amount of empirical studies concerning the possibilities of promoting access/resource discovery and knowledge organization (Vander Wal, 2005).

Irrespective of the nature of folksonomy, some inaccuracies such as typos, lack of plural/singular control, and the presence of lexical and grammatical variants are inherent to any tagging activity (Guy & Tonkin, 2006). In addition, a potential communal benefit arising from social tagging systems depends on a high level of accumulation and overlap of "units of interest" (users, information resources, or tags). Another challenge refers to the increasing flow of new resources on the Web and the [low] probability that the same resource is likely to be tagged by more users and that a significant amount will be found by others (Oh, 2008).

Additionally, it is argued that a bookmarking system receives the adjective "social" merely because the tagging activity is easily done by using a "social software", which is a term that simply means the asynchronous and collective distribution of [any] kind of knowledge (Boeije *et al.*, 2009). This is the antithesis of a previous statement from Golder and Huberman (2006) who claim that such systems actually stimulate associative movements among their users and help them establish groups.

With regard to the roles of information sharing undertaken by individuals, Talja (2002) divides the academic community into four groups - super-sharers, sharers, occasional sharers, and non-sharers - depending on the extent and intent in which participants engage in collective searching and information exchange activities. Although the original empirical data was collected from communities-of-practice, Talja (2002, p.4) identifies the following types of information sharing:

> a) Strategic: information sharing as a conscious strategy to maximize efficiency in a research group;
>
> b) Paradigmatic: information sharing as a means of establishing a novel and distinguishable research approach or area within a discipline or across disciplines;
>
> c) Directive: information sharing between professors and students; and
>
> d) Social: information sharing as a relationship- and community-building activity.

Talja (2002), and previously Haythornthwaite and Wellman (1998), address their conclusions in agreement with the critiques made by some designers who create technology-intensive information systems. According to those designers, individuals are seen as socially disembodied, i.e., by disregarding issues such as "[…] power, gender, socioeconomic status, differential resources, or complex bundles of interactions and alliances" (Haythornthwaite & Wellman, 1998, p.1102).

However, Furnas *et al.* (2006) take into account that a set of resources tagged by different people (with a particular tag in common) represents a collective image of these resources as they are understood by that community. This argument allows the connection to social theories.

Social Representation Theory: Dimensions and constituent processes

The dynamics of group exchanges - as in a social class or in a given culture - makes "familiar the unfamiliar", which allows consensus, creation of knowledge and, therefore, the construction of social representations (Moscovici, 1978).

Under the influence of those specific collective "choirs", or an unique universe of discourse, Moscovici (1978) identifies, for members of a given community, three dimensions that shape the concept of representation and provide content and meaning to what is represented. In addition, given the social characteristic of the process, these dimensions set "[...] social boundaries separating groups" (Santos, 1994, p.136,

our translation)[5], being defined as follows (Moscovici, 1978; Alves-Mazzotti, 1994, Santos, 1994):

a) *Information*: is related to the organization, quantity and quality of knowledge that a group has about an object;

b) *Field of representation*: refers to the idea of an image, a social model, and a concrete and limited body of propositions related to a particular aspect of what is being represented. It implies a hierarchical set of elements, formulated judgments, claims and some sort of arrangement; and

c) *Attitude*: exposes the overall orientation towards the object of social representation, usually on two opposite points (favorable, unfavorable), or even by intermediate positions between these extremes. It is a preconceived opinion rooted in group relationships, as well as the reorganization and reshuffling of the individual's experience concerning the object.

The term "object" requires further clarification. Although not all things can be included in the Theory, Marková (2006, p.202, our translation)[6] states that:

[…] any object or phenomenon, irrespective of being physical (a kitchen), interpersonal (friendship), mythological (the Loch Ness monster) or socio-political (democracy), can become an object of social representation […]. [The] Social Representation and Communication Theory considers any kind of representation. It peruses and builds theories about those social phenomena that have become, for no specific reason, a public concern. These phenomena, which are investigated and discussed, are those that ignite tension and trigger a reaction.

This investigation understands scientific knowledge - encapsulated as resources/items available on the Internet - as a social phenomenon and, therefore, as a latent object of social representation. When the representation process starts, the individuals' reference framework (their values and classification structures) is sustained by the social/group rules, both from an objective and subjective standpoint of the "object" (Moscovici, 1978). Such reference framework and group rules underpin the two fundamental mechanisms of SRT: anchoring and objectification.

Anchoring is:

[a] process that arises our curiosity and alters something troublesome and unfamiliar in our particular system of categories and fits it to a paradigm of a category that we consider appropriate […]. When a certain object or idea is compared with the paradigm of a category, it acquires characteristics from that category and it is readjusted to fit it […]. Anchoring is, therefore, to classify and name something (Moscovici, 2009, p.61, our translation)[7].

Anchoring happens at the private domain of comparisons, interpretations and categorizations, while objectification takes place in a given community/group by the transition of such concepts or ideas to schemes or to concrete images which - by the generality of their use and overall consensus - become would-be reflections of reality (Alves-Mazzotti, 1994).

Objectification has two essential movements: naturalization, which sets the imagined into the cognitive; and classification, which organizes and fixes in scope such stimuli and arranges them, preferably, to a pre-existing schema, i.e., into a socially defined framework. The classification conveys the unfamiliar to a familiar domain placing the object within a defined context, "[…] which means to add a label to those that are already in use, to broaden the existing class tree" (Moscovici, 1978, p.131, our translation)[8], or to assign or not (to the object) the characteristics of a given category.

According to Moscovici (2009), a figurative nucleus arises from those mechanisms, and it is assumed as being a structure of images that reproduces a

composite of ideas, which are revealed by the words that (often) express those ideas. The presence of the figurative nucleus strengthens the role of language in the SRT. In fact, there is a correspondence between the most frequently used words of a language and the core themes inferred from the figurative nucleus, which establishes a relationship between the language and the social representation.

As Moscovici recognizes the mediating force of language so do Talja *et al.* (2005). According to these authors, the constructionist metatheory evokes language and mediation driving components to contemporary studies on information retrieval and knowledge organization. The Constructionist Theory perceives the language as having a significant role in the social construction of "meaning" through the notions of discourse, utterances and vocabularies. Within this theory, the concept of cognition is replaced by conversations; and a conversation is recognized as a *sine qua non* condition for the constitution of the social world, knowledge, and identities (Talja *et al.*, 2005).

Methods

Understanding the tag as the "discourse" structure, it was defined as a unit of register (or an entity of meaning) related to "[...] a content segment considered as the basic unit to be categorized and counted" (Bardin, 2010, p.130, our translation)[9]. In order to organize and perceive the meaning of the unit of register Bardin (2010) establishes the concept of registration unit. In a collaborative tagging system, the context unit has three dimensions: the whole set/*corpora* of tags; the set of items tagged; and users/taggers.

The context units were obtained (free of charge) from datasets provided by a social tagging system aimed to promote and develop the sharing of scientific references among researchers <http://www.citeulike.org>. This database, covering the period from 2006 to March 2012, was processed in MS-SQL™ Server 2008 to exclude non-valid data and identify unique resources/items, tags and users.

This process of exclusion resulted in a countable set of 16,941,749 lines, each corresponding to an input tag per resource/item and per user. From those sets of lines, the ones that had the following contents were excluded: *no-tag*, **file-import%*, *imported%* and *bibtex-import*. Lines containing numbers were discarded keeping only the numbers '2' and '3' to avoid the exclusion of terms such as '2D' and '3D'. Any line containing a tag with less than two characters (a letter, a symbol) was also excluded. The final set, hereafter referred to as "research data", consisted of 14,895,884 lines in which 2,744,129 univocal items were identified, 717,928 were unique tags, and 72,097 were unique users.

Another step of adjustment of the research data helped to establish relations between identification (id) of the item posted/tagged with the unique code given to each user; the code that identifies the groups in which each user participates; and the tag(s) defined by the user(s) as a result of the tagging/posting activity.

The presence or absence of tags offers the possibility to apply a quantitative (statistical) approach and a measurement of weights. Thus, when considering the methods proposed by Bardin (2010), we chose to adapt the "relationship analysis technique". In this technique, the frequency of appearance of the tags (*registration units/entities of meaning*) is "[...] based on the principle that the higher the frequency of the elements, the greater their importance, [and] the co-occurrence (or non-co-occurrence) of two or more elements [reveals] an association or a dissociation process in the mind of the speaker" (Bardin, 2010, p.258, our translation)[10]. In the present research, the co-occurrence shall mean that a given tag is used by one (or more) different user(s) to categorize an item; equivalence indicates that one (or more) similar tag(s) is used by different users to name different items; and association means that different tag(s) is used by different users to identify a given item.

Based on the research data, the following groups of users were identified, as shown in Table 1.

More than half the users (69.11%) did not participate in any group and 2.77% of users belonged as

[9] *"[...] ao segmento de conteúdo a considerar como unidade de base, visando à categorização e a contagem frequencial".*

[10] *"[...] assenta no princípio de que quanto maior for a frequência dos elementos, maior será a sua importância* [e] *a coocorrência (ou a não coocorrência) de dois ou mais elementos* [revela] *a associação ou dissociação no espírito do locutor".*

sole individuals in their own groups. The data is in agreement with the study of Pfeiffer *et al.* (2008) who claim that, in any system shaped for the scientific-academic community, the tagging activity is effective only for private purposes. Roughly one third of the registered groups (31.77%) had about two to four members, which reinforces the research of Wheelan (2009) concerning the productivity of small groups.

Considering the proposal to identify the elements and dimensions of SRT in a "social space" (though virtual), the research data was reduced to select some groups for in-depth analysis. The groups containing users who participated in these groups only and who had articles/items tagged were maintained for further analysis. We chose this procedure to prevent possible bias that could occur if the user was associated with more groups and the same item (or tag) from being posted in distinct groups. As a result, one hundred groups were found, 1,117 unique users and 740,562 items. About 55% of these groups contained only one or two users.

When ordering the list of groups by number of users, it was found that the amount of users was recurrent from the sixth group on (eight users), which made it difficult to define consistent criteria for further cutbacks. Therefore, we chose to analyze only the six groups with a larger number of members. These groups were listed in descending order from the total of taglines of each user (Table 2).

Of these six groups, three - with more items and tags (G159TU21, G264TU15 and G238TU16) - were isolated and the following procedures were applied:

a) Verification, user to user, of reuse of own tags within the period of existence of the group (reuse was calculated by the frequency of use of tag(s));

b) Exclusion, user by user, of duplicate items;

c) Exclusion, user by user, of duplicate tags in order to generate the set of unique tags. No word sense disambiguation procedure was used;

d) Pointing out, user by user, the existence of equivalent tags;

Table 1. Range of users and respective groups: Total and percentage (2006-Mar/2012).

Range of users in the groups	Total number of users in the range	% of users in the range	Number of groups according the range of users	% of groups
Over 1000	1,813	2.51	1	0.02
Between 100 and 250	1,479	2.05	10	0.23
Between 51 and 99	1,832	2.54	29	0.66
Between 10 and 50	7,504	10.41	385	8.70
Between 5 and 9	3,877	5.38	596	13.47
Between 2 and 4	3,769	5.23	1,406	31.77
One user	1,999	2.77	1,999	45.16
Non-grouped users	49,824	69.11		
Total	72,097	100.00	4,426	100.00

Source: By authors (2012).

Table 2. Groups and amount of users: Organized by total of taglines (2006 - Mar/2012).

Groups	Amount of users who tagged items	Total of users' taglines	Total of items in each group
G159TU21	8	1,934	603
G264TU15	13	700	133
G238TU16	9	459	128
G68TU40	34	390	75
G208TU17	15	204	52
G80TU37	21	156	23

Source: By authors (2012).

P.Z. MARCHIORI *et al.*

e) Pointing out, group by group, the quantity of items and unique tags, the total of items per users (and their percentage regarding the total of items);

f) For shared tagging activity (involving more than one user), the quantity of items, the number of tags and unique tags, total of items per user (and their percentages regarding the total of items) were verified.

The aim of the analysis was to evaluate the reuse of tags by the same user, as an indicator of the anchoring mechanism, and reuse of tags by other users of the group as an indicator of the objectification mechanism. We sought for evidence of figurative nucleus and the following SRT dimensions: information, field of representation and attitude, assuming that these are the ones that provide content and meaning to what is represented.

Results and Discussion

Of the three groups selected for analysis, G159TU21 showed constant activity for six years and there were eight active members (Table 3). Since a code was automatically created to identify each user in the original dataset, the last four characters of this code were used to make reference to a given user.

Three of the eight members of the group contributed with 83.46% of total unique tags and 581 items (88.30%). Of the three, the user '1975' and the user 'a1b8' showed 82.00% of reuse. The percentage of reuse was calculated using the following formula:

$\%Reuse = (1 - \dfrac{TU}{TT}) * 100.00\%$. Several anomalies were found in the group's set of tags, such as the use of symbols (asterisk) [*diss; *rulewalker]; junction terms [cluseringgene], use of symbols for junction terms [analysis_tool; assembly-quality], use of plural/singular [classifier; classifiers], existence of misspellings [alignlent]; symbols indicating hierarchies [essemble; essemble_clustering; essemble_network] and functional tags [to_read]. The functional tags represent, according to Golder and Huberman (2006), an action intended or carried out by the individual who performs the tagging activity. This action can refer to the organization or the performance of a task, for example. The monitoring of the tagging activity showed that the users had not used affective tags - usually adjectives - as defined by Lu et al. (2010), which convey affective or judgmental utterances.

The co-occurrence of tags among items tagged in common was identified among six of the eight members of the group with the highest incidence among users who could be considered super-sharers. In fact, users '1975' and 'a1b8' tagged 24 items in common with 15 co-occurrences of tags (assembly; alignment; breakpoint; human). These two users also shared a common item that was linked by the co-occurrence of the tag human. Another pattern identified was a cluster between users 'ddf2' and '6a75', who shared four items, with the co-occurrence of two tags (eqtl; malaria) for one of the items. Users 'ddf2' and '9df0' shared an item using a tag which displays a plural/singular anomaly

Table 3. G159TU21: Users, tags and items tagged (2007-2012/Mar).

User	Tags			Items tagged	
	Total (TT)	Unique (TU)	% reuse	Total	%
1975	888	160	81.98	269	40.88
a1b8	479	88	81.63	176	26.75
ddf2	479	191	60.13	136	20.67
6a75	138	63	54.35	56	8.51
9df0	13	9	30.77	7	1.06
ef0a	11	3	72.73	10	1.52
4173	10	10	0.00	3	0.46
948e	2	2	0.00	1	0.15
Total	2,020[1]	526		658	100.00

Notes: [1]86 taglines have no content, therefore, totaling 1,934 lines.
Source: By authors (2012).

(*network*; and *networks*, respectively). The tag *network* was also used for another item/resource tagged by three users ('a1b8'; 'ddf2' and '1975'). This item was tagged again by another user in the same cluster (user '6a75') but using another tag. There were also three clusters by association in which two or more users tagged the same resource, but with different tags. In this case, user 'ddf2' was the only one who showed up in those clusters. Another cluster with three users ('a1b8', '6a75' and '1975') tagged one item in common. Users '6a75' and '1975' assigned the tag *comparisons* and *comparative*, respectively, which indicated a "variation of the word" anomaly.

Even if other users were less active, the movements of construction, communication and relationship among the individuals, identified this group as having a "social sharing style" (Talja, 2002).

The analysis of the G264TU15 revealed a distinguished characteristic, which was at first considered as a coincidence: the task-oriented search for Internet resources within a given period was divided among the thirteen participants (Table 4).

The dynamics of the tagging activity demonstrated that some rules were probably defined by the participants, except for two users. It followed a pattern of ten unique items tagged per participant. There was also evidence of the establishment of another pattern for the number of tags per user (total ratio of tags and

unique tags). In this case, however, it might be a coincidence. Even so, the search pattern and input of items on the system is not negligible. There was no evidence of the existence of super-sharer users and we observed that the whole group activity lasted two months.

No items were tagged by more than one user, even if a significant degree of equivalence of tags was identified, i.e., identical tags were used by different group members. In this particular group, the participants probably cut and pasted parts of text/title/abstract into the system's text box. These actions led to the appearance of "noise" in the set of tags such as definite/indefinite articles and connectives. The group seemed to have adopted a strategic sharing style, i.e., to increase or maximize the efficiency of a given task (Talja, 2002).

The G238TU16 group had nine members and one of the participants ('7e4e') was identified as a super-sharer, contributing with 80.97% of the tags in the group and 79.69% of the total items tagged. This user was the only one who showed consistent tagging activity throughout the group's existence (which lasted three years) (Table 5).

There were no items tagged in common among members of this group, but we found the equivalence of tags in a couple of cases: two tags were used by two users [*politics*] (or three, if the disambiguation of *policy* is

Table 4. G264TU15: Users, tags and items tagged (Nov/Dec 2011).

User	Tags			Items tagged	
	Total (TT)	Unique (TU)	% reuse	Total	%
7769	51	42	17.65	10	7.52
3dca	47	33	29.79	10	7.52
5609	41	39	4.88	10	7.52
9d52	59	48	18.64	10	7.52
32f5	62	41	33.87	10	7.52
50f3	53	51	3.77	11	8.27
7f48	110	56	49.09	12	9.02
ac19	65	41	36.92	10	7.52
b19e	43	35	18.60	10	7.52
7d97	34	31	8.82	10	7.52
72f7	50	46	8.00	10	7.52
56aa	42	38	9.52	10	7.52
1151	43	41	4.65	10	7.52
Total	700	542		133	100.00

Source: By authors (2012).

Table 5. G238TU16: Users, tags and items tagged (May/2009-Feb/2012).

User	Tags			Items tagged	
	Total (TT)	Unique (TU)	% reuse	Total	%
9e26	10	10	0.00	2	1.56
88c9	23	16	30.43	5	3.91
7f24	1	1	0.00	1	0.78
2e08	16	13	18.75	3	2.34
41de	13	12	7.69	2	1.56
729d	20	11	45.00	7	5.47
c7c4	15	8	46.67	5	3.91
d694	3	3	0.00	1	0.78
7e4e	430	67	84.42	102	79.69
Total	531	141		128	100.00

Source: By authors (2012).

accepted). The tag *feminism* was used by two users and two participants were connected by the terms (and junction terms) [*social_networks*; *socialnetworking*]. Three participants used terms with semantic closeness [*minority_youth*; *youth_cultures*; *[y]oung_people*; *youth*]. The presence of this super-sharer corresponds to the "directive style" as the social behavior of the group. This kind of behavior is prevalent during activities of professors and students (Talja, 2002).

Final Considerations

The results - when the cutbacks were defined - endorsed the research proposal. It was found that the appropriation of the social tagging system by individuals resulted in the tagging activity in the groups they belong. Through tagging resources available on the Internet, these individuals perform dynamic and dialectical relationships in which their frame of experience is reflected by the tag assignment, taken as a structure of meaning that is potentially shareable with other users. With regard to the elements and dimensions that shape the social representation of "knowledge as an object/ phenomenon", as defined in the scope of the investigation, the data suggests that the reuse of one's own tags resembles the anchoring mechanism of the SRT.

On the other hand, the extra SRT dimensions and mechanisms could not be supported by the results. The reuse of tags by participant(s) in another group can indicate the presence of the objectification mechanism. Some tag reuse was perceived, in fact, to occur among users themselves. But, when it occurred between users, it took place extensively and more explicitly only in subgroups involving two users/researchers. What could be identified as objectification, however, may also be the result of decreased cognitive effort, in which the ease of use of a previously submitted tag in the system does not require new attempts by another user.

Regarding the other dimensions of the Social Representation Theory, the defined dataset cutbacks allowed the analysis of some evidence, as follows.

About the SRT modeling dimension of information, the research data showed that the following elements helped to organize the information selected by the user: the actual record of items; the entry of tag(s) assigned to the item(s); and the submission of items and tags in the group in which one participates. The social representation component also happened in the groups analyzed, as a result of a given item that had been tagged by different users; different items tagged by users; and the reuse of one's own tags and/or tags from other users.

The modeling dimension of the field of representation is revealed within the context of the groups through the selection of concepts that comprehends the subject of interest shared by the participants and these concepts are expressed by the total set of tags placed in the group. However, no affective tags were identified in any of the groups' collection of

tags. Their existence could have helped to define the user's "judgment/opinion" on the item(s).

As for the dimension of attitude, the data analysis enabled to the inference about the choice of a given item despite a variety of others available on the Internet. This act of selection implies some degree of value by the user. *Pari passu,* a given item by being tagged by another member of the group could indicate a potentially collective guideline concerning the "object".

Another element of the SRT, the figurative nucleus - understood as the use of words that most often reflect the existence of complex consensual ideas within the group - can be recognized by the frequency that certain tags occur in the total set of tags collected by these groups.

However, the research data did not allow a complete verification on how individuals form the groups, i.e., if the individuals know each other in the physical environment; if they use the system just to facilitate/comply with some activity; if the individuals do not know each other in the real world but choose the system to share material of interest as a result of their activities; or if the participants join, by their own free will, a group created by third parties. Nonetheless, it was found that groups use the system in a variety of ways. This group behavior can be summarized as follows:

a) Groups in which the behavior of super-sharers do not necessarily influence the tagging activity of other users;

b) The existence of [a] super-sharer(s) affects the frequency count and, depending on the degree of reuse, the potential stability of the tags in the group as well, with consequences to the social bookmarking systems as a whole;

c) A larger number of participants - within a larger period of time - results in the reuse of items/resources and tags (improving equivalence/reuse and co-occurrence), and;

d) The system usage with the intention of performing a short task within a particular time frame shows some degree of prior organization and, probably, a common goal to be achieved.

With regard to the tags, users of social tagging systems, when performing the tagging activity, are free to define the word they consider representative to "tag" the content of the resource, as well as the quantity of tags they apply to such item, and how to write/input them into the system. A defined set of tags establishes a significant content description to a given resource that can be expanded both from strengthening the tags previously used (via reuse) and/or from another term/tag given to the same item by another user. As a result of this dynamic process, a given resource achieves a "gain scale" if more users retrieve it and choose to tag it in the system. Similarly, the same resource can be tagged again or the same tag can be used for other items and, in short, this scale effect would result in a repository of selected items.

The building of a 'critical mass of users', i.e., the increase in the number of participants in the system community, seems to be an obstacle for further and more conclusive studies focusing on social tagging. Interactions in those systems seems to occur at different levels (cultural, linguistic, knowledge and behavioral) whose boundaries are not easily defined and analyses at a single level tends to oversimplify others. Thus, further studies on the subject and/or those that consider large-scale multiuser systems would demand that domain analysis go beyond the actions of individuals in the real world and their epistemic communities. The interdisciplinary and multidisciplinary relation of Information Science will become even closer to Statistics, Network Theory, Linguistics (and Sociolinguistics), and Social Psychology, both in sharing and complementing the methodological procedures and techniques, as well as through synergic analysis.

References

Alves-Mazzotti, A.J. Representações sociais: aspectos teóricos e aplicações à educação. *Em Aberto*, v.14, n.61, p.60-78, 1994. Disponível em: <http://emaberto.inep.gov.br/index.php /emaberto/article/viewFile/912/818>. Acesso em: 21 maio 2013.

Bardin, L. *Análise de conteúdo*. 4.ed. Lisboa: Edições 70, 2010.

Boeije, R. *et al*. Knowledge workers and the realm of social tagging. In: Hawaii International Conference on System Sciences, 42., 2009, Waikoloa, Big Island, Hawaii. *Proceedings...* Washington, DC: IEEE Computer Society, 2009. p.1-10. doi: 10.1109/HICSS.2009.801

Furnas, G.W. *et al*. Why do tagging systems work? In: Conference for Human-Computer Interaction, 2006, Montreal. *Proceedings...* New York: ACM, 2006. p.36-39. doi: 10.1145/11 25451.11 25462

Furner, J. User tagging of library resources: Toward a framework for system evaluation. In: World Library and Information Congress; IFLA General Conference and Council, 73., 2007, Durban. *Proceedings...* Durban: IFLA, 2007. Available from: <http://archive.ifla.org/IV/ifla73/papers/157-Furner-en.pdf>. Cited: May 21, 2013.

Golder, S.A.; Huberman, B.A. Usage patterns of collaborative tagging systems. *Journal of Information Science*, v.32, n.2, p.198-208, 2006. doi: 10.1177/0165551506062337

Guy, M.; Tonkin, E. Folksonomies: Tidying up tags? *D-Lib Magazine*, v.12, n.1, 2006. doi: 10.1045/january2006-guy.

Haythornthwaite, C.; Wellman, B. Work, friendship, and media use for information exchange in a networked organization. *Journal of the American Society for Information Science*, v.49, n.12, p.1101-1114, 1998. doi: 10.1002/(SICI)1097-4571(1998) 49:12<1101::AID-ASI6>3.0.CO;2-Z

Hjørland, B. Domain analysis in information science: Eleven approaches, traditional as well as innovative. *Journal of Documentation*, v.58, n.4, p.422-462, 2002. doi: 10.1108/00 220410210431136

Ide, N.; Véronis, J. Word sense disambiguation: The state of the art. *Computational Linguistics*, v.24, n.1, p.1-41, 1998. Available from: <http://sites.univ-provence.fr/~veronis/pdf/1998wsd. pdf>. Cited: May 21, 2013.

Lu, C. *et al*. The topic-perspective model for social tagging systems. In: ACM Sigkdd International Conference on Knowledge Discovery and Data Mining, 16, 2010, New York. *Proceedings...* New York: ACM, 2010. p.683-692. doi: 10.1145/1 835804.1835891

Markova, I. *Dialogicidade e representações sociais*: as dinâmicas da mente. Petrópolis: Vozes, 2006.

Moscovici, S. *A representação social da psicanálise*. Jorge Zahar: Rio de Janeiro, 1978.

Moscovici, S. *Representações sociais*: investigações em psicologia social. Petrópolis: Vozes, 2009.

Oh, J.S. Shared interests expressed in a social bookmarking site. *Proceedings of the American Society for Information Science and Technology*, v.45, n.1, p.1-6, 2008. doi: 10.1002/meet. 2008.14504503133.

Pfeiffer, H.D. *et al*. Tagging as a communication device: Every tag cloud has a silver lining. *Proceedings of the American Society for Information Science and Technology*, v.45, n.1, p.1-5, 2008. doi: 10.1002/meet.2008. 1450450135.

Santos, F.S. Representação social e a relação indivíduo-sociedade. *Temas em Psicologia*, v.2, n.3, 1994. Disponível em: <http://pepsic.bvsalud.org/scielo.php?pid=S1413-389X1994000300013& script=sci_arttext>. Acesso em: 21 maio 2013.

Sun, R. (Ed.) *Cognition and multi-agent interaction*: From cognitive modeling to social simulation. New York: Cambridge University Press, 2006. Available from: <http://goo.gl/ HJWwG>. Cited: May 21, 2013.

Talja, S. Information sharing in academic communities: Types and levels of collaboration in information seeking and use. *New Review of Information Behavior Research*, v.3, p.143-160, 2002. doi: 10.1.1.96.163&rep=rep1&type=pdf

Talja, S.; Tuominen, K.; Savolainen, R. "Isms" in information science: Constructivism, collectivism and constructionism. *Journal of Documentation*, v.61, n.1, p.79-101, 2005. doi: 10.1108/00220410510578023>

Vander Wal, T. *Explaining and showing broad and narrow folksonomies*. 2005. Available from: <http://www.vanderwal. net/random/entrysel.php?blog=1635>. Cited: May 21, 2013.

Wheelan, S.A. Group size, group development, and group productivity. *Small Group Research*, v.40, n.2, p.247-262, 2009. doi: 10.1177/1046496408328703